

Reward Processing in Primate Orbitofrontal Cortex and Basal Ganglia

Wolfram Schultz, Léon Tremblay¹ and Jeffrey R. Hollerman²

Institute of Physiology and Program in Neuroscience,
University of Fribourg, CH-1700 Fribourg, Switzerland

¹Present address: INSERM Unit 289, Hôpital de la Salpêtrière,
47 Boulevard de l'Hôpital, F-75651 Paris, France

²Present address: Department of Psychology, Allegheny
College, Meadville, PA 16335, USA

This article reviews and interprets neuronal activities related to the expectation and delivery of reward in the primate orbitofrontal cortex, in comparison with slowly discharging neurons in the striatum (caudate, putamen and ventral striatum, including nucleus accumbens) and midbrain dopamine neurons. Orbitofrontal neurons showed three principal forms of reward-related activity during the performance of delayed response tasks, namely responses to reward-predicting instructions, activations during the expectation period immediately preceding reward and responses following reward. These activations discriminated between different rewards, often on the basis of the animals' preferences. Neurons in the striatum were also activated in relation to the expectation and detection of reward but in addition showed activities related to the preparation, initiation and execution of movements which reflected the expected reward. Dopamine neurons responded to rewards and reward-predicting stimuli, and coded an error in the prediction of reward. Thus, the investigated cortical and basal ganglia structures showed multiple, heterogeneous, partly simultaneous activations which were related to specific aspects of rewards. These activations may represent the neuronal substrates of rewards during learning and established behavioral performance. The processing of reward expectations suggests an access to central representations of rewards which may be used for the neuronal control of goal-directed behavior.

Accumulating evidence suggests that the orbital part of the prefrontal cortex is crucially involved in the motivational control of goal-directed behavior (Damasio, 1994; Rolls, 1996). Patients with lesions of the orbitofrontal cortex are impaired in making decisions about the expected outcome of actions (Bechara *et al.*, 1998). As rewards may constitute basic goals of behavior (Dickinson and Balleine, 1994), the motivational functions of the orbitofrontal cortex may be partly based on the processing of reward information. Monkeys with orbitofrontal lesions respond abnormally to changes in reward contingencies (Iversen and Mishkin, 1970; Dias *et al.*, 1996) and show altered reward preferences (Baylis and Gaffan, 1991). Orbitofrontal neurons respond selectively to rewards or aversive stimuli (Thorpe *et al.*, 1983) and process the relative preference for rewards (Tremblay and Schultz, 1999). The motivational functions of the orbitofrontal cortex may involve the closely connected basal ganglia via fronto-striatal projections (Selemon and Goldman-Rakic, 1985; Haber *et al.*, 1990, 1995; Eblen and Graybiel, 1995). Human diseases such as Parkinsonism and chorea involve the striatum and substantia nigra and lead to deficits in voluntary behavior. According to psychopharmacological, lesioning and electrical self-stimulation studies, dopamine inputs to the striatum are involved in approach behavior, evaluation of rewarding outcomes, reward-related learning and the action of addictive substances such as cocaine and heroin (Fibiger and Phillips, 1986; Wise and Rompre, 1989; Robbins and Everitt, 1996; Wise, 1996).

We were interested in studying the neuronal processing of reward information to advance the understanding of the neuronal basis of voluntary, goal-directed behavior. After investigating the responses of dopamine neurons to primary rewards and reward-predicting stimuli (Schultz, 1998), we aimed to address the question of how information about rewards is integrated into neuronal processes underlying reward-directed behavioral actions. We focused on the orbitofrontal cortex and basal ganglia as presumed key structures for the processing of reward information and investigated how rewards expected at trial end influenced neuronal activity related to various components of delayed response tasks which test some of the typical functions of these structures (Jacobsen and Nissen, 1937; Divac *et al.*, 1967). This article presents a summary of our recently published results and interprets the different reward-related activities found in the orbitofrontal cortex, striatum and dopamine neurons. A summary figure presented at the end of this review compares the different forms of reward-related activity observed (Fig. 9).

Neurophysiological Methods

The activity of single neurons was recorded from extracellular positions with moveable microelectrodes in *Macaca fascicularis* monkeys performing behavioral tasks under computer control. The behavioral apparatus and most recording and evaluation techniques were similar to those described previously (Hollerman *et al.*, 1998). Discharges from single neurons were conventionally amplified, filtered, displayed on oscilloscopes, passed through an adjustable Schmitt-trigger and sampled at 2 kHz by a computer, together with markers from events in the behavioral task, such as visual stimuli, key presses, reward delivery and licks of the tongue detected by an infrared device. Neuronal activities were evaluated after the experiment in relation to behavioral events with statistical procedures based on the Wilcoxon test ($P < 0.01$) (Ljungberg *et al.*, 1992). Magnitudes of neuronal activations were compared between two trial types with the Mann-Whitney *U*-test ($P < 0.01$). Neuronal recording sites were marked towards the end of the experiment with small electrolytic lesions by passing negative currents through the microelectrode. Positions of neurons were reconstructed from 40-mm-thick coronal brain sections stained with cresyl violet or immunoreacted for tyrosine hydroxylase. Electromyograms from the extensor digitorum communis, arm biceps and masseter muscles, and horizontal and vertical electrooculograms were recorded through chronically implanted electrodes together with neuronal activity.

Results

Reward Coding in the Orbitofrontal Cortex

Behavioral Tasks

A modified delayed go/no-go task employed with two monkeys allowed us to compare neuronal processing in liquid-rewarded versus unrewarded trials (Fig. 1, top) (Tremblay and Schultz, 2000a). When the animal kept its right hand relaxed on a resting

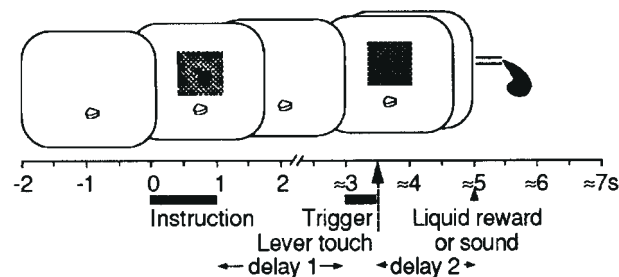
key, one of three colored, fractal instruction pictures appeared on a computer monitor for 1.0 s and indicated one of three trial types, namely rewarded movement, rewarded non-movement and unrewarded movement. A red square trigger stimulus presented at 2.5–3.5 s after instruction onset required the animal to execute or withhold a reaching movement according to the trial type. The trigger stimulus was the same in all three trial types. In rewarded movement trials, the animal released the resting key and touched a small lever below the trigger to receive a small quantity of apple juice after a delay of 1.5 s. In rewarded non-movement trials, the animal remained motionless on the resting key for 2.0 s and received the same liquid reward after a further 1.5 s. In unrewarded movement trials, the animal reacted as in rewarded movement trials, but correct performance was followed not by liquid reward but by a 1 kHz sound after 1.5 s. One of the two rewarded trial types followed each correctly performed unrewarded trial. Animals were required to perform the unrewarded movement trial, as incorrect performance resulted in trial repetition. The sound constituted a conditioned auditory reinforcer as it ameliorated task performance and predicted a rewarded trial, but it was not an explicit reward, hence the simplifying term ‘unrewarded’ movement. Thus, each instruction was the unique stimulus in each trial indicating the behavioral reaction to be performed following the trigger (execution or withholding of movement) and predicting the type of reinforcer (liquid or sound). Each trial contained two delay periods, namely the instruction–trigger delay, during which the animal remembered the type of instruction and prepared for the behavioral reaction (delay 1 in Fig. 1), and the trigger–reinforcer delay, during which the animal could expect the reinforcer (delay 2). Trials lasted 12–14 s, intertrial intervals were 4–6 s.

