Research

Neural coding of reward magnitude in the orbitofrontal cortex of the rat during a five-odor olfactory discrimination task

Esther van Duuren,^{1,2,5} Francisco A. Nieto Escámez,³ Ruud N.J.M.A. Joosten,¹ Rein Visser,¹ Antonius B. Mulder,⁴ and Cyriel M.A. Pennartz²

¹Netherlands Institute for Neuroscience, 1105 BA Amsterdam, The Netherlands; ²Swammerdam Institute for Life Sciences, University of Amsterdam, 1098 SM Amsterdam, The Netherlands; ³University of Almeria, 04120 Almeria, Spain; ⁴Department of Anatomy and Neuroscience, CNCR, VU University Medical Center, 1081 BT Amsterdam, The Netherlands

The orbitofrontal cortex (OBFc) has been suggested to code the motivational value of environmental stimuli and to use this information for the flexible guidance of goal-directed behavior. To examine whether information regarding reward prediction is quantitatively represented in the rat OBFc, neural activity was recorded during an olfactory discrimination “go”/“no-go” task in which five different odor stimuli were predictive for various amounts of reward or an aversive reinforcer. Neural correlates related to both actual and expected reward magnitude were observed. Responses related to reward expectation occurred during the execution of the behavioral response toward the reward site and within a waiting period prior to reinforcement delivery. About one-half of these neurons demonstrated differential firing toward the different reward sizes. These data provide new and strong evidence that reward expectancy, regardless of reward magnitude, is coded by neurons of the rat OBFc, and are indicative for representation of quantitative information concerning expected reward. Moreover, neural correlates of reward expectancy appear to be distributed across both motor and nonmotor phases of the task.

It has been noted for a long time that the magnitude of a primary reinforcer exerts a profound effect on the selection and speed of behavioral responses (Black 1968; Campbell and Seiden 1974; Brown and Bowman 1995; Boysen et al. 2001; Bohn et al. 2003). Likewise, in computational neuroscience, different algorithms for reinforcement learning (RL) consider reward magnitude an important parameter to be gauged and predicted during sensorimotor processing (Sutton and Barto 1981; Schultz et al. 1997). In one of these models, in which glutamate serves as a reinforcing signal guiding synaptic modifications necessary for adapting operant behavior, reward-related information is primarily processed by glutamatergic projection neurons of the orbitofrontal cortex (OBFc), basolateral amygdala, and related limbic areas (Pennartz 1997). The OBFc is known to be involved in the representation of the motivational significance of stimuli and in applying this information to the guidance of goal-directed behavior (Thorpe et al. 1983; Gallagher et al. 1999; Lipton et al. 1999; Schoenbaum et al. 1999, 2003; Yonemori et al. 2000; O’Doherty et al. 2003). Accordingly, rats with OBFc lesions exhibit impairments when stimulus–reward contingencies are altered during olfactory discrimination learning, whereas the initial acquisition of those associations remains unaffected (Ferry et al. 2000; Schoenbaum et al. 2002). Monkeys and humans with damage to the OBFc demonstrate similar deficits: both show impairments in reversal learning and decision-making (Dias et al. 1996; Meunier et al. 1997; Baxter et al. 2000; Manes et al. 2002; Hornak et al. 2004; Izquierdo et al. 2004).

Another consideration for a central role of the OBFc in RL is the neurophysiological evidence for the coding of predictive information regarding upcoming reinforcers. Neurons in primate OBFc process information concerning expected outcomes of be-

havioral responses, showing differential firing activity during the anticipation of various types or amounts of reinforcers (Tremblay and Schultz 1999; Hikosaka and Watanabe 2000; Wallis and Miller 2003; Roesch and Olson 2004; Ichihara-Takeda and Funahashi 2006; Padoa-Schioppa and Assad 2006). In addition, findings in rat OBFc suggest predictive neural coding of both appetitive and aversive outcomes (Schoenbaum et al. 1998). However, a difficulty in interpreting the latter findings is that “go” responses for the rewarding outcome were compared with “go” responses for an aversive outcome, the latter being a response that was erroneously made. Thus, the neural activity seen in the waiting period during these “false alarm” responses may not be related to the expectancy of a particular (aversive) outcome, but, considering OBFc functioning, for example, to a signal reflecting a conflict, error in responding or an internal requirement for cognitive flexibility that will result in adjustment of the behavioral response. A second motivation for this study on coding of reward magnitude was that many previous studies employed only two reward sizes. A limitation of such a design is that it is not possible to characterize the tuning relationship between reward size and neural activity. In the current five-odor olfactory discrimination “go/no-go” task, three odors were associated with a parametrically varied amount of reward, allowing the comparison of neural responsivity between trial types with highly similar behavior (i.e., all correct “go” responses). Thirdly, while previous studies in rats focused on reward representation in an immobile trial phase preceding reward delivery by a short delay, this study also paid attention to the possible occurrence of reward-expectancy correlates during the motor phase of the trial, i.e., when the animal actually performed a “go” response toward the reward site. When effects of the expected reward size could be uncovered during voluntary, operant action, this may help to understand the formation and representation of action–outcome associations (Balleine and Dickinson 1998; Dayan and Balleine 2002), an issue that has remained underexposed in orbitofron-

⁵Corresponding author.

E-mail evduuren@science.uva.nl; fax 31 20 5257709.

Article is online at <http://www.learnmem.org/cgi/doi/10.1101/lm.546207>.

tal cortex studies performed until now, including those in primates.

Results

Behavior

For the analysis, data were used from 24 recording sessions, obtained from seven rats. Animals needed on average ~17 trials for each of the different positively reinforced trial types to reach the criterion of 15 successful trials per reward size; there was no difference in performance level between the different amounts of reward (number of trials needed for 0.05 mL: 16.6 ± 0.4 ; 0.15 mL: 16.9 ± 0.5 ; 0.30 mL: 17.4 ± 0.6). Rarely the number of 15 trials per reward size was not reached during a session. For the nonreinforced and quinine condition, animals made on average 5.8 ± 0.7 and 4.5 ± 0.7 “go” responses, respectively (sucrose versus quinine or versus nonreinforced: $P = 0.000$; paired sampled *t*-test). In the course of a session these erroneous “go” responses (false alarms) were followed by trials in which animals, after sampling odors predictive for these outcomes, began to withhold responses toward or at the fluid well. Withholding was typically followed by odor pokes with durations shorter than the required 2 sec, thus yielding invalid trial types. This indicates that the animals were actually capable of associating an odor with either a positive or negative outcome. Since the minimal number of trials accepted for the analysis of the electrophysiological data was six, neural responses to the various stimuli applied during the nonrewarded trial type were not examined any further. However, sessions in which six or more trials for quinine were present ($n = 9$) were used to determine whether neural responses showing reward magnitude differences were compatible with response patterns for quinine (see below). This was done by comparing neural responses during the quinine trials with responses during the trial type with the same amount of sucrose (0.15 mL).

