Reward-Related Neuronal Activity During Go-Nogo Task Performance in Primate Orbitofrontal Cortex

LÉON TREMBLAY AND WOLFRAM SCHULTZ
Institute of Physiology and Program in Neuroscience, University of Fribourg, CH-1700 Fribourg, Switzerland

INTRODUCTION

One of the least charted territories of the primate cortex appears to be the orbitofrontal part of the frontal lobe. Its functions are defined largely by anatomic connections to brain centers whose functions are better known and by the deficits after lesions in human patients and experimental animals which concern altered and reduced emotional reactions to environmental changes (Butter et al. 1970; Damasio 1994; Hornak et al. 1996). Primates with orbitofrontal lesions show altered reactions to rewarding and aversive events (Baylis and Gaffan 1991; Butter and Snyder 1972; Butter et al. 1969, 1970) and impaired adaptations to changed reinforcement contingencies (Butter 1969; Dias et al. 1996; Iversen and Mishkin 1970; Jones and Mishkin 1972; Passingham 1972; Rosenkilde 1979). Medial orbitofrontal lesions in primates lead to deficits in visual discrimination and matching tests (Bachevalier and Mishkin 1986; Baylis and Gaffan 1991; Kowalska et al. 1991; Mishkin and Manning 1978; Passingham 1975; Voytko 1985). A role in reward processing is suggested by strong inputs from basal amygdala (Porrino et al. 1981; Potter and Nauta 1979) and, transsynaptically, from ventral striatum (Haber et al. 1995) with its reward-related neurons (Nishijio et al. 1988; Schultz et al. 1992). Heavy inputs arise also from medial temporal cortex structures whose roles in reward processes are less known (Barbas 1988, 1993; Carmichael and Price 1995; Seltzer and Pandya 1989; Ungerleider et al. 1989).

Neurophysiological investigations of orbitofrontal cortex addressed mnemonic functions in delayed response paradigms typical for the functions of prefrontal cortex (Jacobsen and Nissen 1937). Orbitofrontal neurons showed smaller activations during the delay periods of spatial and object matching tasks compared with dorsolateral prefrontal cortex but responded to delivery of juice reward at the end of the trial (Niki et al. 1972; Rosenkilde et al. 1981). Orbitofrontal neurons discriminated between primary and conditioned appetitive and aversive stimuli and were activated specifically in extinction or reversal trials (Thorpe et al. 1983). Neurons in the caudally adjoining orbitofrontal taste area showed specific gustatory and olfactory responses that were modified in relation to the animal’s satiation (Rolls and Baylis 1994; Rolls et al. 1989, 1996). Thus orbitofrontal neurons may respond to rewards in a manner appropriate for reinforcing behavioral reactions.

To better understand neuronal mechanisms underlying the motivational control of behavior, we investigated the neuronal processing of reward information in brain structures participating in the control of behavior. After the characterization of different forms of reward processing in primate striatum (caudate nucleus, putamen, and ventral striatum) (Apicella et al. 1991, 1992; Hollerman et al. 1998; Schultz et al. 1992), we searched for inputs that could possibly contribute to striatal reward-related activations. One of the principal candidates is the orbitofrontal cortex, which strongly projects to the ventral striatum and medial caudate (Arikuni and Kubota 1986; Eblen...
METHODS

Two Macaca fascicularis monkeys (A: male, 5.4 kg; B: female, 3.2 kg weight) served for the study. The activity of single neurons was recorded with moveable microelectrodes during performance of a behavioral task while monitoring arm and mouth muscle activity, licking movements, and eye movements. Electrode positions were reconstructed from small electrolytic lesions on 40-μm-thick, cresyl violet-stained histological brain sections. Most methods were similar to those described in detail for the recordings in the striatum, and the present animal A had served also for the study of striatum in the same task (animal B of Hollerman et al. 1998).

Behavioral procedures

Animals were seated in a primate chair and contacted an immovable, touch-sensitive resting key. Visual stimuli of $13 \times 13^\circ$ were presented as instruction or trigger stimuli on a 13-in computer monitor. A small transparent response lever was positioned centrally in a transparent vertical wall in front of the monitor immediately below the position of the visual stimuli. A 1-kHz sound with ~68 dB intensity served as conditioned reinforcer. Small quantities of apple juice (0.15–0.20 ml) delivered by a solenoid valve served as rewards. A closed-circuit video system served to continuously supervise limb movements from above. Animals were fluid- and partly food-deprived to motivate correct task performance and daily cooperation considerably. An auditory cue was presented for 2.0 s while the animal maintained its hand on the resting key, and liquid reward was delivered after a further 1.5 s.