A modified spatial delayed response task employed with two monkeys permitted us to compare neuronal activity between different liquid or food rewards (Fig. 1, bottom) (Tremblay and Schultz, 1999). When the animal kept its hand on the resting key, one of two instruction pictures appeared for 1.0 s on the computer monitor at a fixed, left or right position. Each instruction picture contained two types of information: (i) its position indicated the target of arm movement performed in reaction to a trigger stimulus which appeared at 2.5–3.5 s after instruction onset; and (ii) its visual aspects indicated which one of two liquid or food rewards would be delivered for correct behavior at trial end. The trigger stimulus consisted of two identical red squares covering exactly both the left and right positions of the instruction stimuli. Thus, the trigger indicated the time of responding without indicating the reward or the movement target. Touching the lever on the side previously indicated by the instruction resulted 1.5 or 2.0 s later in delivery of the corresponding liquid or food reward. The behavioral task differed in two aspects from conventional spatial delayed response tasks: (i) the instruction predicted the reward in addition to indicating the spatial position of the future movement; and (ii) each trial contained two delays to separate the working memory and movement preparation period (delay 1 in Fig. 1) from the reward expectation period (delay 2).

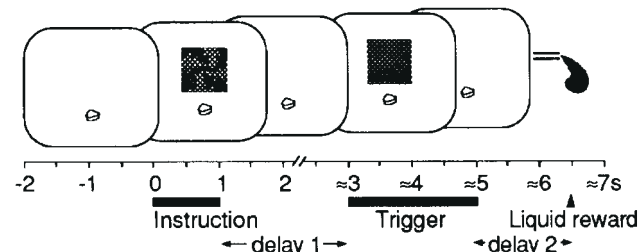
In the delay task each instruction picture indicated specifically one reward at trial onset. We used five different pairs of instruction pictures in liquid trials and one set of three pictures in food trials to assess the influence of visual features on neuronal responses. Only two instruction pictures with their associated two liquid or two food rewards were presented in a given block of trials. Liquid rewards were grenadine and apple juice in one animal, and orange and grape juice in a second

Delayed go-nogo task

Rewarded or unrewarded movement



Rewarded nonmovement



Spatial delayed response task

Left or right movement

Reward 1 or reward 2

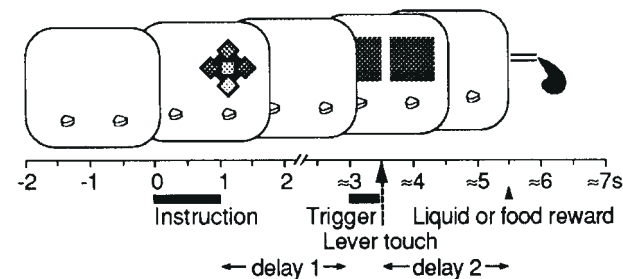


Figure 1. Behavioral tasks. The monkey sat with its right hand immobile on an immovable resting key and faced a computer monitor positioned behind a transparent wall in which a nearly transparent lever was mounted centrally. Top: delayed go/no-go task. The task consisted of three trial types alternating semi-randomly, namely movement trials rewarded by liquid, movement trials reinforced by a conditioned sound and nonmovement trials rewarded by liquid. 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 following instruction onset (delay 1), the red square trigger stimulus appeared at the center of the monitor. In rewarded and unrewarded movement trials, 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 following lever touch (delay 2). In nonmovement trials, 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. Bottom: spatial delayed response task. An initial instruction picture indicated the left or right movement target and the liquid or food reward delivered at trial end. After 2.5–3.5 s (delay 1) two identical squares appeared and elicited the movement from the resting key to the left or right target lever indicated by the instruction. Correct performance was rewarded after 1.5 or 2.0 s (delay 2) with a drop of liquid or a morsel of food. Only two liquid or two food rewards were used in a given trial block (reward 1 and 2).

animal. Food rewards became available in a box located to the right of the monitor following computer-controlled opening of its door (40 × 40 mm frontal aperture). Foods were raisins (most preferred), small apple morsels (intermediate preferred)

and sugar-honey cereal (least preferred). Rewards and target positions alternated randomly in each block of trials, with maximally three consecutive identical trials.

Reward preferences were assessed in separate blocks of choice trials in the spatial delay task before or after recording from each neuron. In these trials two different instructions for two rewards appeared simultaneously at randomly alternating left and right target positions, allowing the animal to touch the lever of its choice following the trigger stimulus.

Types of Task-related Activity

Orbitofrontal neurons in area 11, rostral area 13 and lateral area 14 displayed only three major types of event relationships in the present experiments, although the employed delayed tasks comprised many more behavioral events. These three types consisted of responses to instructions, activations preceding reinforcers and responses to reinforcers (Fig. 9, top). Neurons with instruction responses were widely distributed, being significantly more frequent in the medial than in the lateral parts of the explored region ($P < 0.05$; chi-square test). Neurons with activations preceding reward were predominantly found in rostral area 13 and were significantly more frequent in the posterior than the anterior orbitofrontal cortex ($P < 0.001$). Neurons responding to reinforcers were significantly more frequent in the lateral than the medial orbitofrontal cortex ($P < 0.01$). In addition, a few neurons showed activations preceding the instructions, during the entire instruction–trigger delay or following the trigger stimulus. Nearly all activations in the orbitofrontal cortex showed pronounced relationships to rewards. This was true for the responses to instructions, activations preceding reinforcers and responses to reinforcers which differed between rewarded and unrewarded trials in the go/no-go task and distinguished between different rewards in the spatial task.

Reward Versus No Reward

In the delayed go/no-go task we tested how task-related activity in 188 of 505 tested neurons (37%) differed between rewarded and unrewarded trials (Tremblay and Schultz, 2000a). Rewarded nonmovement trials served as controls for movement relationships.

Instruction responses occurred in 99 of 188 task-related neurons (54%). Two-thirds of the 99 responses reflected the type of reinforcer in occurring preferentially in rewarded movement trials, rewarded nonmovement trials or both rewarded trial types irrespective of the execution or withholding of movement, but not in unrewarded movement trials (38 neurons; Fig. 2A). Conversely, 22 responses occurred preferentially in unrewarded movement trials but not in any rewarded trial type. Only three neurons responded preferentially in both types of movement trial irrespective of the type of reinforcer. Instruction responses in 35 of the 99 neurons occurred unselectively in all three trial types.

Activations preceding the reinforcer occurred in 51 of the 188 task-related neurons (27%). These activations began well before the liquid reward or the conditioned auditory reinforcer, during the trigger–reinforcer interval or even earlier, and terminated within 0.5–1.0 s after the reward or auditory reinforcer was delivered (Fig. 2B). Activations remained present until the liquid or sound reinforcer was delivered and subsided immediately afterwards, even when these events occurred before or after the usual time. Thus, delayed reinforcement prolonged the activations and earlier reinforcement terminated them. These activations were time-locked to the reinforcer and not the in-

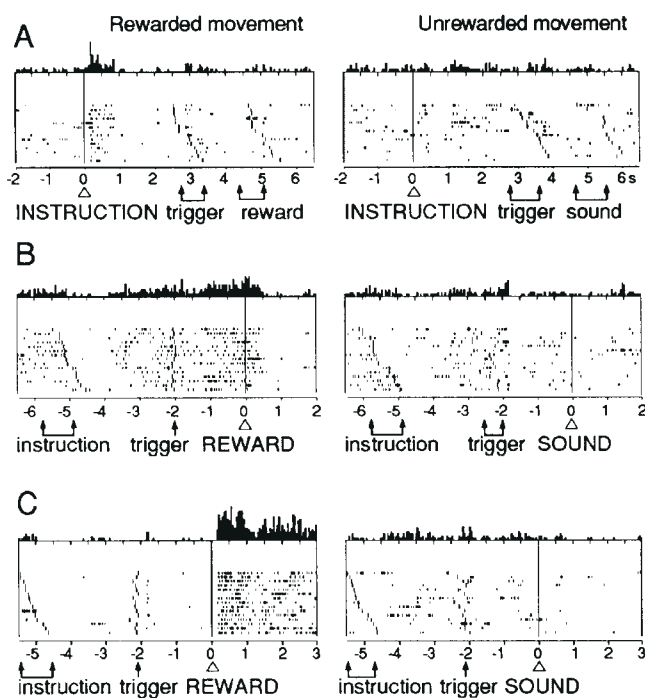


Figure 2. Activations of three orbitofrontal neurons differentiating between rewarded and unrewarded movement trials during performance in the delayed go/no-go task. (A) Instruction response. (B) Activation preceding reward. (C) Reward response. Perievent time histograms are composed of neuronal impulses shown as dots below them. Each dot denotes the time of a neuronal impulse, and distances to instruction onset (A) or reward onset (B, C) correspond to real-time intervals. Each line of dots shows one trial. Trials shown to the left and right in (A–C) alternated semi-randomly during the experiment. They are separated for analysis according to trial types and rearranged according to instruction–trigger (A) or instruction–reinforcer (B, C) intervals. Non-movement trials are not shown.