Examination of the movement time showed no significant differences between the positively reinforced trial types (0.05 mL: 1.43 ± 0.04 sec; 0.15 mL: 1.35 ± 0.03 sec; and 0.30 mL: 1.37 ± 0.04 sec). A comparison between these trial types and the nonreinforced and quinine trials revealed a significant difference: movement time during positively reinforced trials taken together (1.38 ± 0.03 sec; sample sizes per reward magnitude: 0.05 mL: $n = 336$; 0.15 mL: $n = 343$; 0.30 mL: $n = 356$) was significantly shorter compared to nonrewarded and quinine trial types (nonrewarded “go” responses: 1.50 ± 0.06 sec, $n = 138$; quinine “go” responses: 1.74 ± 0.13 sec, $n = 107$). However, when the entire sequence of odor sampling and moving to the fluid well was taken into consideration (overall response time), rats performed significantly faster for the 0.15 mL and 0.30 mL trial types (3.46 ± 0.06 sec and 3.51 ± 0.05 sec, respectively) as compared to the 0.05 mL sucrose trials (3.68 ± 0.06 sec).

Histology

Histological verification of the tetrode positions (Fig. 1) showed that the recording sites ranged between 2.7 mm and 4.2 mm anterior to bregma, and were limited to the ventral and lateral orbital regions of the OBFC. Recording depth ranged from ~3 mm to 5.5 mm below cortical surface (Paxinos and Watson 1996).

Electrophysiology

General overview

During the 24 recording sessions that were mainly performed on consecutive days, a total of 894 single units were recorded in the OBFC. The number of single units recorded per session ranged

Figure 1. Representative histological section showing the localization of tetrode recording sites. Black arrowheads in the upper part of the section indicate sites at which tetrodes have entered the brain. Several tracks are partially visible, including one track with an endpoint marked by a lesion (*). Recordings in all rats were localized in the ventral and lateral regions of the OBFC (areas VO and LO), between 2.7 and 4.2 mm anterior from bregma (Paxinos and Watson 1996).

from 24 to 61 (mean $\pm$ SEM: 38.9 ± 2.2) with mean firing rates ranging from 0.06 to 48.4 spikes/sec. Of these 894 neurons, 141 (16%) showed 176 statistically significant changes in firing rate correlated to one or more main events in the task (Table 1). These neural correlates consisted of responses during the following four events or phases: odor sampling, behavioral activity preceding the nose entry into the fluid well, the waiting period, and the delivery of reinforcement (Fig. 2). The remaining neurons failed to demonstrate any significant task-related modulation as revealed by the statistically assessed histograms and were not examined further. For the analysis of neural correlates, only positively rewarded trial types were taken into consideration.

During odor sampling, a total of 44 neural responses were found (25%), all consisting of an increase in firing (Fig. 2A). Many of the cells showing this correlate (45.5%, $n = 20$) started firing when the nose was already in the sampling port but before odor was presented, activity that might reflect the anticipation of odor delivery or the onset of behaviors related to odor sampling. All responses during odor sampling typically peaked within 1500 msec after odor onset. Because the task was not designed to determine whether differential responses during odor sampling were due to different sensory inputs (odor identity) or expectancy for different reward magnitudes, the results are inconclusive regarding the coding of expected reward magnitude during this phase. Therefore the responses observed during odor sampling were not analyzed further.

In “go” trials, the presentation of the odor stimulus was followed by the animal’s movement from the sampling port toward the fluid well and the subsequent waiting period in the fluid well before reinforcement delivery. In both phases, neurons demonstrated significant changes in firing. In the task phase, where the animal moved from the sampling port to the fluid well, 33 correlates (19%) consisting of an enhancement in firing rate were found (Fig. 2B). During the waiting period, 47 responses (27%) were found, one of which showing a decrease in firing

Table 1. Numbers of behavioral correlates

	Total	Discriminating between two sizes	Discriminating between three sizes	Total Discriminating
Odor sampling	44	—	—	—
Alone	35	—	—	—
& movement period	4	—	—	—
& waiting period	3	—	—	—
& reward delivery	2	—	—	—
Movement Period	33	7	10	17
Alone	28	6	8	14
& odor sampling	4	1	2	3
& waiting period	0	0	0	0
& reward delivery	1	0	0	0
Waiting Period	47	18	3	21
Alone	21	8	1	9
& odor sampling	3	2	0	2
& movement period	0	0	0	0
& reward delivery	23	8	2	10
Reward delivery	52 (42)	19	10	29
Alone	26 (19)	9	7	16
& odor sampling	2 (1)	1	0	1
& movement period	1	0	0	0
& waiting period	23 (21)	9	3	12
Total	176 (122)	44	23	67

Overview of the number of significant behavioral correlates observed during the task in a population of 141 units. Only correlates for the positively rewarded trial types were taken into consideration for calculating the overall number of responses. Responses indicated by "Alone" occurred only to the specific event indicated on the line above; the other responses correlated to two events ("dual correlates"). Units with dual correlates account for the fact that the total number of correlates (176) is higher than the total number of units (141). Responses during odor sampling were not examined in relation to reward magnitude (indicated by —; see text). Numbers of correlates after reward delivery between parentheses indicate the number of cells that were used to examine differential firing to the various reward sizes (the fifth type of reward response was left out of the analysis; see text).

activity, whereas the other 46 consisted of an enhancement in firing (Fig. 2C).

During the reinforcement delivery phase, all animals displayed similar behavior at the fluid well: for every amount of sucrose solution, their snouts remained in the fluid well until the end of trial. In response to positive reinforcement, 52 correlates were found (29%), consisting of five different response types. The first type of response consisted of a transient increase in firing rate, peaking within 500 msec after reinforcement delivery (27%, $n = 14$) (Fig. 2D), and the second type consisted of an increase in firing rate that peaked between 500 and 1500 msec after delivery (31%, $n = 16$). The third type of response consisted of an increase in firing rate starting almost immediately after the delivery of the reward, remaining during the entire period the animal visited the fluid well (15%, $n = 8$). The fourth type was a rarely encountered transient decrement in firing rate (8%, $n = 4$), whereas the fifth type comprised all other responses showing enhancements in firing rate starting at least 3 sec after reward delivery and with variable duration (19%, $n = 10$).