In the computer-controlled delay go-nogo task, the animal kept its right hand relaxed on the resting key and a fractal picture appeared on the screen for 1 s (Fig. 1). It served as an instruction, indicating whether the animal should execute or withhold a movement in response to an upcoming trigger stimulus and whether it would receive a liquid reward or a conditioned auditory reinforcer. Three instruction pictures were used for three trial types, comprising rewarded movement, rewarded nonmovemnt, or unrewarded movement. Thus each instruction served as a preparatory signal that the animal could remember and use for preparing the upcoming reaction (what), whereas the trigger determined the time of the behavioral reaction (when) without providing additional information about the nature of the required reaction. The trigger stimulus consisted of the same red square in each trial type and appeared at a random 1.5–2.5 s after instruction offset. In rewarded-movement trials, the animal released the resting key, touched the lever and received the liquid reward 1.5 s later (Fig. 1, top). The trigger stimulus extinguished on lever touch in correctly performed trials or 1.5 s after lever release in erroneous trials without key release or lever touch. In rewarded-movement trials, a small quantity of liquid reward, and in unrewarded-movement trials the reinforcing sound, were presented at 1.5 s after lever touch. In nonmovement trials (middle), the same trigger stimulus was presented for 2.0 s while the animal maintained its hand on the resting key, and liquid reward was delivered after a further 1.5 s.

FIG. 1. Behavioral task. Monkey sat with its right hand immobile on the immovable resting key and faced a computer monitor positioned behind a transparent wall in which a nearly transparent lever was mounted centrally. Task consisted of 3 trial types alternating semirandomly. All trials began with a 2-s control period during which the monitor was blank, followed by a 1-s presentation of a fractal instruction picture at monitor center immediately above the lever. After a random delay of 2.5–3.5 s after instruction onset, the red square trigger stimulus appeared at the center of the monitor. In rewarded (top)- and unrewarded-movement trials (bottom), the trigger elicited the movement and disappeared when the animal touched the lever after release of the resting key or stayed on for 2.0 s in erroneous trials without key release or lever touch. In rewarded-movement trials, a small quantity of liquid reward, and in unrewarded-movement trials the reinforcing sound, were presented at 1.5 s after lever touch. In nonmovement trials (middle), the same trigger stimulus was presented for 2.0 s while the animal maintained its hand on the resting key, and liquid reward was delivered after a further 1.5 s.

release with an erroneous movement. Unrewarded-movement trials required the same behavioral reaction as rewarded-movement trials, but the liquid drop was replaced by a sound of 100-ms duration (Fig. 1, bottom). The sound served as signal of correct task performance and, as compared with no sound, improved the animals’ correct task performance and daily cooperation considerably. Animals needed to perform this trial type correctly before advancing to a trial reinforced by liquid. To maintain the motivation of the animal, every correct unrewarded-movement trial was followed by one of the two rewarded trial types, thus predicting an upcoming reward. Thus the sound did not constitute an immediate reward but served as reinforcer and predicted a reward in the following trial, thus qualifying it as a secondary reinforcer.

The three trial types alternated semirandomly, with the consecutive occurrence of same trial types being restricted to three rewarded-movement trials, two nonmovement trials, and one unrewarded-movement trial. Thus a movement trial was followed by any trial type with a probability of 0.33, a nonmovement trial was followed by a movement trial type with a probability of 0.75, and an unrewarded-movement trial was followed by a rewarded trial type with a probability of 1.0, as long as trials were performed correctly. Trials lasted 11–13 s, intertrial intervals were 4–7 s. In free-liquid trials, animals received small quantities of liquid without performing in any behavioral task.
and in the absence of phasic stimuli. Intervals between drops were irregular and >11 s.

Data acquisition

After behavioral conditioning, animals were implanted under deep pentobarbital sodium anesthesia and aseptic conditions with two cylinders for head fixation and a stainless steel chamber permitting vertical access with microelectrodes to the left frontal lobe. The dura was left intact. Teflon-coated, multistranded, stainless steel wires were implanted into the right extensor digitorum communis and biceps brachii muscles for electromyographic (EMG) recordings. In animal A, Ag-AgCl electrodes were implanted into the outer, upper, and lower canthi of the orbits for the recording of electrooculograms (EOG). (In animal B, EOGs were recorded with an Iscan infrared oculometer.) The implant was fixed to the skull with stainless steel screws and several layers of dental cement.