struction or trigger stimuli, thus being related to the upcoming reward rather than constituting delayed responses to the preceding events. All of these activations reflected the type of reinforcer, occurring mostly in both liquid-rewarded trial types but not in sound-reinforced trials (41 neurons), or only in one of the rewarded trial types (six neurons). Only four activations preceded preferentially the conditioned sound in unrewarded movement trials.

Responses following the delivery of the reinforcer were found in 67 of the 188 task-related neurons (36%) (Fig. 2C). These responses were unlikely to be related to mouth movements, as licking movements occurred also at other task periods during which these neuronal responses were not observed. All of these responses reflected the type of reinforcer, occurring mostly in both liquid-rewarded trial types but not in sound-reinforced trials (62 neurons), or only in rewarded movement trials (two neurons). Only three responses occurred preferentially after sound reinforcement in unrewarded movement trials.

Different Rewards

In the spatial delayed response task we tested how task-related orbitofrontal neurons discriminated between liquid rewards (294 of 1095 neurons, 27%) and between food rewards (138 of 329 neurons, 42%) (Tremblay and Schultz, 1999).

All three principal types of orbitofrontal activations discriminated between liquid rewards and between food rewards (Fig. 3). Activations occurred exclusively or were significantly higher for one liquid reward than the other liquid rewards in

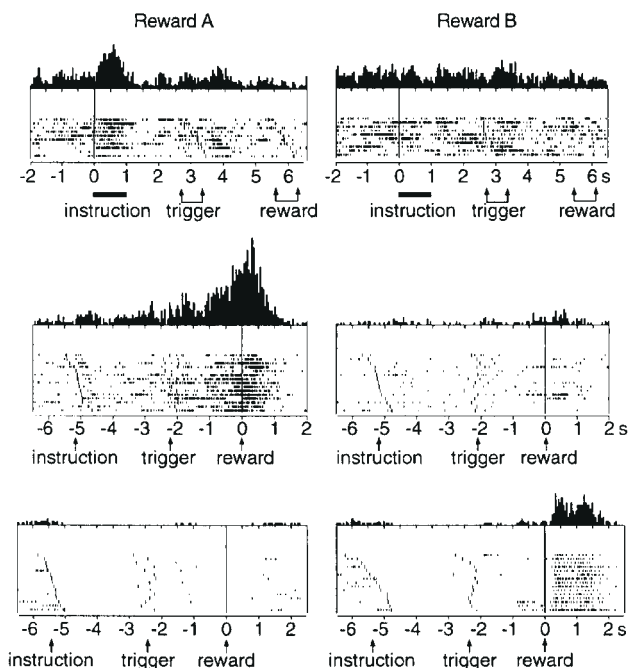


Figure 3. Reward-discriminating activity in the three principal types of activity of orbitofrontal neurons during performance in the spatial delayed response task. Top: response to instruction. Middle: activity preceding liquid reward. Bottom: response to reward. Activity occurred only in trials using reward A (orange juice) but not reward B (top and middle), or only with reward B (grape juice) but not reward A (bottom). Data shown at the top, middle and bottom are from three different neurons, respectively. Trials with reward A and B alternated randomly during the experiment. They are separated for analysis according to trial types and rank-ordered according to instruction–reward intervals.

150 of 218 instruction responses (69%), 65 of 160 activations preceding reward (41%) and 76 of 146 reward responses (52%). These differences were unrelated to eye or licking movements before or after reward delivery, which varied inconspicuously between trials using the two rewards. By contrast, only seven of 218 instruction responses (3%) discriminated between left and right movement targets, and none of the 50 neurons tested with several instruction sets showed significantly different responses to different instructions indicating the same reward. Thus about one-half of reward-related activations of orbitofrontal neurons discriminated between different liquid and food rewards. The instruction responses in orbitofrontal neurons reflected much better the predictable rewards than the spatial or visual features of the instructions in the present task situation.

Reward Preference

We used different food rewards in the spatial delayed response task to test the relationships of 65 reward-discriminating task-related neurons to the animals' preferences of rewards (Tremblay and Schultz, 1999).

We assessed reward preferences behaviorally in a choice version of the spatial delayed response task in which two instruction pictures instead of one were shown simultaneously, above the left and right levers, respectively. Each picture was associated with a different reward. Following the trigger stimulus, animals chose one of the rewards by touching the corresponding lever. We used three food rewards (A–C) but only presented two of them in a given trial block (A–B, B–C or A–C). Animals showed clear preferences in every comparison, choosing reward

A over B, B over C and A over C in 90–100% of trials, independent of the side of lever touch. Thus reward B was chosen less frequently when reward A was available but more frequently when reward C was available in a given trial block. Apparently the preference for reward B was relative and depended on the reward with which it was being compared.

Neuronal activity was studied in the standard delay task with single instruction pictures indicating a single, preferred or non preferred food reward. Two rewards alternated randomly in each trial block. The activity of 40 of 65 reward-discriminating orbitofrontal neurons did indeed depend on the relative preference of food reward. Activations were significantly higher with the relatively more preferred reward as compared with the less preferred reward, no matter which of the different reward combinations was used (Fig. 4). Or, neuronal activations with an intermediately preferred reward were significantly higher when this reward was the relatively preferred one as compared with trials in which this reward was the nonpreferred one (reward B in Fig. 4, top versus bottom). Neuronal activity reflected the preferred reward with 12 of 28 instruction responses, eight of 20 pre-reward activations and seven of 17 reward responses. Conversely, activations reflecting less preferred rewards were observed with six of 28 instruction responses, five of 20 pre-reward activations and two of 17 reward responses. Thus, the reward discrimination occurred in some orbitofrontal neurons on the basis of relative preference rather than physical properties.

Reward Unpredictability

We studied neuronal responses to liquid reward in the delayed go/no-go task in which the reward was predicted by the task stimuli and compared them with responses in free liquid trials in which the same reward was delivered at unpredictable times outside of any behavioral task (Tremblay and Schultz, 2000a).

A total of 76 neurons responding to reward were tested in both the go/no-go task and free liquid trials. Of these, 27 neurons were activated in free liquid trials but failed to respond to liquid in the task (Fig. 5). Of the remaining neurons, 46 responded indiscriminately to the liquid during task performance and in free liquid trials, and three responded only in the task but not in the free liquid trials. Thus reward responses in some orbitofrontal neurons depended on the temporal unpredictability of reward.

Reward-related Activity in the Striatum

We tested neuronal processing in rewarded versus unrewarded trials in slowly discharging neurons in the caudate nucleus, putamen and ventral striatum including the nucleus accumbens with two monkeys, using the same delayed go/no-go task as for the study of the orbitofrontal cortex (1487 neurons in the striatum; Fig. 1, top) (Hollerman *et al.*, 1998). The investigation followed earlier studies in which we tested in two monkeys neuronal activity in temporal relation to reward in a symmetrically reinforced delayed go/no-go task without unrewarded movement trials [1173 neurons in the caudate and putamen (Apicella *et al.*, 1991, 1992); 420 neurons in the ventral striatum (Schultz *et al.*, 1992)]. Slowly discharging neurons comprise 95% of the striatal neuronal population. Statistically significant task-related activations occurred in 991 task-related neurons of the total of 3080 tested neurons (32%).