Neural activity during the movement and waiting period

After determining the various neural responses to task events, the question arose whether expectancy for the upcoming reinforcement was represented within these different types of neural correlates. Given the absence of changes in sensory input and the animal's overall immobility, neurons that showed responses within the waiting period in the fluid well may code the pre-

dicted outcome. However, to exclude the possibility that these responses were due to anticipatory licking of the animal instead of expected outcome per se, neural responses ($n = 22$) recorded during sessions in which lick detection was available were examined in relation to the extensions and retractions of the tongue in the fluid well. This showed that these responses did not covary with licking behavior since in these cases anticipatory licking was either completely absent, started after the neural response, or, when overlapping in time with the neural response, continued during reward consumption, when the neural response had already ended. Furthermore, a comparison between neural responses in this period between (incorrect) "go" responses during quinine trials versus "go" responses during sucrose trials revealed that 10 out of 11 neurons recorded showed differential firing for the sign of the response outcome during this period. Therefore, the present data are consistent with neural coding of reward expectancy in a subset of OBFC neurons during the waiting period.

Expectancy for reward, however, might also occur earlier in the task sequence, for example, before or during the movement period of the "go" response. To examine whether neural responses observed during the movement phase were related to the "go" movements of the animal per se or whether they reflected a truly goal-directed action in the context of the task, neural responses obtained during the correct "go" trials were compared with those during performance of the same behavior within the intertrial interval. This showed that these correlates were task dependent for 30 out of the 33 neurons exhibiting significant responses in this period: they were not observed when the behavioral sequence ("leave odor port and go to reward site") was performed during the intertrial interval (Fig. 3). In three units, a significant response occurred during the intertrial interval at a similar time point as during correct "go" trials, but these activations were not as strong as the responses observed during trials. In addition, a comparison between neural responses during (incorrect) "go" responses during quinine trials versus "go" responses during sucrose trials revealed that during the movement period, 12 out of 14 neurons recorded demonstrated differential firing depending on the response outcome.

Neural activity related to different magnitudes of reward

Whether neurons showed significant differences in response to the three different reward magnitudes was examined for three phases, namely, the responses occurring during the movement period preceding the fluid poke, the waiting period, and the period starting with reward delivery. A total number of 67 (55%) of the 122 correlates found during these three task phases demonstrated statistically significant differences between different amounts of reward, which was either between two (66%) or all three (34%) different reward sizes. The proportion of neurons

Figure 2. Overview of behavioral correlates observed during task performance. Peri-event time histograms and raster plots showing examples of all main behavioral correlates observed during the task. Examples from four different units recorded in four rats demonstrating correlates related to (A) odor sampling, synchronized on onset of odor presentation; (B) movement activity preceding nose entry into the fluid well, synchronized on fluid well entry; (C) the waiting period of 1500 msec in the fluid well, synchronized on onset of waiting; and (D) delivery of reinforcement, synchronized on reward delivery onset. These histograms, as well as the following histograms in Figures 3, 5, 6, are presented with a bin size of 100 msec. In all raster plots, individual consecutive trials are represented as horizontal lines, with the first trial at the top row. Horizontal scales show time (sec), vertical scales firing rate (Hz).

showing differential activity varied for the three task periods and ranged from 45% to 69% (Table 1).

Within the group of 33 neurons demonstrating task-specific neural correlates during the movement period preceding the fluid poke, 17 neurons (52%) demonstrated statistically significant differential firing across different reward sizes, with seven and 10 neurons discriminating between two and three different sizes, respectively (Fig. 3). Some neurons displayed firing activity that increased with incremental reward size (18%, $n = 3$ out of 17), whereas other neurons displayed the inverse relationship (29%, $n = 5$) or demonstrated the largest or smallest response to the middle reward size (53%, $n = 9$) (Fig. 4A). During the waiting period, in which 47 neurons showed significant responses, 21 units (45%) discriminated significantly between different magnitudes of reward (Fig. 5). A total number of 18 of these 21 neurons showed differential activity between two different reward sizes, three neurons between three sizes. A similar heterogeneity within the tuning of the neurons was found as during the movement period: units showed a rise in firing activity with increasing reward size (29%, $n = 6$ out of 21), whereas other neurons displayed the inverse relationship (24%, $n = 5$) or showed the largest or smallest activation to the middle reward size (48%, $n = 10$) (Fig. 4B).

Besides expressing neural correlates related to reward magnitude in the periods preceding reinforcement delivery, OBFC neurons may code aspects related to the actual reward magnitude. After examining whether neural responses ($n = 10$) in the period of reward delivery were related to the tongue movements of the animal during licking, none of the responses could be

ascribed to licking behavior. Furthermore, due to the different nature of the responses observed in this period, determining whether neurons demonstrated discriminatory responses to the different magnitudes of reward was done using different criteria for the various response types. For the transient response types, only the time periods in which the response occurred were compared; for the sustained responses, the entire period with enhanced firing was tested. Responses of the fifth type were left out of the analysis due to the variability of the responses within this group of neurons. Of the 42 neurons showing responses of the remaining four types, 29 neurons (69%) showed significant differential firing across reward sizes, with 19 and 10 neurons discriminating between two and three different reward sizes, respectively (Fig. 6). Again, all types of response profiles were demonstrated: increasing (21%, $n = 6$ out of 29) or decreasing (14%, $n = 4$) firing activity with increasing reward size, or the largest or smallest response to the middle reward size (66%, $n = 19$) (Fig. 4C).

An additional Kolmogorov-Smirnov test ($P < 0.05$) was performed for the responses that peaked between 500 and 1500 msec (Fig. 6A, second response type), since this subgroup often showed a response time course varying with reward magnitude. Taking firing rate values for successive bins as variable, this test revealed that for all neurons of this second type ($n = 16$), the time course of the response differed either between two (53%) or three (47%) reward sizes as well.

Discussion

The present experiment was performed to examine whether and how information concerning reward magnitude is coded in rat

Figure 3. Differential firing in relation to expected reward magnitude during the movement period. (A) Example of one unit demonstrating an increase in firing rate after leaving the odor port and before nose entry into the fluid well. Activity is synchronized on nose entry into the fluid well; red marks indicate offset of the odor poke. This single unit discriminated between the two smallest and the largest amounts of reward: there was no significant difference in firing between the two lowest reward sizes, but both responses were significantly higher compared to the largest amount of reward. (B) Activity of the same unit during the performance of the same behavioral sequence as in A, but now in the intertrial interval (ITI).