Glass-insulated, platinum-plated tungsten microelectrodes stuck inside a metal guide cannula served to record extracellularly the activity of single neurons, using conventional electrophysiological techniques. Histological inspections revealed that the tips of all guide cannulas ended above the most dorsal parts of orbitofrontal cortex. Although guide cannulas damaged more tissue than solid microelectrodes, they permitted use of thin microelectrodes, causing very little damage to the areas investigated. Discharges from neuronal perikarya were converted into standard digital pulses by means of an adjustable Schmitt-trigger. EMGs and horizontal and vertical EOGs were collected during neuronal recordings. EMGs were converted into standard digital pulses by a Schmitt-trigger. Licking movements were recorded as a standard digital pulse when the tongue interrupted an infrared light beam at the liquid spout.

Pulses from neuronal discharges and EMGs were sampled together with digital signals from the behavioral task by a computer, together with analogue signals from EOGs. Only data from neurons sampled by the computer for ≥30 trials using all three trial types are reported. All data from neurons suspected to covary with some task component, and occasionally from unmodulated neurons, were stored uncondensed on computer disks.

Data analysis

Onset, duration, magnitude, and statistical significance of increases of neuronal activity were assessed with a specially implemented sliding window procedure based on the nonparametric one-tailed Wilcoxon signed-rank test (Apicella et al. 1992), using a 2-s control period immediately before the instruction, and a time window of 250 ms that was moved in steps of 25 ms through the period of a suspected change. For activations preceding the instruction, the control period was placed individually for each neuron toward trial end at a position without obvious neuronal changes. Magnitudes of activations were expressed as percentage above control period activity. Peak activity was determined from the 500-ms interval with maximum neuronal activity. Depressions of activity are not reported.

Latencies, durations, and magnitudes of neuronal activations were calculated for blocks of trials and compared among the three trial types using ANOVA with post hoc Fisher’s PLSD test (P < 0.05). Magnitudes of activations were compared between trial blocks with the two-tailed Mann-Whitney U test on the basis of impulse counts in individual trials, normalized for durations of comparisons (P < 0.01). Neuronal activations were considered as preferential for one or two trial types when they were statistically significant (Wilcoxon test), and their magnitudes were significantly higher than in the other trial types (Mann-Whitney U test). These included statistically significant activations occurring selectively in only one or two trial types but not in the other trial types (Wilcoxon test). Activations either preceded or followed individual task events. They were considered to follow a task event when their onset and peak latencies were <500 ms after an event and when their peak activation was closer to the preceding rather than the subsequent event.

We evaluated movement parameters in terms of reaction time (from trigger onset to release of resting key), movement time (from key release to touching the response lever), and return time (from lever touch back to touch of resting key) and compared them using the Kolmogorov-Smirnov test (P < 0.001). We assessed differences in distributions of neuronal activations in orbitofrontal cortex with the χ² test, using four equidistant mediolateral levels and three rostrocaudal levels (sections a-b, c-d and e-f of Fig. 14).

RESULTS

Behavior

Both animals showed >95% correct task performance throughout the experiment (monkey A: 99.0, 99.6, and 98.0%; monkey B: 98.0, 97.0, and 99.6% for rewarded-movement, rewarded-nonmovement, and unrewarded-movement trials, respectively). Unrewarded movements did not lead to immediate reward but were followed by a conditioned auditory reinforcer and a subsequent rewarded trial. Reaction times in both animals were significantly shorter in rewarded as compared with unrewarded-movement trials, although both trials involved reaching from the same starting position toward the same response lever (Table 1). Movement times differed inconsistently. Return times were significantly longer in rewarded as compared with unrewarded-movement trials, as both animals kept pressing the response lever after the reaching movement until the liquid reward was delivered, whereas they immediately returned to the resting key after lever press in unrewarded-movement trials. All movement differences concerned predominantly the timing of movement. Major differences in patterns of arm muscle activity or visible postural differences were not observed in electromyographic and video recordings between rewarded- and unrewarded-movement trials.

Eye movements were very similar in the three trial types and failed to show systematic differences between rewarded and unrewarded movements (Fig. 2). The instruction elicited an ocular saccade to a relatively fixed position on each instruction picture unless the gaze was already there. The trigger stimulus in both movement trials elicited a very similar saccade to the

---

**TABLE 1. Movement parameters in rewarded- and unrewarded-movement trials**

<table>
<thead>
<tr>
<th>Monkey</th>
<th>Reaction Time</th>
<th>Movement Time</th>
<th>Return Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A (ms)</td>
<td>B (ms)</td>
<td>A (ms)</td>
</tr>
<tr>
<td>Rewarded</td>
<td>328 ± 2</td>
<td>295 ± 3</td>
<td>634 ± 4</td>
</tr>
<tr>
<td>Unrewarded</td>
<td>452 ± 8*</td>
<td>434 ± 8*</td>
<td>358 ± 5*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3095 ± 15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>673 ± 19*</td>
</tr>
</tbody>
</table>

All values are given in means ± SE in milliseconds. Values were obtained from 688 trials in monkey A and from 1,060 trials in monkey B. These were measured during neuronal recordings in blocks of maximally 15 trials of the same type. *P < 0.001 against rewarded movements; Kolmogorov-Smirnov test.
response lever. In no cases were differences of neuronal activity between trial types clearly related to differences in eye movements.