Types of Task-related Activity

Striatal neurons displayed a considerable variety of task relation-

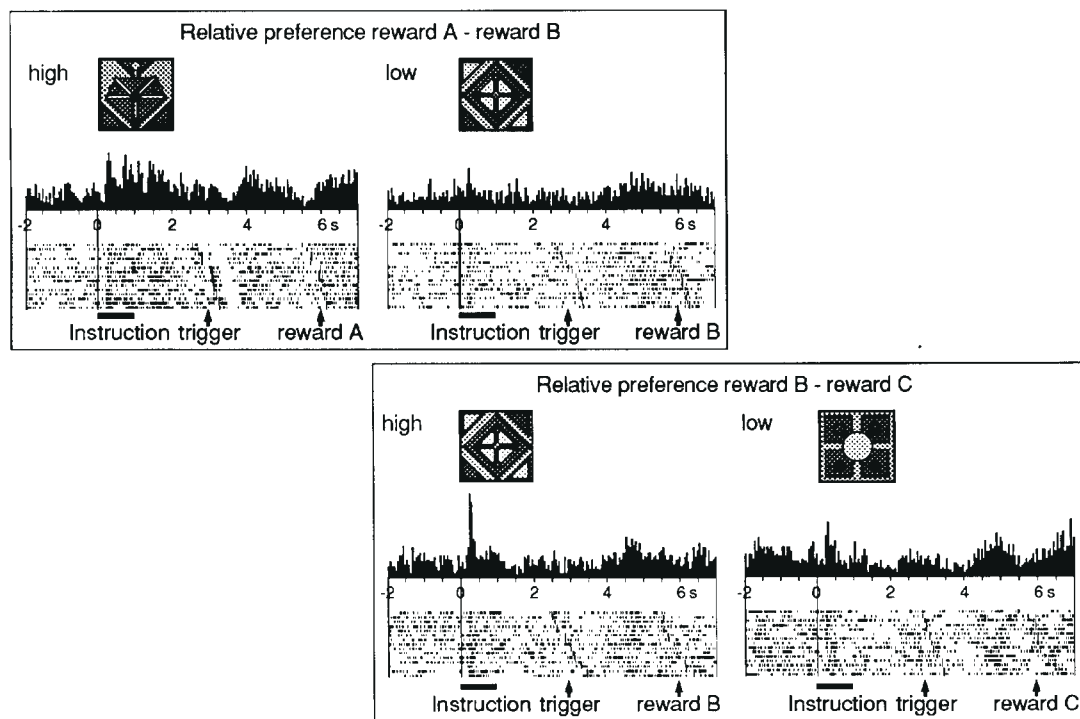


Figure 4. Coding of relative reward preference in an orbitofrontal neuron during performance in the spatial delayed response task. Activity increased in response to the instruction indicating the relatively more preferred food reward (raisin in top, apple in bottom). Although reward B was physically identical in the top and bottom panels (apple), its motivational value differed dependent on the other available reward (low in the top, high in the bottom), as expressed by the animal's choice behavior assessed in separate trials. Different combinations of reward were used in the two trial blocks (top: A = raisin versus B = apple, bottom: B = apple versus C = cereal). Trials alternated randomly within each block between rewards during the experiment. They are separated for analysis according to trial types and rank-ordered according to instruction–trigger intervals. The instruction pictures for each reward are shown above histograms.

ships. These consisted of responses to instructions, sustained activations preceding the trigger stimulus and movement, responses to the trigger stimulus, activations immediately preceding or accompanying the arm movement, sustained activations preceding immediately the reinforcers and responses following the reinforcers (Fig. 9, middle). We studied reward processing with the two types of changes immediately preceding or following the reinforcers and with the remaining task relationships.

Direct Reward Relationship

In the different versions of the delayed go/no-go task we tested how many of a total of 991 task-related striatal neurons showed changes in activity in direct temporal relation to reward (Apicella *et al.*, 1991, 1992; Schultz *et al.*, 1992) or a conditioned auditory reinforcer (Hollerman *et al.*, 1998).

One type of reward relationship consisted of activations preceding the reward in both rewarded trial types irrespective of the go or no-go reaction. These activations were found in 164 of the 991 task-related striatal neurons (17%). They were more frequent in the ventral striatum (43 of 117 task-related neurons; 37%) as compared with the dorsal regions of the caudate and putamen (87 of 615 neurons; 14%). Activations began usually >1 s before the liquid reward (mean 1200 ms), remained present until the reward was delivered and terminated 0.5–1.0 s afterwards, even when the reward occurred before or after the usual time (Fig. 6). Some of these activations discriminated between different reward liquids. They were usually absent in sound-reinforced trials, with the exception of 11 of 259 task-related

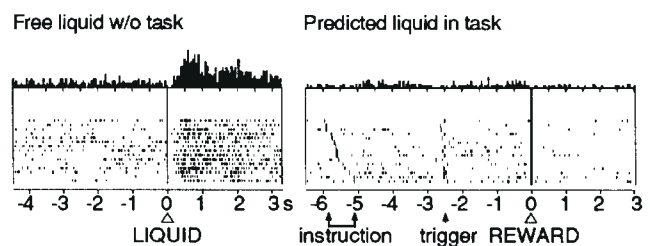


Figure 5. Relationship of reward response to reward unpredictability in an orbitofrontal neuron. The reward response occurred in free liquid trials, in which the liquid occurred at unpredictable times, but not during performance in the delayed go/no-go task (rewarded movement trials; reward response occurred also in nonmovement trials and is not shown). Task trials are rearranged according to instruction–trigger intervals. Free liquid trials were run in separate blocks.

neurons (4%) which showed mostly weak activations preceding preferentially the conditioned auditory reinforcer.

The second type of reward relationship consisted of responses following the delivery of reward in both rewarded trials irrespective of the go or no-go reaction. These responses were found in 168 of the 991 task-related striatal neurons (17%). They were more frequent in the ventral striatum (46 of 117 task-related neurons; 39%) than in the dorsal regions of the caudate and putamen (69 of 615 neurons; 11%). Only five of 259 task-related striatal neurons (2%) responded preferentially to the conditioned reinforcing sound.

Reward Dependence

In the delayed go/no-go task with unrewarded movement trials we tested how activity in 259 task-related striatal neurons differed between rewarded and unrewarded trials (Hollerman *et al.*, 1998). Rewarded nonmovement trials served as controls for movement relationships.

The prediction of liquid reward at trial end influenced all forms of task relationship. In one form, movement preparatory activity depended on the reward expected at trial end. Relationships to movement preparation were seen with transient responses to the initial instructions and with sustained activations during the instruction–trigger delay which occurred frequently only in movement but not in nonmovement trials (54 and 50 of

the 259 task-related neurons, 20 and 19%, respectively). Nearly all of these neurons showed transient instruction responses or sustained delay activations that occurred selectively either in rewarded (33 and 41 neurons) or in unrewarded movement trials (19 and eight neurons, respectively) (Fig. 7A). A purely reward-predicting property was seen in 18 and 10 of the 259 task-related neurons (7 and 4%) which showed transient instruction responses and sustained delay activity, respectively, in both rewarded trial types irrespective of the movement, but not in unrewarded movement trials.

Activations related to the initiation and execution of movement showed comparable dependence on reward. Many activations following the trigger stimulus or occurring immediately before or during the arm movement were seen only in movement but not nonmovement trials (61 of the 259 task-related neurons, 24%). More than two-thirds of these neurons were activated selectively either in rewarded (26 neurons) or unrewarded movement trials (18 neurons) (Fig. 7B), the remaining neurons being activated in both movement trial types irrespective of the reward. Pure reward relationships were seen with 10 of the 259 task-related neurons (4%) which showed post-trigger activations in both rewarded trial types irrespective of the movement.