OBFc. A total number of 141 out of 894 neurons recorded (16%) demonstrated a behavioral correlate of firing activity during several events in the task, that is, odor sampling, the movement period from odor port to the fluid delivery well, the waiting period in the fluid well, and the reward delivery period. During the movement and waiting period, respectively, 52% and 45% of the neurons showing a significant behavioral correlate demonstrated differential firing across varying reward sizes. After reward delivery, 69% of the neurons showed differential responses toward the various reward sizes.

Overall, the percentage of neurons showing significant correlations during this olfactory discrimination task (16%) was lower than in a previous study with an eight-odor olfactory discrimination task (Schoenbaum and Eichenbaum 1995). Apart from differences in the behavioral paradigms, this might be due to the conservative statistical assessment in the present study, to methodological differences in unit recording and isolation, or to the fact that recording sites in the latter study were located more laterally in the OBFc and also comprised the ventral part of the agranular insular cortex, whereas in the present study locations of recordings were limited to the ventral and lateral orbital regions.

Testing the ability of animals in a T-maze task to discriminate between the different reward sizes used in this behavioral task demonstrated that animals were able to discriminate the various reward sizes. In the five-odor discrimination task, the scores on overall response time revealed that animals responded significantly slower for the smallest amount of reward, indicating that animals learned to discriminate at least between the smallest and the two larger amounts of reward. The finding of this difference not being present within the movement time might be explained by the limited number of trials. The differences could have turned out to be significant when the number of trials per

session had been larger. However, other possible explanations include a higher degree of stereotypy or habit-based performance of “go” responses, once initiated, than applies to the odor sampling phase. That the movement time was highly similar across different reward sizes makes it likely that the discriminatory neural responses during the movement period were not due to differences in motor-related aspects, including the speed of responding.

Neural responses during the reward delivery phase: Magnitude effects

As concerns actual reward magnitude, the results indicate that neurons in rat OBFc may, indeed, code information related to this reinforcement parameter, showing differential firing across various reward sizes during reward consumption. Although this discriminative firing appears not to be due to licking behavior, there is the possibility of modulation of these responses by sensory-motor processes (e.g., tasting or ingestion), insofar as this is distinguishable from reward processing per se. The observation that the OBFc is involved in the processing of information related to the magnitude of the actual reinforcement is in line with the fMRI result in humans that distinct regions within the OBFc represent the magnitude of monetary reward (O’Doherty et al. 2001). Furthermore, it fits in with earlier findings concerning neural coding of other aspects of reinforcers in the OBFc, such as qualitative properties or motivational value (Thorpe et al. 1983; Tremblay and Schultz 1999; Hikosaka and Watanabe 2000; Wallis and Miller 2003; Roesch and Olson 2004; Ichihara-Takeda and Funahashi 2006; Padoa-Schioppa and Assad 2006). Together with the present result, this implies that at least some major aspects of reinforcers that might be influential for the selection of the most appropriate response strategy are represented within the OBFc.

Figure 4. Overview of the response profiles to reward size during different task phases, that is, responses in (A) the movement period, (B) the waiting period prior to reinforcement delivery, and (C) reward delivery. Only profiles of units with single behavioral correlates (“alone” in Table 1) and with significant differences between at least two reward sizes are shown. In each graph, different units are represented by different colors and symbols. Some neurons changed their firing pattern monotonically with increasing or decreasing reward size, whereas others showed the largest or smallest response to the middle reward size. On the horizontal scale, reward size (in milliliters) is plotted; the vertical scale displays the peak firing rate of individual units in association with different reward sizes, normalized to the response to the 0.05-mL reward.

Neural coding of reward expectancy

In a previous study by Schoenbaum et al. (1998), neurons in the OFC were suggested to code expected outcomes, since they demonstrated discriminatory firing during the delay period in which animals awaited either positive or negative reinforcement, which

is a result confirmed in this study. However, a confounding factor in the study by Schoenbaum et al. (1998) was the comparison between “go” responses for both positive and negative reinforcers, the latter being erroneously made. The possibility of such a confounding factor is underscored by a study of Ramus and Eichenbaum (2000) on olfactory recognition memory, who reported units in rat orbitofrontal cortex showing firing rate changes during odor sampling that occurred specifically prior to erroneous “go” responses. The advantage of the parametric design in the present study for the examination of reward expectancy is that very similar “go” responses can be compared, since the impending outcomes are all appetitive. Indeed, we found no significant difference in movement time across the three positively rewarded trial types, while the waiting period was fixed in duration (1.5 sec). Thus, the present design may provide more certainty about the nature of the observed neural correlate, since no change or switch in response strategy is required during these trials.

Nonetheless, it remains possible that other cognitive pro-

Figure 5. Differential firing in relation to expected reward magnitude during the waiting period. Example of one single unit. Activity, synchronized on nose entry into the fluid well, did not differ between the two lowest amounts of reward, but both differed significantly from the largest amount of reward. Highest activity was demonstrated for the smaller reward sizes. Vertically aligned red marks indicate reinforcement delivery at 1500 msec.

Figure 6. Differential firing to reward magnitude after reward delivery. Examples of two different units showing differential firing toward the different reward sizes after reward delivery. (A, left) A unit is shown demonstrating firing that differed significantly between the smallest and the two largest rewards, with the largest response for the larger reward sizes. (B, right) A different unit is shown demonstrating differential activity between the two smallest and the largest amounts of reward, with the largest response for the lower reward sizes.

cesses besides reward expectancy, like working memory or attention, although not obviously required for correct task performance, are reflected within the waiting or movement period. An additional remark is that, although learning of the discrimination problem by the animals was obvious from both “go”/“no-go” decisions, movement and overall response time for all five odors, it would be difficult to actually show learning-related changes in single-unit firing. Animals learned to discriminate positively versus negatively reinforced odors within a few trials (average for quinine trials: 4.5 ± 0.7), and the number of spikes during the waiting period per trial appears insufficient to permit a robust analysis of the evolution of the neural representation of stimulus–reward contingencies across these few trials. Resolving this issue will necessitate a new study in which learning is temporally spread out across more trials. Despite this, it does not seem likely that the reward-size differences in neural activity would be attributable to carryover effects from previous sessions. First, new odors were used in every new session, and care was taken to avoid constant reinforcement contingencies of the odor families used here. Second, interference by previously established contingencies would be expected to prevent the emergence of a consistently slower response to the odor predicting the

lowest amount of reward as compared to the high-reward contingencies.