Mouth movements were not a part of the task contingencies. Tongue contacts with the spout occurred over relatively long periods in each trial. They began sporadically before and after the instruction, were unrelated to instruction onset, became more frequent during the trigger-reward interval, were reproducible and maximal after reward delivery, and occurred occasionally during the intertrial interval (Figs. 7 and 9). They also occurred sporadically in unrewarded-movement trials.

Neuronal database

A total of 505 orbitofrontal neurons with mean spontaneous discharge rate of 6.3 imp/s (range 0.4–38.5 imp/s) was tested during task performance. Of these, 188 neurons (37%) exhibited 260 statistically significant task-related activations. Three major task relationships were found, namely responses to instructions, activations preceding reinforcers, and responses to reinforcers (Table 2). A few neurons showed activations preceding the instructions or after the trigger stimulus.

Responses to instructions

Instruction responses occurred in 99 of the 188 task-related neurons (54%) (Table 2). Many responses reflected the type of reinforcer. They occurred preferentially in both rewarded trials irrespective of the execution or withholding of movement but not in unrewarded-movement trials (Fig. 3A) or, conversely, only in unrewarded-movement trials (Fig. 3B). Ten neurons responded preferentially in nonmovement trials (Fig. 3C). Only three neurons responded in both movement trial types irrespective of the type of reinforcer. Although some responses lasted for >1 s, only four neurons showed statistically significant sustained activations lasting until trigger onset or beyond (Fig. 3D). Instruction responses in 35 of the 99 neurons occurred unselectively in all three trial types. Responses had mean latencies ranging from 155 to 179 ms and durations of 459–562 ms in the different trial types. Response magnitudes amounted to about fourfold increases of activity (mean magnitudes

TABLE 2. Numbers of orbitofrontal neurons differentially influenced by the type of reinforcement

<table>
<thead>
<tr>
<th>Trial Type</th>
<th>Instruction</th>
<th>Trigger</th>
<th>Reinforcement</th>
<th>Instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Following</td>
<td>Preceding</td>
<td>Following</td>
<td>Preceding</td>
</tr>
<tr>
<td>Rewarded movement</td>
<td>8</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Nonmovement</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Unrewarded movement</td>
<td>22</td>
<td>0</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Reward (irrespective of</td>
<td>20</td>
<td>1</td>
<td>1</td>
<td>41</td>
</tr>
<tr>
<td>movement)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movement (irrespective of</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>reinforcer)</td>
<td>36</td>
<td>2</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Nonpreferential</td>
<td>99 (53)</td>
<td>4 (2)</td>
<td>17 (9)</td>
<td>51 (27)</td>
</tr>
</tbody>
</table>

Total number of task-modulated neurons (n = 188) is inferior to the sum of table entries (n = 260) because of multiple-task relationships. Activations listed under Reward occurred in both rewarded-movement and nonmovement trials. Activations listed under Movement occurred in both rewarded- and unrewarded-movement trials. Trial type with activations preceding instructions refers to the preceding, not the current, trial. Values in parentheses are percentages.
of 286–320% above control activity). None of these parameters varied significantly among the three trial types ($P > 0.05$; ANOVA).

Activations preceding reinforcers

Of the 188 task-related neurons, 51 (27%) showed activations that began well before the liquid reward or the conditioned auditory reinforcer and terminated 0.5–1.0 s after these events (Table 2). Activations in 41 neurons occurred in both liquid-rewarded trial types but not in sound-reinforced trials (Fig. 4A), a few others being restricted to one rewarded trial type. Twenty-one of the 41 neurons responded also to reward delivery. Some, usually weak, activations preceded only the reinforcing sound (Fig. 4B).

Most activations began in the trigger-reinforcer interval, occasionally <1 s before reinforcement (3 neurons; Fig. 5A) but usually earlier (15 neurons; Fig. 5B). Other activations began before the trigger (17 neurons; Fig. 5C). Some activations had rather long time courses, showing sluggish onset times, lasting during major portions of the trial and returning shortly to baseline after reinforcement (6 neurons; Fig. 5D).