Reward Coding by Dopamine Neurons

We tested dopamine neurons in several behavioral tasks. (i) In free liquid trials tested in four monkeys, fluid-deprived animals received a small quantity of apple juice at the mouth (0.15 ml) in the absence of reward-predicting phasic stimuli and without performing in any behavioral task (Mireniewicz and Schultz, 1994; Hollerman and Schultz, 1998). Intervals between rewards were irregular and >10 s. (ii) When testing self-initiated movements in two monkeys, animals had one hand on a touch-sensitive, immovable resting key below a food box (Romo and Schultz, 1990). A cover prevented vision into the interior of the box while permitting access of the hand from below. At a self-chosen moment, animals released the resting key without any phasic external stimuli and reached into the box for obtaining a small morsel of apple. (iii) In reaction time tasks tested in four monkeys, stimuli consisted of a light emitted by a diode or a sound emitted by a small loudspeaker mounted slightly below eye level in front of the animals (Ljungberg *et al.*, 1992; Mireniewicz and Schultz, 1994). Following the

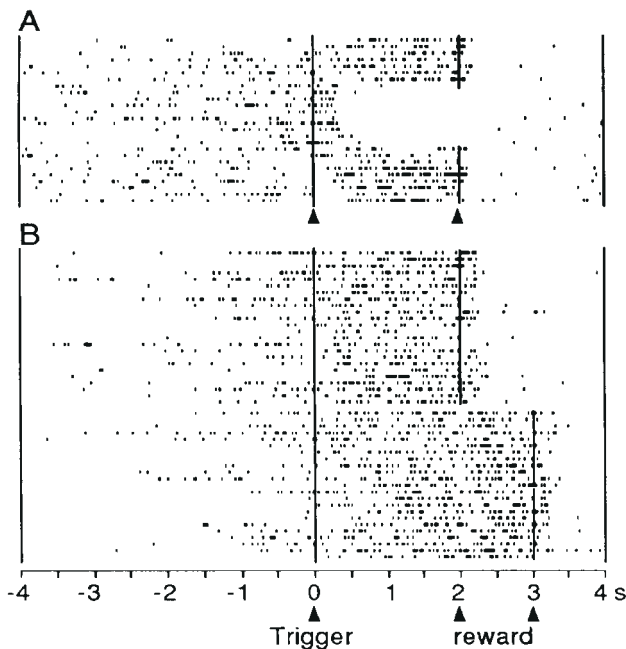


Figure 6. Activations preceding reward in two striatal neurons during performance in the delayed go/no-go task. (A) Activations terminated earlier when reward was delivered earlier than usually, at about the time of the trigger stimulus. Top and bottom trials were recorded with reward delivered at the usual time. (B) Activations lasted longer when reward was delayed by 1.0 s (bottom). In (A) and (B) the original sequence of trials is displayed from top to bottom. Only activities in no-go trials are shown.

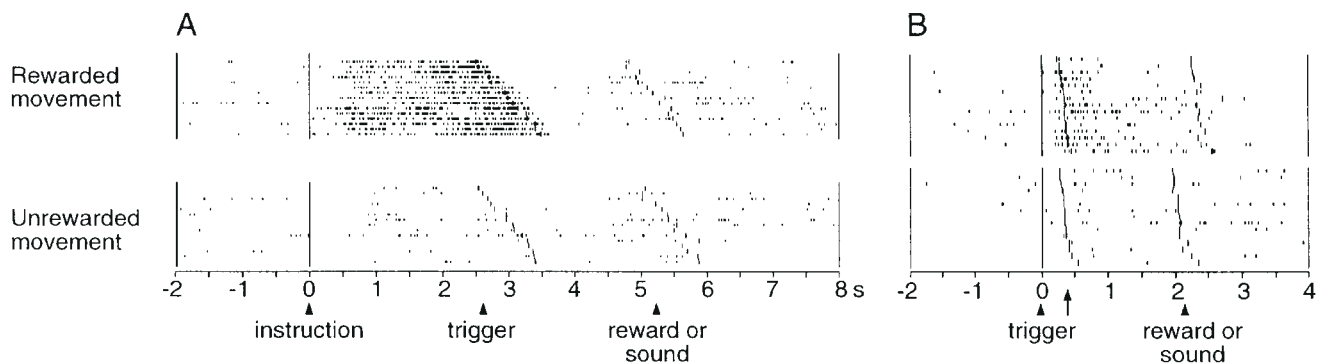


Figure 7. Movement-related activations in two caudate neurons depending on expected reward. (A) Selective activation during the movement preparatory instruction–trigger delay in rewarded but not unrewarded movement trials (top versus bottom). (B) Activations following the movement trigger stimulus in rewarded but not unrewarded movement trials (top versus bottom). Note that the trigger stimulus was identical in all trial types. Neuronal activity is referenced to instruction onset (A) and movement trigger (B). The small arrow next to the trigger line in (B) indicates movement onset (release of resting key). No activations were seen in nonmovement trials (not shown). In (A) and (B) trials alternated semi-randomly between the three trial types. They are separated for analysis according to trial types and rank-ordered according to instruction–trigger intervals (A) and trigger–movement onset intervals (reaction time, B).

stimulus, animals released the resting key, touched a small lever mounted immediately below the stimulus and received apple juice. (iv) In an aversive, active avoidance task tested in two monkeys, animals released a small key following conditioned visual or auditory stimuli to avoid a mild air puff to the hand (Mireniewicz and Schultz, 1996). (v) In a spatial delayed response task tested in two monkeys, we used small light-emitting diodes of different colors, forms and spatial position as instruction and movement-triggering stimuli (Schultz *et al.*, 1993). The other aspects of the task were similar to those described for testing orbitofrontal neurons (Fig. 1, bottom). This task was learned via two intermediate tasks which served to introduce the spatial and temporal delay components. (vi) In a visual discrimination task tested in two monkeys, two colored pictures were presented side by side on a computer monitor, and animals touched a lever below one of the pictures to receive a drop of liquid after a delay of 1.0 s (Hollerman and Schultz, 1998).

More than 2000 dopamine neurons were tested with free primary liquid reward in the absence of any specific task, with primary food reward during self-initiated movements, in reaction time tasks using visual and auditory stimuli, and during the learning of reaction time, delayed response and visual discrimination tasks. A sample of 314 dopamine neurons were also tested with primary aversive stimuli and conditioned visual and auditory stimuli in an active avoidance task.

Dopamine neurons in and close to the pars compacta of the substantia nigra and ventral tegmental area (groups A8, A9 and A10) were characterized by their typical electrophysiological properties, namely polyphasic, initially negative or positive waveforms with relatively long durations (1.8–5.5 ms) and low basal discharge rates (0.5–8.0 imp/s), which contrasted with those of pars reticulata neurons of the substantia nigra (70–90 imp/s and <1.1 ms duration) and neighboring fibers (<0.4 ms duration).

Effective Stimuli

Dopamine neurons responded to two types of reward-related events in a similar manner. They were activated by (i) primary liquid and food rewards and (ii) by conditioned, reward-predicting, visual and auditory stimuli (Fig. 8). Thus, ~80% of dopamine neurons showed rather homogeneous, short latency activations following free liquid (Mireniewicz and Schultz, 1994; Hollerman and Schultz, 1998) and when touching food during self-initiated movements (Romo and Schultz, 1990). Most of these dopamine neurons were also activated by conditioned, reward-predicting, visual and auditory stimuli but lost their response to the primary rewards (Ljungberg *et al.*, 1992; Mireniewicz and Schultz, 1994). During learning of the reaction time and delayed response tasks, dopamine neurons were activated by the primary reward whenever reward contingencies changed with a new learning step, whereas reward responses occurred only exceptionally when learning was completed (Ljungberg *et al.*, 1992; Schultz *et al.*, 1993; Mireniewicz and Schultz, 1994). A similar result was obtained when the same dopamine neurons were studied both during and after learning of new pictures in the discrimination task. Here 50% of dopamine neurons were activated by reward during the rising phase of the learning curve, whereas only 12% of them were activated when the curve approached an asymptote (Hollerman and Schultz, 1998). By contrast, primary or conditioned visual or auditory aversive stimuli were rather ineffective, activating <15% of dopamine neurons (Mireniewicz and Schultz, 1996).