Reward expectancy: Timing and magnitude effects

Discriminatory activations toward different magnitudes of reward were found to occur during the waiting period as well as during the movement period, the latter likely being an expectancy-related correlate during the “go” response that to our knowledge has not been described earlier in rat or primate OFC. In addition to the rat OFC, expectancy-related activity was previously found in rat basolateral amygdala (BLA), with neurons demonstrating selective firing during a delay period prior to reinforcement or during predictive odor cues after learning (Schoenbaum et al. 1998, 1999; for review, see Holland and Gallagher 2004). In addition to these two types of expectancy-related activity, the results of the present study indicate that at least in the OFC, reward expectancy may also be coded during the movement phase of the “go” response. Although this neural correlate occurred during motor behavior, it was absent when the animal executed the same movement outside the context of the trials, thus excluding a pure motor confound. The presence of

this expectancy-related signal matches the results from lesion studies indicating the involvement of the OFC in the flexible guidance of goal-directed behavior (Dias et al. 1996; Meunier et al. 1997; Hornak et al. 2004; Izquierdo et al. 2004) and is also consistent with a role for OFC in encoding motor set as well as action-outcome relationships (Baxter et al. 2000). Especially when reward contingencies are changing, information representing these relationships should be available during task performance, particularly before and during the execution of a behavioral response. The idea that a reward expectancy signal is available during ongoing behavior is supported furthermore by the finding that neurons in the primate OFC may code both long- and short-range reward expectancy (Hikosaka and Watanabe 2004). It is suggested by these investigators that during behavior these different OFC signals serve to adjust motivational levels across different temporal ranges, namely, toward an immediate or a more distant outcome. Altogether, the available results support the notion that reinforcement parameters such as magnitude and quality are coded across different temporal phases along task trials, corroborating the hypothesis that OFC neurons collectively code a matrix of reward parameters as a function of the delay toward the moment of outcome. Future recordings will be needed to study the function of neurons with expectancy-related firing during movement periods in more detail, for example, during different motor behaviors or reversal learning.

A previous study by Roesch et al. (2006) already demonstrated neurons within rat OFC discriminating between two different reward sizes. However, apart from the fact that reward size was not parametrically varied (only two different reward sizes were used), their results are not directly comparable with the results from the present study. Additional differences concern the method of reward application and location of reward delivery. Roesch et al. (2006) applied a single bolus of 0.05 mL as a small reward and an additional drop 500 msec after the first bolus as the large reward, with different reward sizes administered at different locations. In contrast, in the present study all reward sizes were delivered in a single dose and at a common location.

The finding of the present study that information concerning the magnitude of reward is represented in rat OFC is in line with one of the predictions from the HSAT (Hebbian Synapses with Adaptive Thresholds) model of RL (Pennartz 1997), in which the OFC, together with other related limbic areas, functions as a brain center for processing and predicting reinforcement, as well as directing Hebbian changes in synaptic strength in brain areas targeted by the OFC. The presence of information comparing the actual versus the expected magnitude of reinforcers within the OFC was modeled to direct behavior of a neural network performing parallel distributed processing toward the most profitable outcome in an efficient and flexible manner. In particular, the observation that about one-half of the OFC neurons (23 out of 52) demonstrating responses during reward delivery also showed expectancy-related activity during the waiting period is in line with the model, since in the course of learning, neurons within the reinforcement processing module of the HSAT network that are initially activated by the reward will become activated during the expectancy period preceding reward as well. Besides the OFC, other brain areas emitting glutamatergic fibers may serve as candidate regions for the reinforcement processing posited by this model, such as medial prefrontal cortex (Pratt and Mizumori 2001), anterior cingulate cortex, and dorso-lateral prefrontal cortex (Watanabe 1996; Leon and Shadlen 1999; Tremblay and Schultz 1999; Shidara and Richmond 2002). That the OFC neurons in the present study exhibited a variety of different tuning curves to reward magnitude—either during the expectancy phase or actual reward delivery—constitutes a result that was not directly predicted by the computational model. This

finding is in line with results of Wallis and Miller (2003) demonstrating that neurons within monkey OFC code expected reward magnitude both parametrically and nonparametrically. Such a diversity of tuning is as compatible with coding in a parallel-distributed network as is a monotonic relationship between firing activity and reward magnitude. Adopting a wider perspective on RL than prescribed by any particular model, it should be mentioned that reward variables, including magnitude, are also represented in the striatum (Hollerman et al. 1998; Hassani et al. 2001; Cromwell and Schultz 2003). Although recent results by Pasupathy and Miller (2005) showed that learning-related changes in firing activity in a cue-saccade association task occurred earlier in striatum than prefrontal cortex, the actual anatomical loci mediating stimulus–response and action–outcome learning remain to be firmly established. The distribution of the information across several cortical and subcortical areas suggests a broader computational system by which RL can be mediated than a basic actor–critic architecture trained by temporal difference learning that is composed of the striatum and mesencephalic dopamine system (for review, see Montague et al. 2004).

Materials and Methods

Subjects

Data were collected from seven male Wistar rats (Harlan CPB; Horst, The Netherlands), weighing 325–425 g at the time of surgery. Animals were socially housed in standard macrolon cages, weighed and handled daily, and kept under a reversed 12-h light/dark cycle (lights off at 7:00 a.m.) with food available *ad libitum* (standard rat chow; Hope Farms, The Netherlands). Water deprivation started overnight prior to behavioral testing to motivate the animals to perform the task. During performance animals received a maximum of 7.5 mL of fluid, and after finishing the session there was, with a variable delay, free access to water for a 2-h period. After surgery, the animals were housed individually under the same conditions. All experiments were carried out in accordance with the National Guidelines for Animal Experimentation.

Behavior

Apparatus

Behavioral testing was performed in a Plexiglas operant chamber that was placed in a sound-attenuated and electrically shielded box, with behavioral events and data collection controlled by a computer. The recording chamber (40 × 37 × 41.5 cm) had a black interior with straight walls and a front panel that contained on the right side an odor sampling port beneath a light indicating trial onset, and on the left side a fluid delivery well. The lower part of the front panel was placed at an angle of 101° with respect to the grid floor, with an additional angle higher up of 161° with respect to the lower part of the front panel to ensure that during recordings, when animals were attached to the recording equipment, ample space was left for the rat to put its snout into the odor sampling port. To detect the odor and fluid poke responses made by the animals, both the odor sampling port and fluid delivery well were equipped with an infrared beam transmitter and detector. During six recording sessions, licking behavior in the fluid well was monitored with the use of an electronic circuit with fluid contact causing a current of 150 nA and a concomitant change in voltage.