Activations remained present until the liquid or sound reinforcer was delivered and subsided immediately afterward, even when these events occurred before or after the usual time (Fig. 6A). Prereward activations occurred also when liquid was delivered at regular intervals in free-liquid trials outside the task, in all 10 neurons tested with task and free reward (Fig. 6B). Prereward activations in 10 neurons adapted rapidly to the last timing of reward relative to the trigger stimulus. The prereward activation in Fig. 6C began earlier after the trigger stimulus when reward had been delivered earlier in a preceding

![Figure 3](http://jn.physiology.org/)  
**FIG. 3.** Different trial selectivities of responses to instruction stimuli of 4 orbitofrontal neurons. A: response in both rewarded trial types but absence of response in unrewarded-movement trials. B: response restricted to unrewarded-movement trials reinforced by a conditioned sound. C: response restricted to nonmovement trials. D: 1 of the rare examples of sustained activation during the instruction-trigger interval, occurring in rewarded-movement trials and, to a lesser extent, in nonmovement trials. Perievent time histograms are composed of neuronal impulses shown as dots below. Each dot denotes the time of a neuronal impulse, and distances to instruction onset correspond to real-time intervals. Each line of dots shows 1 trial. Trials alternated semirandomly during the experiment and are separated according to trial types and rearranged according to instruction-trigger intervals.

![Figure 4](http://jn.physiology.org/)  
**FIG. 4.** Selective activations preceding reinforcers in three orbitofrontal neurons. A: activation preceding the delivery of liquid reward in the 2 rewarded trial types but not before the reinforcing sound in unrewarded-movement trials. B: weak activation preceding the reinforcing sound in unrewarded-movement trials. Trials are rank-ordered according to instruction-reinforcer intervals.
trial block (compare 4th with 2nd trial block after reward had been shifted to an earlier time in the 3rd compared with the 1st block).

Prereward activations occurred during and immediately after the trigger-reward interval during which licking movements were also frequent (Fig. 7). However, the activations were absent during other trial periods and in intertrial periods in which licking movements occurred occasionally. Activations also were absent in unrewarded-movement trials that showed considerable licking activity.

Responses to reinforcers

Of the 188 task-related neurons, 67 (36%) responded to the delivery of a reinforcer (Table 2). Responses in 62 neurons occurred in both liquid-rewarded trial types irrespective of the movement and not in unrewarded-movement trials (Fig. 8, A and B). Twenty-one of the 62 neurons also showed prereward activations. Very few neurons were further selective for rewarded-movement trials. A few responses occurred only after sound reinforcement in unrewarded-movement trials and not after liquid reward. Fourteen of the 67 neurons responded also to the instructions.

Latencies of reward responses in rewarded-movement and nonmovement trials ranged from 70 to 1,590 ms (<100 ms in 25 and 26 neurons, 100–300 ms in 13 and 17 neurons, >300 ms in 24 and 19 neurons in the 2 trial types, respectively; means of 298 and 322 ms). Durations of reward responses in these trials ranged from 120 to 2,320 ms (means of 633 and 651 ms). Response magnitudes amounted to about fivefold increases of activity (means of 381 and 418% above control activity). None of these parameters varied significantly among the two rewarded trial types ($P > 0.05$; ANOVA).

We delivered reward earlier or later than usually to further characterize the responses. Responses in all nine neurons tested followed the reward to the new time (Fig. 9A) and were increased in magnitude in four of them. Thus responses occurred to the reward and were not delayed trigger responses.

Reward responses were restricted to the period after reward...
delivery, although licking movements also occurred before reward delivery and in unrewarded-movement trials (Fig. 9, B and C). Thus reward responses appeared to be unrelated to mouth movements.

The relationship of reward responses to the solenoid noise associated with reward delivery was tested in 14 responding neurons by blocking the liquid tube while maintaining the solenoid noise in free-liquid trials. Eight of these neurons failed to respond to the solenoid noise alone, suggesting a true reward response (Fig. 10), whereas the other six neurons also responded without the liquid.

Responses to free-liquid versus task reward

A total of 76 neurons responded to reward in the behavioral task, in free-liquid trials or in both situations. Of these, 46
neurons responded to the liquid during task performance in both rewarded trial types and in free-liquid trials in the absence of any specific task (Fig. 11A). Three neurons responded to liquid in the task but not in free-liquid trials (Fig. 11B). By contrast, 27 neurons failed to respond to liquid in the task but were activated in free-liquid trials (Fig. 11C). Sixteen of them showed instruction responses in both rewarded trials (Fig. 11D).