Novel, physically salient or particularly intense attention-

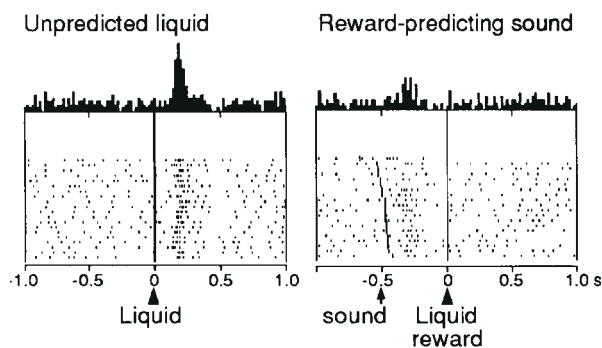


Figure 8. Response transfer from unpredictable, primary liquid reward to conditioned, reward-predicting stimulus in a dopamine neuron. Left: the animal received a small quantity of apple juice at irregular intervals without performing in any behavioral task. Right: the animal performed in an operant lever-pressing task in which it released a touch-sensitive resting key and touched a small lever in reaction to an auditory trigger signal. The dopamine neuron lost its response to the primary reward and responded to the reward-predicting sound.

grabbing stimuli were also effective in activating dopamine neurons. However, the elicited responses usually consisted of activation-depression sequences (Ljungberg *et al.*, 1992). It should be noted that novel or particularly interesting stimuli can also be rewarding, as they sustain approach behavior (Fujita, 1987). Nonrewarding stimuli which physically resembled reward-related stimuli in terms of color, form, size or spatial position induced an activation-depression generalizing response in a somewhat lower fraction of neurons (65%, instead of the 75% with reward-related stimuli) (Mireniewicz and Schultz, 1996).

Reward Prediction Error

A closer inspection of the reward responses suggested that dopamine neurons were very sensitive to the unpredictability of rewards (Schultz *et al.*, 1993, 1997; Schultz, 1998). They were only activated by rewards that were not predictable in time by phasic stimuli, and showed no response to fully predictable rewards (Fig. 8). If, however, a fully predictable reward was omitted because of an error of the animal or an intervention by an experimenter, dopamine neurons were depressed in their activity at the exact time at which the reward would have occurred (Fig. 9, bottom). However, the more general prediction of reward in an experimental context did not seem to influence their responses. This obvious temporal aspect in reward prediction was tested by suddenly modifying the time of reward delivery (Hollerman and Schultz, 1998). An earlier reward (0.5 s) induced an activation for a few trials at the earlier time, and a delayed reward induced a depression at the old reward time and an activation at the new time. Thus dopamine neurons appear to code the temporal discrepancy between the occurrence and the prediction of reward, which is also termed an error in the temporal prediction of reward.

Discussion

These results demonstrate a considerable variety of neuronal activities coding different aspects of rewarding events. Neurons in the orbitofrontal cortex showed three principal forms of reward-related activity, namely responses to reward-predicting instructions, activations during the expectation period immediately before reward, and responses following the reward (Fig. 9, top). Neurons in the striatum were also activated in relation to the expectation and detection of reward (Fig. 9, middle). Other striatal neurons showed activities related to the preparation,

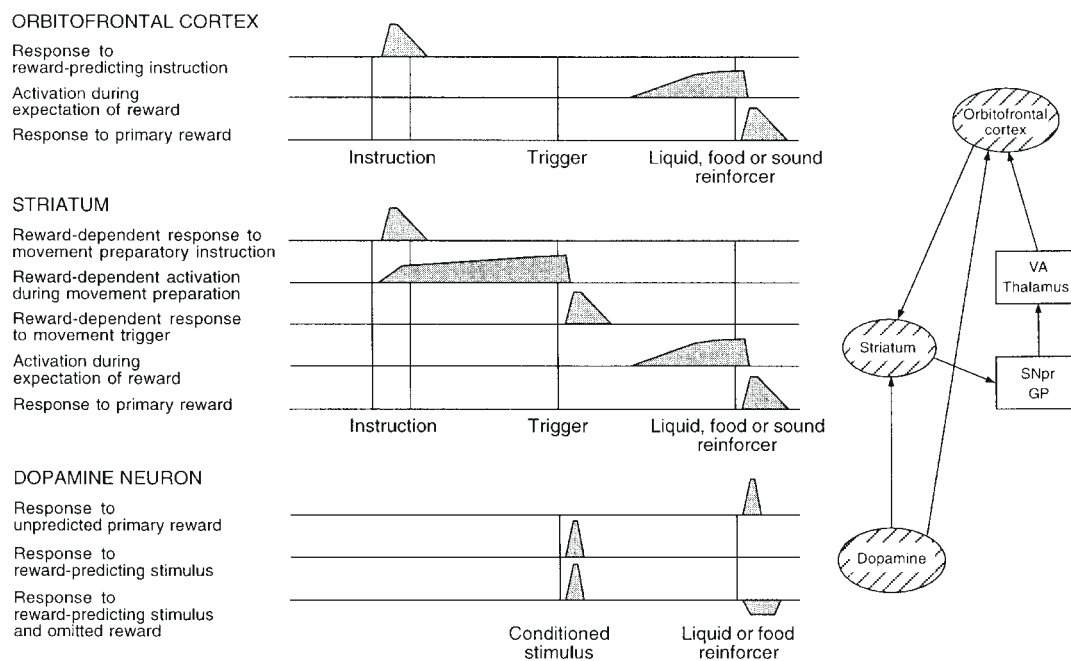


Figure 9. Schematic overview of the forms of reward processing found in fronto-striatal and dopamine systems. Top: neurons in the orbitofrontal cortex showed three main forms of task-related changes in delay tasks, all of which were related to the upcoming or past delivery of reward. Middle: in the striatum, all forms of task-related activations in delay tasks depended frequently on reward expected at trial end rather than a conditioned auditory reinforcer. In addition, striatal neurons showed two forms of activations related directly to the time of reward. Bottom: dopamine neurons were activated by unpredictable reward and by reward-predicting conditioned stimuli but were depressed by omitted rewards. In the orbitofrontal cortex and striatum, the different forms were found in different neuronal populations, whereas the same dopamine neurons showed any of the three forms depending on the situations in rewards and conditioned stimuli occurred or were omitted. All displays were constructed from population histograms and represent average changes in single neurons. Right: schematic connectivity of involved structures. Brain areas investigated in the present studies are shaded. VA thalamus, ventroanterior thalamus; SNpr, substantia nigra pars reticulata; GP, globus pallidus.

initiation and execution of movements, and many of these activities depended on reward being delivered at trial end, rather than an auditory reinforcer. Whereas different orbitofrontal and striatal neurons showed different reward-related activities, dopamine neurons displayed rather homogeneous population responses to rewards and reward-predicting stimuli. These responses depended on event unpredictability, as neurons were activated by unpredictable rewards, showed no changes with fully predictable rewards and were depressed when rewards were omitted (Fig. 9, bottom). Thus, the investigated cortical and basal ganglia structures showed multiple, heterogeneous activities that were related to the reward components of the tasks and occurred simultaneously in these structures. This variety of reward signals was revealed because of the temporal resolution provided by neurophysiological techniques (covariation in time with different, consecutively occurring task events), the employed different reward schedules (reward versus no reward, different rewards) and the different forms of reward occurrence (predictable versus unpredictable, delivery versus omission).