The delivery of odors was regulated by a system of solenoid valves and flow meters (cf. Schoenbaum 2002) with separate delivery lines for each odor to prevent mixture of odors in the system. To be able to deliver different types of fluid reinforcement (i.e., quinine and sucrose solutions), separate fluid delivery lines were present. Fluid delivery was gravity-driven, with a tap and valves controlling the flow and amount of fluid delivered. The odorants (Tokos BV) were separated into different families, that is, herbal, floral, woody, citrus, and fruity. In order to have

distinct odors in a set of five odors used in a single discrimination session, each set contained one odor from each family. Care was taken to ensure that no single family of odors was preferentially associated with a particular trial outcome (positive versus negative and in terms of reward size).

Shaping and behavioral paradigm

After the animals were habituated to the recording chamber, they were gradually trained on the behavioral procedure of the five-odor olfactory discrimination “go-no/go” task. A set of five different odors was used for any discrimination session: three odors associated with a particular amount of a positive reinforcement (10% sucrose in water, i.e., 0.05, 0.15, and 0.30 mL), one odor with no reward (nonreinforced condition), and one odor with a negative reinforcement (0.15 mL of a 0.015 M quinine solution in water). A pilot test in a T-maze was performed to examine the animal’s ability to discriminate between the volumes of sucrose solution used. This test demonstrated that animals selected the larger amount of reward when offered a choice between two alternatives (e.g., 0.05 versus 0.15 mL or 0.15 versus 0.30 mL), indicating they were able to discriminate between the different reward sizes. The odors coupled to the nonreinforced condition and quinine were used to determine whether the animal was actually capable of associating an odor with a particular outcome, as should be visible by withholding responses toward or at the fluid well after sampling the odors that were predictive of these two outcomes.

During shaping, animals were initially trained to make a nose poke in the odor sampling port (“odor poke”), which was sufficient to immediately obtain reinforcement by visiting the fluid well. At this stage only two different odors were used, associated with 0.05 and 0.15 mL of sucrose solution. At the next stage, a third odor was introduced, associated with 0.30 mL of sucrose solution. In this phase, animals were trained to make an odor poke with a minimal duration of 2 sec: the animals had to wait 1 sec in the odor sampling port before the odor was presented in order to have their body stationary during cue sampling, and odor sampling itself was required to last at least 1 sec. In the final stage of shaping, the odors coupled to the nonreinforced condition and quinine were introduced together with a waiting period of 1500 msec in the fluid well following the nose poke in the well (“fluid poke”) before reinforcement was delivered. The behavioral sequence comprising the departure from the sampling port to the fluid well, including nose entry and waiting period in the well, is referred to as the “go” response.

The criterion for behavioral performance was set at 15 trials per positively rewarded trial type since it was difficult to achieve a reliable acquisition of many more trials than 15 due to the larger reward volumes used in this task. Once they reached this criterion, animals were implanted with a headstage containing an array with individually movable tetrodes (“hyperdrive”), and recordings were started. In the first recording session, the set of odors that was used during shaping was presented again, in order to retrain the animals with familiar odors. Provided that the performance was back at presurgical level, a new set of five odors was given the next session. Each time the discrimination had been learned, as visible by withholding responses toward or at the fluid well during quinine or nonreinforced trial types (usually within one session), a new set of five odors was presented.

During the task, odors were presented in a pseudorandom order. After the trial light switched on, the animal had 15 sec to make an odor poke. If no odor poke was made, the light turned off and an intertrial interval (with a variable duration of 10–25 sec) started. Whenever an odor poke was made, the trial light switched off after 500 msec, followed 500 msec later by the presentation of an odor. After retraction of the animal’s nose out of the odor sampling port or whenever a maximal duration for odor sampling (10 sec) was exceeded, odor presentation was terminated. Premature retraction from the odor sampling port (odor poke shorter than the minimal duration of 2 sec) resulted in the intertrial interval. Whenever the animal received reinforcement in the fluid well, it had to consume the fluid within 10 sec, after which the fluid well was drained by a vacuum line and the in-

tertrial interval started. An incorrect (“go”) response after sampling an odor predictive of quinine or the nonrewarded contingency had no further programmed consequences.

Surgery and electrophysiology

Animals were anaesthetized with 0.08 mL/100 g Hypnorm i.m. (0.2 mg/mL fentanyl, 10 mg/mL fluanison) and 0.04 mL/100 g Dormicum s.c. (midazolam 1 mg/kg) and mounted in a Kopf stereotaxic frame. Additional local anesthesia (Xylocaine spray; 10%, Astra) was also given. Body temperature was maintained at 37.5°C using a heating pad. After exposure of the cranium, five small holes were drilled into the cranium to accommodate surgical screws, one of which serving as ground. Another larger hole was drilled over the OBFC in the left hemisphere (3.2 mm anterior, 3.2 mm lateral to bregma according to Paxinos and Watson 1996). After opening of the dura, the base of a hyperdrive (an array of 12 individually drivable tetrodes [13 µm nichrome wire; Kanthal] and two reference electrodes, spaced apart by at least 310 µm) was lowered onto the exposed cortex (Wilson and McNaughton 1993; Gray et al. 1995; Pennartz et al. 2004). The hyperdrive was anchored to the screws with dental cement. Immediately after surgery, all tetrodes were advanced 1 mm into the brain; in the course of the next 3 d, the tetrodes were gradually lowered until the OBFC was reached. Animals were allowed to recover at least 7 d before re-exposure to the behavioral paradigm and the start of the recordings. Recording sessions were initiated when all tetrodes were estimated to have reached the OBFC and the animal had obtained its presurgical performance level. In order to record different units during each recording session, all tetrodes were lowered at the start of a recording day with increments of 40 µm. Depending on the amount of neural activity, individual tetrodes could be advanced further. Once the tetrodes were lowered, the animal was left to rest in his home cage for at least 2 h in view of unit recording stability, after which the experimental session started.

Electrophysiological recordings were performed using a Cheetah recording system (Neuralynx). Signals from the individual leads of the tetrodes were passed through a low noise unity-gain field-effect transistor preamplifier, insulated multiwire cables, and a 72-channel commutator (Dragonfly) to digitally programmable amplifiers (gain 5000 times; band-pass filtering 0.6–6.0 kHz). Amplifier output was digitized at 32 kHz and stored on a Windows NT station. A 1-msec data sample was taken whenever the signal crossed a preset voltage boundary, so that the width of a window containing the recorded spike was captured in 32 data points. The occurrence of task events in the behavioral chamber was recorded simultaneously, and the behavior of the animals was recorded on videotape.