**Activations preceding instructions**

Some neurons showed an interesting type of activation that was partly also related to reinforcement. Of the 188 task-related neurons, 22 (12%) showed activations which began slowly and at varying times after the reinforcer of the preceding trial, showed their peak <500 ms before the instruction and terminated abruptly afterward (Table 2). According to their sluggish onset, they appeared to precede the upcoming instruction rather than after the past reinforcer. Activations in 6 of the 22 neurons occurred preferentially after both rewarded trial types and not after unrewarded-movement trials, whereas in 2 neurons they occurred preferentially after unrewarded trials (Fig. 12).

**Population activity of major reinforcement-related activations**

The histograms of Fig. 13 display averaged activity from neurons showing responses to instructions (A), activations

![Image](https://example.com/image1)

**Fig. 10.** Influence of interruption of liquid reward flow on reward response in an orbitofrontal neuron. Response was lost in the absence of reward delivery, suggesting a true response to reward rather than to the associated noise of the solenoid liquid valve, which opened audibly when reward liquid was delivered (“liquid with solenoid noise”). For “solenoid noise only,” the tube between solenoid and animal’s mouth was closed, and the solenoid noise occurred alone without delivering liquid. Data are from free-liquid trials, ordered chronologically from top to bottom.

![Image](https://example.com/image2)

**Fig. 11.** Relationship of reward responses to task performance in 4 orbitofrontal neurons. A: reward response in both rewarded trial types in behavioral-task and free-liquid trials without any behavioral task. B: reward response in behavioral-task but not in free-liquid trials. C: reward response occurring in free-liquid trials but not in behavioral task. D: response to reward in free-liquid trials, and response to instruction but not to liquid in behavioral task. Baseline activity was increased in free-liquid trials in C and D. Task trials alternated semirandomly during the experiment and are separated according to trial types and rearranged according to instruction-trigger intervals. None of the neurons in A–D responded in unrewarded-movement trials. Free-liquid trials were run in separate blocks.
preceding liquid reward (B), and responses after liquid reward (C) in both rewarded trial types. Note that averaging of activity over long task periods reduces temporally disperse activations peaks. Therefore population responses appear lower than averages of individual activations.

Positions of neurons

Histological reconstructions showed that rostral area 13, entire area 11, and lateral area 14 of orbitofrontal cortex were explored (Fig. 14). Neurons with instruction responses were distributed widely, being significantly more frequent in medial as compared with lateral parts of the explored region \( (P < 0.05) \). Neurons with prereward activations were found predominantly in rostral area 13, where they were significantly less frequent in its very anterior part \( (P < 0.001) \). Neurons responding to reinforcers were significantly more frequent in lateral than medial orbitofrontal cortex \( (P = 0.01) \).

DISCUSSION

These data show that neurons in orbitofrontal cortex process rewards in three principal forms in a delayed go-nogo task, as transient responses to reinforcer-predicting instructions, sustained activations preceding reward, and transient responses after reward. A few neurons were activated before the initial instruction signal in relation to the reward situation in the preceding or expected upcoming trial. In contrast to other prefrontal areas, few orbitofrontal neurons showed activations related to behavioral reactions in this task. These data support the notion that orbitofrontal cortex constitutes an important component of reward circuits in the brain.

Processing of reinforcement information

Processes tested by the behavioral task. Delayed response tasks typically assess the functions of prefrontal cortex in the temporal organization of goal-directed behavior, working memory, and preparation of responding (Bauer and Fuster 1976; Fuster 1973; Jacobsen and Nissen 1937; Kubota et al. 1974; Niki et al. 1972; Rosenkilde et al. 1981). Go-nogo tasks test the inhibition of overt behavioral responses and are deficient after orbitofrontal lesions (Iversen and Mishkin 1970). Performance in these tasks depends on reinforcement and thus makes them suitable for investigating the role of reinforcement in goal-directed behavior. To compare primary reward with conditioned reinforcement, we added a trial type to the standard delayed go-nogo task in which movement was reinforced with a conditioned tone instead of liquid. To differentiate movement preparation from reinforcer expectation, we introduced a second delay that separated the behavioral response from the reinforcer.

Responses to reward-predicting environmental stimuli. According to animal learning theory (Dickinson 1980), the instructions in our task were associated with specific reinforcers through a Pavlovian procedure and had an occasion-setting function determining the movement or nonmovement reaction. Most instruction responses differentiated between liquid and sound, but very few responses differentiated between the behavioral reactions irrespective of the type of reinforcer. Thus orbitofrontal neurons reported environmental stimuli more in
association with reinforcement than behavioral reaction. These instruction responses occurred in orbitofrontal areas influenced by the medial temporal cortex (Morecraft et al. 1992).