Reward Signals

Predominating Reward Relationships in Orbitofrontal Neurons

One amazing finding in the presently employed tasks was the predominant reward-related nature of orbitofrontal task-related activations. These neurons showed a limited spectrum of covariations with behavioral events compared with dorsolateral and ventrolateral prefrontal cortex, where many neurons process, respectively, spatial positions and visual features of

environmental objects (the 'where' or 'how', and the 'what') (Kubota and Niki, 1971; Fuster and Alexander, 1971; Funahashi *et al.*, 1990; Wilson *et al.*, 1993; Ungerleider *et al.*, 1998). These functions may be derived from inputs from posterior parietal cortex and inferotemporal cortex (Petrides and Pandya, 1988; Pandya and Yeterian, 1996). As motivational processes determine the probability and intensity of behavioral reactions, orbitofrontal neurons coding motivational aspects of goal-directed behavior would influence 'why' a behavioral reaction is performed (Fig. 10). These activities may result from trans-synaptic inputs from the striatum with the prominent relationships to reward expectation described in this report, and from inputs from the amygdala and rostral and medial temporal lobe (Barbas, 1988; Carmichael and Price, 1995; Cavada *et al.*, 2000; Öngür and Price, 2000). The apparent oversimplicity of this scheme is demonstrated by the observation that spatial and visual features are not treated in complete separation in the prefrontal convexity (Rao *et al.*, 1997; Owen *et al.*, 1998) and that the delay activity in some dorsolateral prefrontal neurons is sensitive to reward information (Watanabe, 1996; Hikosaka and Watanabe, 2000).

The paucity of activity during the entire instruction-trigger delay (delay 1 of Fig. 1) was rather surprising, as such activations are frequently found in the dorsolateral prefrontal cortex, where they reflect spatial working memory or movement preparation (Fuster, 1973; Kubota *et al.*, 1974; Funahashi *et al.*, 1993). They were also reported from the orbitofrontal cortex (Rosenkilde *et al.*, 1981; Hikosaka and Watanabe, 2000). As the employed tasks used only a single delay period between instruction and reward rather than the two in our experiments, some of the reported

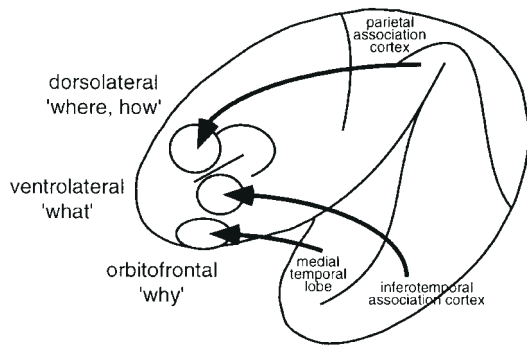


Figure 10. Highly schematic display of separate functions of primate prefrontal cortex. The dorsolateral part (dorsal to and including sulcus principalis) may primarily process the spatial position of goal objects. The ventrolateral part may primarily process the visual features of objects. The degree of specialization may depend on the particular processes using spatial and object information. The present results sustain the notion that the orbitofrontal part may be strongly involved in processing the motivational features of objects presented in goal-directed behavior.

orbitofrontal delay activities might reflect the reward expectation found with our second delay.

Orbitofrontal neurons discriminated very well between rewards and conditioned reinforcers, and between different liquid or food rewards. Many of these discriminations appeared to reflect the motivational value relative to the available alternative reward, as inferred from the animals' preference in overt choice behavior. Apparently these orbitofrontal neurons did not primarily code the pure physical aspects of the reward objects. Physical properties of rewards may be coded in more caudal orbitofrontal areas where neurons discriminate well between various tastes and smells of rewards (Rolls *et al.*, 1990; Critchley and Rolls, 1996; Rolls, 2000).

Striatal Activities in Comparison

Similar to orbitofrontal neurons, slowly firing neurons in the striatum showed activities preceding or following rewards, some of which also discriminated between different reward liquids. However, striatal neurons showed a large spectrum of other behavioral relationships (Schultz *et al.*, 1995). These neurons responded to signals instructing for arm and eye movements, and were activated during the delay periods of tasks testing short-term memory and the preparation of externally instructed or self-initiated arm and eye movements (Hikosaka *et al.*, 1989a; Alexander and Crutcher, 1990; Johnstone and Rolls, 1990; Apicella *et al.*, 1992; Schultz and Romo, 1992). They were also activated during the initiation and execution of arm and eye movements (Crutcher and DeLong, 1984; Hikosaka *et al.*, 1989a; Crutcher and Alexander, 1990; Montgomery and Buchholz, 1991; Romo and Schultz, 1992). Our data suggest now that a considerable proportion of these activities depended on liquid reward delivered at trial end rather than a conditioned auditory reinforcer (Hollerman *et al.*, 1998). A comparable result was obtained in a task using a fixed ratio response schedule in which some striatal neurons responded only to the stimulus closest to the final reward (Shidara *et al.*, 1998). Eye movement-related activations in the striatum showed a comparable dependence on liquid reward rather than no reward, and their directional tuning was strongly modified by the spatial positions of rewarded as compared with unrewarded stimuli (Kawagoe *et al.*, 1998). These data suggest that the expectation of reward influenced heavily striatal activities related to the behavior producing these rewards.

Dopamine Responses in Comparison

Dopamine neurons showed a much smaller spectrum of task-related activities than orbitofrontal and striatal neurons. They were driven by primary rewards, reward-predicting stimuli and attention-inducing stimuli. The responses were similar between different neurons and between different stimuli eliciting them. By contrast, different neurons in the orbitofrontal cortex and striatum showed different responses to the same events or responded to different events. The strong reward relationships of dopamine responses were similar to orbitofrontal responses, although responses to certain attention-inducing stimuli were also seen in dopamine neurons. Dopamine neurons failed to show sustained activations preceding rewards which occurred in orbitofrontal and striatal neurons.

Reward Unpredictability

Most reward-related activations in the orbitofrontal cortex and striatum occurred in relation to well-predictable task events. However, some orbitofrontal neurons responded to reward only when it occurred unpredictably outside of any behavioral task. The dependence on event predictability was the rule for dopamine neurons, being activated by unpredictable reward, uninfluenced by predictable reward and depressed by omitted reward. Thus, a subset of orbitofrontal neurons and most dopamine neurons appeared to code an error in reward prediction.

Distributed Processing of Reward Features

Learning

Rewards have an important function in learning, as they increase the frequency and intensity of behavior leading to rewards. This function depends on the unpredictability of rewards, as formal learning theories and reinforcement learning models indicate that only surprising rewards have the capacity to support learning (Rescorla and Wagner, 1972; Sutton and Barto, 1981). Neuronal responses signalling selectively unpredictable rewards would be particularly effective as teaching signals in learning networks. The rather homogeneous reward response in the majority of dopamine neurons would be broadcast as a global reinforcement signal along divergent projections to large numbers of neurons in the striatum and frontal cortex, where it could exert an immediate or long-term influence on the efficacy of other inputs to these neurons or modify synaptic strengths further downstream (Schultz, 1998). By contrast, orbitofrontal reward responses occur in selected groups of cortical neurons which, unlike dopamine neurons, project in a highly specific and selective manner to postsynaptic neurons. These responses might serve as selective teaching signals for certain groups of neurons, rather than having a global reinforcing effect.

In contrast to the potential function of teaching signals, dopamine responses did not appear to discriminate between different rewards. (Orbitofrontal responses to unpredictable rewards have not yet been tested with different rewards.) However, correct learning about specific rewards requires discrimination between different rewards. Reward-discriminating neurons described above have been found in the anterior orbitofrontal cortex. They have been reported also in the posterior, gustatory and olfactory parts of the orbitofrontal cortex (Rolls *et al.*, 1990; Critchley and Rolls, 1996; Rolls, 2000), striatum (Cromwell *et al.*, 1998) and amygdala (Nishijo *et al.*, 1988). These neurons detected rewards irrespective of any prediction, indicating that their responses may not constitute ideal teaching signals. As reward unpredictability and reward

discrimination may be processed separately, the optimal use of rewards for learning may involve the simultaneous processing of both reward aspects in different brain structures. Deficiencies in any of these systems would lead to slowed, inaccurate or redundant learning.