Data analysis

Single units were isolated by off-line cluster cutting procedures (BBClust/MClust 3.0). Before a cluster of spikes was accepted as a single unit, several parameters and graphs were checked visually, namely, the averaged waveform across the four leads, the cluster plots showing spike parameter distributions such as peak amplitudes across the four dimensions, the autocorrelogram, and the spike interval histogram. Since the absence of spike activity during the refractory period (2–3 msec) is indicative for good isolation, units of which the autocorrelogram and the spike interval histogram revealed activity during this period were removed from the analysis.

Correlations between events in the task and changes in firing rate were examined by constructing peri-event time histograms (PETHs) and statistically assessed with the nonparametric Wilcoxon matched-pairs signed-rank (WMPSR) test ($P < 0.01$) with bin resolutions of 100 and 1500 msec. The 1500-msec bin resolution was used to examine the significance of neural correlates during broader periods such as the entire waiting period of 1500 msec, whereas the 100-msec bin resolution was used to statistically examine the more exact time course of the response. Neural responses were considered significant if firing rates, quantified per bin, were significantly different from a control (base-

line) period during the intertrial interval. This control period consisted of five bins, and any of the bins tested for a significant change in firing during the trial was required to differ significantly from each of these five control bins. In addition, responses had to be significant for both bin resolutions to be considered as such. This procedure enhanced conservativeness in the identification of neural correlates of task events, and, by using a nonparametric test, avoids a number of difficulties and assumptions associated with some other tests that are related to the nature of spike timing distributions and their deviations from baseline activity.

Once the WMPSR test indicated a significant deviation in firing rate with respect to baseline, the nonparametric Kruskal-Wallis test ($P < 0.05$), followed by a post hoc Mann-Whitney U-test ($P < 0.05$) were used to assess response differences between PETHs pertaining to different odor–reward magnitude pairs.

Behavioral data were analyzed using SPSS for Windows (version 12.0.1). Unless otherwise stated, results are expressed as mean $\pm$ SEM values. Movement Time was defined as the interval between nose retraction from the odor port and nose entry into the fluid well, whereas the Overall Response Time was defined as the duration of the entire sequence of odor sampling and moving to the fluid well. The mean response times per reward magnitude were obtained from all different trial types within all the sessions that were used for analysis. These measures were compared across different trial types with the nonparametric Kruskal-Wallis test ($P < 0.05$), followed by a post hoc Mann-Whitney U-test ($P < 0.05$).

Histology

The final position of the tetrodes was marked by passing a 10-sec, 25 μ A current through one of the leads of each tetrode in order to induce a lesion and initiate gliosis. The next day, ~24 h after lesioning, animals were perfused transcardially using a 0.9% saline solution followed by 10% formalin. After removal from the skull, the brain was stored in a 10% formalin solution for several days before sectioning. Brain sections (40 μ m) were cut using a vibratome and Nissl-stained to reconstruct the tetrode tracks and their final positions.

Acknowledgments

This work was supported by NWO Grant 903-47-084, NWO grant 918.46.609, and BSIK (SenterNovem) grant 03053. We thank Geoffrey Schoenbaum for providing information about the behavioral setup, Bruce McNaughton for his help with the use of tetrode arrays, and David Redish and Peter Lipa for providing the cluster cutting software. Furthermore, we thank Eunjeong Lee and Ton Put for their contribution to the data analysis and graphical illustrations, respectively; and our colleagues at the electronic and mechanical workshop of the Netherlands Institute for Neuroscience for their excellent technical assistance.