Our approach was based on experiments in which neurons in dorsolateral prefrontal cortex discriminated between instruction stimuli predicting liquid reward versus no reinforcement in a less complex task (Watanabe 1990, 1992). Preceding work had shown that orbitofrontal neurons discriminate between appetitive and aversive visual stimuli (Thorpe et al. 1983).

The presently observed reward relationships in both rewarded trial types would argue against relationships to visual stimulus features of these instruction responses. In addition, the adjoining report demonstrates that reinforcement-related trial selectivities were maintained when novel visual instructions were learned despite considerable differences in visual features (Tremblay and Schultz 2000). Also, selectivities in orbitofrontal neurons remained related to rewards when multiple instruction sets were used in a spatial delayed response task (Tremblay and Schultz 1999). Thus the trial selectivities were more likely due to differences in reinforcement than visual features.

**EXPECTATION OF REWARD.** Sustained activations preceding reinforcement occurred mostly in trials rewarded with liquid irrespective of the behavioral reaction and were largely absent in trials reinforced by the sound. This suggests a relationship to reward and not to the end of trial message contained in the reinforcers. The activations began typically around the time of the trigger stimulus as the last signal preceding reward and terminated immediately after reward was delivered, irrespective of its time of occurrence. They apparently reflected the expectation of reward by coding the occurrence of reward but not its precise moment. These prereward activations occurred in orbitofrontal areas influenced by the medial temporal cortex (Morecraft et al. 1992).

The expectation of reward evoked by a conditioned appetitive stimulus is a major component of the central motivational state underlying approach behavior (Bindra 1968; Dickinson 1980). Although the instruction stimuli in the present task are associated with reward and have occasion-setting properties in instrumental behavior, the trigger would have a better reward-predicting property because of temporal proximity. In line with this reasoning, most sustained activations followed the trigger rather than preceded it.

The present differential prereinforcement activations resembled activities discriminating between expected appetitive and aversive reinforcers in rat orbitofrontal cortex (Schoenbaum et al. 1998). They were somewhat more variable than reward expectation-related activations in primate striatum (Apicella et al. 1992; Hikosaka et al. 1989; Hollerman et al. 1998; Schultz et al. 1992; Shidara et al. 1998).

**REWARD RESPONSES.** Many orbitofrontal neurons detected the delivery of liquid reward in both rewarded trials irrespective of the behavioral reaction, whereas only few neurons responded to sound reinforcement. Most reward-driven neurons also responded to liquid outside the task, suggesting a relationship to the primary appetitive event and not to a particular reinforcing function or an end of trial signal. These reward responses occurred in orbitofrontal areas influenced by the amygdala (Morecraft et al. 1992). Earlier studies reported similar orbitofrontal responses to liquid reward (Niki et al. 1972; Rosenkilde et al. 1981), which discriminated against aversive liquids (Thorpe et al. 1983), whereas neurons in more caudal parts of area 13 responded to gustatory and olfactory stimuli (Rolls et al. 1990, 1996; Schoenbaum and Eichenbaum 1995; Thorpe et al. 1983). Reward responses also were found in dorsolateral prefrontal cortex (Watanabe 1989) and striatum (Apicella et al. 1991; Bowman et al. 1996; Hikosaka et al. 1989; Shidara et al. 1998).

Some orbitofrontal neurons only responded to liquid outside the task. This may be because of the fact that the liquid was not predicted by any phasic stimulus. More than half of these neurons responded to the instruction during task performance as also reported by others (Matsumoto et al. 1995). Apparently
the response was transferred to the earliest liquid-predicting stimulus as in midbrain dopamine neurons (Mirenowicz and Schultz 1994). The unpredictable occurrence of reinforcement is a necessary condition for acquiring new behavioral responses (Rescorla and Wagner 1972). By contrast, the detection of fully predicted reward is necessary to prevent extinction of established behavior. Thus the orbitofrontal responses to unpredicted reward may play a role in reward-directed learning, whereas the responses to predicted reward may function in maintaining established task performance.

EXPECTATION OF INSTRUCTION. Preinstruction activations reflected the expectation of instructions acquired from the experience in the task schedules. Previous studies reported preinstruction activations in striatal and cortical neurons that were unconditional on trial type (Apicella et al. 1992), changed with regularly alternating trial types (Hikosaka et al. 1989), or reflected the employed dimensions in discriminations (Sakagami and Niki 1994). The present preinstruction activations apparently were related to the possible type of upcoming trial. As correct unrewarded trials were invariably followed by a rewarded trial type, activations preferentially following unrewarded trials may reflect the expectation of a rewarded trial. By contrast, as rewarded trials could follow each other in our asymmetric trial schedule, it is less certain which kind of expectation was reflected by activations occurring preferentially after rewarded trials.