References

- Balleine, B.W. and Dickinson, A. 1998. Goal-directed instrumental action: Contingency and incentive learning and their cortical substrates. *Neuropharmacology* **37**: 407–419.
- Baxter, M.G., Parker, A., Lindner, C.C.C., Izquierdo, A.D., and Murray, E.A. 2000. Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *J. Neurosci.* **20**: 4311–4319.
- Black, R.W. 1968. Shifts in magnitude of reward and contrast effects in instrumental and selective learning: A reinterpretation. *Psychol. Rev.* **75**: 114–126.
- Bohn, I., Gierler, C., and Hauber, W. 2003. Orbital frontal cortex and guidance of instrumental behaviour in rats under reversal conditions. *Behav. Brain Res.* **143**: 49–56.
- Boysen, S.T., Berntson, G.G., and Mukobi, K.L. 2001. Size matters: Impact of item size and quantity on array choice by chimpanzees (*Pan Troglodytes*). *J. Comp. Psychol.* **115**: 106–110.
- Brown, V.J. and Bowman, E.M. 1995. Discriminative cues indicating reward magnitude continue to determine reaction time of rats following lesions of the nucleus accumbens. *Eur. J. Neurosci.* **7**: 2479–2485.
- Campbell, A.B. and Seiden, L.S. 1974. The effect of relative and absolute reinforcement magnitude on operant responding. *Physiol. Behav.* **12**: 843–849.
- Cromwell, H.C. and Schultz, W. 2003. Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J. Neurophysiol.* **89**: 2823–2828.
- Dayan, P. and Balleine, B.W. 2002. Reward, motivation, and reinforcement learning. *Neuron* **36**: 285–298.
- Dias, R., Robbins, T.W., and Roberts, A.C. 1996. Dissociation in prefrontal cortex of affective and attentional shifts. *Nature* **380**: 69–72.
- Ferry, A.T., Lu, X.C., and Price, J.L. 2000. Effects of excitotoxic lesions in the ventral striatopallidal-thalamocortical pathway on odor reversal learning: Inability to extinguish an incorrect response. *Exp. Brain Res.* **131**: 320–335.
- Gallagher, M., McMahan, R.W., and Schoenbaum, G. 1999. Orbitofrontal cortex and representation of incentive value in associative learning. *J. Neurosci.* **19**: 6610–6614.
- Gray, C.M., Maldonado, P.E., Wilson, M.A., and McNaughton, B.L. 1995. Tetrodes markedly improve the reliability and yield of multiple single-unit isolation from multi-unit recordings in cat striate cortex. *J. Neurosci. Methods* **63**: 43–54.
- Hassani, O.K., Cromwell, H.C., and Schultz, W. 2001. Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *J. Neurophysiol.* **85**: 2477–2489.
- Hikosaka, K. and Watanabe, M. 2000. Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. *Cereb. Cortex* **10**: 263–271.
- Hikosaka, K. and Watanabe, M. 2004. Long- and short-range reward expectancy in the primate orbitofrontal cortex. *Eur. J. Neurosci.* **19**: 1046–1054.
- Holland, P.C. and Gallagher, M. 2004. Amygdala–frontal interactions and reward expectancy. *Curr. Opin. Neurobiol.* **14**: 148–155.
- Hollerman, J.R., Tremblay, L., and Schultz, W. 1998. Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J. Neurophysiol.* **80**: 947–963.
- Hornak, J., O'Doherty, J., Bramham, J., Rolls, E.T., Morris, R.G., Bullock, P.R., and Polkey, C.E. 2004. Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. *J. Cogn. Neurosci.* **16**: 463–478.
- Ichihara-Takeda, S. and Funahashi, S. 2006. Reward-period activity in primate dorsolateral prefrontal and orbitofrontal neurons is affected by reward schedules. *J. Cogn. Neurosci.* **18**: 212–226.
- Izquierdo, A.D., Suda, R.K., and Murray, E.A. 2004. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *J. Neurosci.* **24**: 7540–7548.
- Leon, M.I. and Shadlen, M.N. 1999. Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron* **24**: 415–425.
- Lipton, P.A., Alvarez, P., and Eichenbaum, H. 1999. Crossmodal associative memory representations in rodent orbitofrontal cortex. *Neuron* **22**: 349–359.
- Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M., and Robbins, T. 2002. Decision-making processes following damage to the prefrontal cortex. *Brain* **125**: 624–639.
- Meunier, M., Bachevalier, J., and Mishkin, M. 1997. Effects of orbital frontal and anterior cingulate lesions on object and spatial memory in rhesus monkeys. *Neuropsychologia* **35**: 999–1015.
- Montague, P.R., Hyman, S.E., and Cohen, J.D. 2004. Computational roles for dopamine in behavioural control. *Nature* **431**: 760–767.
- O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., and Andrews, C. 2001. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat. Neurosci.* **4**: 95–102.
- O'Doherty, J., Critchley, H., Deichmann, R., and Dolan, R.J. 2003. Dissociating valence of outcome from behavioural control in human orbital and ventral prefrontal cortices. *J. Neurosci.* **23**: 7931–7939.
- Padoa-Schioppa, C. and Assad, J.A. 2006. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**: 223–226.
- Pasupathy, A. and Miller, E.K. 2005. Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* **433**: 873–876.
- Paxinos, G. and Watson, C. 1996. *The rat brain in stereotaxic coordinates*. Academic Press, New York.
- Pennartz, C.M.A. 1997. Reinforcement learning by hebbian synapses with adaptive thresholds. *Neuroscience* **81**: 303–319.
- Pennartz, C.M.A., Lee, E., Verheul, J., Barnes, C.A., and McNaughton, B.L. 2004. The ventral striatum in off-line processing: Ensemble reactivation during sleep and modulation by hippocampal ripples. *J. Neurosci.* **24**: 6446–6456.
- Pratt, W.E. and Mizumori, S.J.Y. 2001. Neurons in rat medial prefrontal cortex show anticipatory rate changes to predictable differential rewards in a spatial memory task. *Behav. Brain Res.* **123**: 165–183.

- Ramus, S.J. and Eichenbaum, H. 2000. Neural correlates of olfactory recognition memory in the rat orbitofrontal cortex. *J. Neurosci.* **20**: 8199–8208.
- Roesch, M.R. and Olson, C.R. 2004. Neuronal activity to reward value and motivation in primate frontal cortex. *Science* **304**: 307–310.
- Roesch, M.R., Taylor, A.R., and Schoenbaum, G. 2006. Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron* **51**: 509–520.
- Schoenbaum, G. 2002. Olfactory learning and the neurophysiological study of rat prefrontal functioning. In *CRC Series: Methods in chemosensory research* (eds. S.A. Simon and M.A.L. Nicolelis), pp. 371–427. CRC Press, New York.
- 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.
- Schoenbaum, G., Chiba, A.A., and Gallagher, M. 1998. Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nat. Neurosci.* **1**: 155–159.
- Schoenbaum, G., Chiba, A.A., and Gallagher, M. 1999. Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. *J. Neurosci.* **19**: 1876–1884.
- Schoenbaum, G., Nugent, S.L., Saddoris, M.P., and Setlow, B. 2002. Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. *Neuroreport* **13**: 885–890.
- Schoenbaum, G., Setlow, B., Saddoris, M.P., and Gallagher, M. 2003. Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from the basolateral amygdala. *Neuron* **39**: 855–867.
- Schultz, W., Dayan, P., and Montague, P.R. 1997. A neural substrate of prediction and reward. *Science* **27**: 1593–1599.
- Shidara, M. and Richmond, B.J. 2002. Anterior cingulate: Single neuronal signals related to the degree of reward expectancy. *Science* **296**: 1709–1711.
- Sutton, R.S. and Barto, A.G. 1981. Towards a modern theory of adaptive networks: Expectation and prediction. *Psychol. Rev.* **88**: 135–170.
- Thorpe, S.J., Rolls, E.T., and Maddison, S. 1983. The orbitofrontal cortex: Neuronal activity in the behaving monkey. *Exp. Brain Res.* **49**: 93–115.
- Tremblay, L. and Schultz, W. 1999. Relative reward preference in primate orbitofrontal cortex. *Nature* **398**: 704–708.
- Wallis, J.D. and Miller, E.K. 2003. Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur. J. Neurosci.* **18**: 2069–2081.
- Watanabe, M. 1996. Reward expectancy in primate prefrontal neurons. *Nature* **382**: 629–632.
- Wilson, M.A. and McNaughton, B.L. 1993. Dynamics of the hippocampal ensemble code for space. *Science* **261**: 1055–1058.
- Yonemori, M., Nishijo, H., Uwano, T., Tamura, R., Furuta, I., Kawasaki, M., Takashima, Y., and Ono, T. 2000. Orbital cortex neuronal responses during an odour-based conditioned associative task in rats. *Neuroscience* **95**: 691–703.

Received January 25, 2007; accepted in revised form April 5, 2007.

Neural coding of reward magnitude in the orbitofrontal cortex of the rat during a five-odor olfactory discrimination task

Esther van Duuren, Francisco A. Nieto Escámez, Ruud N.J.M.A. Joosten, et al.

Learn. Mem. 2007, **14**:

Access the most recent version at doi:[10.1101/lm.546207](https://doi.org/10.1101/lm.546207)

References

This article cites 50 articles, 10 of which can be accessed free at:
<http://learnmem.cshlp.org/content/14/6/446.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).