DELAY ACTIVITY. Sustained activations of dorsolateral prefrontal neurons during the instruction-trigger delay probably reflect working memory or movement preparation (Funahashi et al. 1993). Sustained activations in orbitofrontal neurons occurred rarely in the instruction-trigger delay in our task. This contrasted sharply with the frequent occurrence of sustained delay activity in the striatum in an identical conditional delayed go-nogo task (Hollerman et al. 1998) and in spatial delayed response tasks in dorsolateral prefrontal cortex (cf. Funahashi et al. 1993). Sustained activations were presently frequent in the second, trigger-reward delay, where they may reflect the expectation of reward. An earlier spatial delayed response task used only the initial instruction-trigger delay and reported sustained delay activity in 25% of tested orbitofrontal neurons (Rosenkilde et al. 1981). As that delay ended close to the reward, some of the activations might reflect the expectation of reward. More sustained mnemonic and movement preparatory activity conceivably may occur in orbitofrontal neurons in behavioral tasks involving more elaborate memory demands and behavioral reactions.

Comparison with other reward-processing brain systems

The prominent relationships to reinforcingers would allow the orbitofrontal cortex to be a major component of the reward system of the brain. A comparison with reward signals in closely related brain structures may help to assess the potential contributions of orbitofrontal activities to the motivational control of goal-directed behavior.

STRIATUM. Orbitofrontal neurons appear to process reward information in many similar ways as neurons in caudate nucleus, putamen, and ventral striatum. Striatal neurons are activated during the expectation of reward, respond to reward delivery, and discriminate between primary appetitive liquid and sound reinforcement (Apicella et al. 1991, 1992; Hikosaka et al. 1989; Hollerman et al. 1998; Schultz et al. 1992). However, striatal neurons show a larger variety of behavioral relationships than orbitofrontal neurons, including the expectation of external stimuli and the preparation, initiation and execution of movement (cf. Schultz et al. 1995). Many of these activities depend on the expectation of reward as opposed to secondary reinforcement (Hollerman et al. 1998). The similarity between orbitofrontal and striatal reward-related activations may suggest that orbitofrontal inputs induce the striatal reward signals. Many striatal reward-related activations occur in areas with heavy orbitofrontal projections, in particular the ventral striatum (Apicella et al. 1991, 1992; Arikuni and Kubota 1986; Bowman et al. 1996; Ebben and Graybiel 1995; Haber et al. 1996; Schultz et al. 1992; Selemon and Goldman-Rakic 1985; Shidara et al. 1998), although they also are found in more dorsal striatal regions with fewer orbitofrontal inputs (Hikosaka et al. 1989; Yeterian and Pandya 1991).

AMYGDALA. Neurons in different nuclei of amygdala respond selectively to primary foods and liquids and to conditioned stimuli associated with rewards (Nishijo et al. 1988). Amygdala neurons show sustained activations preceding behavioral reactions in a delayed response task (Nakamura et al. 1992). Without an interval between behavioral reaction and reward, some of these activations might reflect an expectation of reward, which was confirmed in rats (Schoenbaum et al. 1998).

DOPAMINE NEURONS. Dopamine neurons show entirely different forms of reward processing. They show phasic, but not sustained, activations after unpredicted rewards and conditioned, reward-predicting stimuli, and they are depressed when a predicted reward is omitted (Ljungberg et al. 1992; Mirenowicz and Schultz 1994; Romo and Schultz 1990; Schultz et al. 1993). Dopamine responses appear to report the discrepancy between an expected and an actually occurring reward (Schultz et al. 1997) and thus have the formal characteristics of reinforcement signals for acquiring new behavioral reactions (Rescorla and Wagner 1972).

We thank B. Aebischer, J. Corpataux, A. Gaillard, A. Pisani, A. Schwarz, and F. Tinguely for expert technical assistance.
The study was supported by Swiss National Science Foundation Grants 31-28591.90, 31.4331.95, and NFP38.4038-43997. L. Tremblay received a postdoctoral fellowship from the Fondation pour la Recherche Scientifique of Quebec.

Present address of L. Tremblay: INSERM Unit 289, Hôpital de la Salpetrière, 47 Boulevard de l’Hôpital, F-75651 Paris, France.

Address reprint requests to W. Schultz.

Received 18 February 1999; accepted in final form 29 November 1999.

REFERENCES