Established Task Performance

Activities of neurons in all three structures investigated here reflected the expectation of predictable reward (Fig. 9). In the first form, neurons showed phasic activations following the occurrence or depressions during the omission of predictable rewards. Thus orbitofrontal and striatal neurons were activated by predictable rewards, and dopamine neurons were depressed by omitted rewards. In the second form, dopamine and some orbitofrontal and striatal neurons responded phasically to signals predicting rewards. In the third form, orbitofrontal and striatal neurons showed sustained activations during the expectation immediately preceding a predictable reward. In the fourth form, striatal activations related to the preparation and execution of movements were specifically modulated by expected rewards. Similar activities have been reported for the dorsolateral prefrontal cortex (Watanabe, 1996). Thus various aspects of predictable rewards were processed simultaneously in different brain structures. The optimal use of predictable rewards for controlling behavior would be based on combining this heterogeneous information.

From Reinforcers to Goals of Intentional Behavior

Activities following Rewards

One of the main functions of rewards is to act as positive reinforcers. In initial learning situations, neuronal mechanisms of reinforcement could only use responses following the occurrence of reward, as established reward predictions do not exist. Any activity following reward may contribute to these mechanisms. Responses following reward were presently found in the orbitofrontal cortex, striatum and dopamine neurons. The responses of dopamine and orbitofrontal neurons may constitute effective teaching signals because of their relationship to reward unpredictability, a requirement for appetitive learning (Rescorla and Wagner, 1972). By contrast, rewards occur predictably in established tasks and during learning-related modifications of reward predictions. Only neurons responding to predictable rewards would be activated in these situations. Responses which discriminate between different rewards could be essential for the selectivity of learning. This report showed such responses in the orbitofrontal cortex, and have also been reported for the posterior orbitofrontal cortex (Rolls *et al.*, 1990; Critchley and Rolls, 1996; Rolls, 2000), striatum (Cromwell *et al.*, 1998) and amygdala (Nishijo *et al.*, 1988).

Activations Related to the Expectation of Predictable Rewards

A higher form of reward processing was found with activations which preceded the occurrence of predictable rewards in orbitofrontal, striatal and dopamine neurons. The occurrence of activations before reward in each trial suggests that these neurons had access to stored representations of the upcoming reward. The representations were evoked by the presentation of instruction and trigger stimuli which had gained a predictive value for the future occurrence of reward through the past experience of the animal.

Activations preceding rewards were found in two forms, namely phasic responses to reward-predicting stimuli in orbito-

frontal, striatal and dopamine neurons, and sustained activations in orbitofrontal and striatal neurons which lasted between the last event before the reward and the reward itself. Similar phasic and sustained reward-predicting activations were reported for orbitofrontal and amygdala neurons (Nishijo *et al.*, 1988; Schoenbaum *et al.*, 1998), and for striatal neurons during oculomotor tasks (Hikosaka *et al.*, 1989b) and in relation to the distance to the predicted reward (Shidara *et al.*, 1998). Neurons in the orbitofrontal cortex and striatum adapted their reward prediction-related activities when learning involved the modification of existing reward predictions (Tremblay *et al.*, 1998; Tremblay and Schultz, 2000b).

The activations related to the expectation of a predictable reward provide information about the upcoming event. This information allows subjects to identify objects of vital importance, discriminate valuable from less valuable objects before they appear, mentally evaluate various ways of reacting, compare the gains and losses from each possible reaction, and select and prepare behavioral reactions. The neuronal processing of predictive reward information would be an enormous advantage for the control of behavior and would increase the likelihood of obtaining reward objects.

Activations Reflecting Representations of Goals?

Rewarding outcomes can be considered as goals of voluntary behavior when the behavior is intentionally directed at obtaining the reward. According to motivational theory, at least two basic criteria should be fulfilled before calling an action 'goal-directed' (Dickinson and Balleine, 1994): (i) there should already be representations of the outcome at the time the behavior is executed (knowing the outcome when performing the action), and (ii) the behavior should be based on knowledge about the contingency between choice and outcome (knowing the causal relationship between the action and the outcome obtained by that action). Behavioral choices are made according to the motivational value of rewards (quality, quantity and probability of reward). Monkeys can apparently make decisions about the outcome of behavior. They show consistent choices among different outcomes irrespective of spatial positions or visual features of cues, as observed with video games (Washburn *et al.*, 1991), food versus cocaine discrimination (Nader and Woolverton, 1991) and nutrient rewards (Baylis and Gaffan, 1991; Tremblay and Schultz, 1999).

The present experiments revealed two forms of reward processing that may reflect the neuronal coding of goals of intentional behavior. First, the expectation of reward had pronounced influences on all forms of task-related activity in the striatum (Hollerman *et al.*, 1998). Reward expectation modified the responses to instructions and the sustained activations related to the preparation of movement. Previous studies related these activities to parameters of upcoming movements (Hikosaka *et al.*, 1989a; Alexander and Crutcher, 1990), to execution versus withholding of movement (Apicella *et al.*, 1992) and to stimulus-triggered versus self-initiated movements (Schultz and Romo, 1992). Similar reward influences were seen on activations following the trigger stimulus which may be related to stimulus detection, movement initiation and movement execution (Aldridge *et al.*, 1980; Hikosaka *et al.*, 1989a; Montgomery and Buchholz, 1991; Romo and Schultz, 1992; Gardiner and Nelson, 1992). Comparable reward influences were seen with arm and eye movements in the dorsolateral prefrontal cortex and caudate nucleus (Watanabe, 1992a,b, 1996; Kawagoe *et al.*, 1998). These data suggest that these

neurons already have access to representations of the future reward at the time they are coding the preparation and execution of the action leading to this outcome. Thus striatal and prefrontal neurons could fulfill criterion (i) of goal-directed behavior stated above. However, the data do not reveal whether these neurons would code the causal relationship between the action and its outcome.

In the second form of reward processing in the context of goal-directed behavior, reward discriminations in some orbitofrontal neurons reflected the motivational value of reward relative to available alternatives, as inferred from the preferences animals showed during choice behavior (Tremblay and Schultz, 1999). These neurons showed higher or lower reward-related activations associated with the more preferred reward among two rewards, irrespective of which combination of rewards was used. By contrast, very few orbitofrontal activations reflected physical reward characteristics irrespective of motivational value, although such relationships probably exist in posterior, gustatory and olfactory parts of the orbitofrontal cortex (Rolls *et al.*, 1990; Critchley and Rolls, 1996; Rolls, 2000). These results are consistent with results of lesions of the orbitofrontal cortex. Human patients make errors in decisions about outcomes of actions (Bechara *et al.*, 1998) and monkeys change their reward preferences (Baylis and Gaffan, 1991).

The coding of reward preferences in the orbitofrontal cortex suggests that neurons have access to representations of outcomes relative to other outcomes. The neuronal activations might fulfill criterion (ii) of goal-directed behavior stated above, as the representations seem to take into account the animal's own choice. However, it is unclear whether the activations reflect the relative motivational value of the outcome irrespective of any action or concern also the contingencies of the choice behavior leading to the preferred outcome. A further, interesting aspect in the coding of action–outcome relationships is found in neurons of the cingulate motor area which become active selectively when monkeys change to a different movement after the quantity of reward diminished with the current movement (Shima and Tanji, 1998). The preference-related activations of orbitofrontal neurons may serve to identify the reward with the relatively higher or lower value when a decision is made between alternatives. These activations would be advantageous for neuronal mechanisms involved in making decisions about immediate behavioral choices leading to rewarding goals, as they indicate the most profitable outcome among available alternatives. Processing of information relative to other objects is also found in neurons of the supplementary eye field which code spatial positions of eye movement targets relative to other objects (Olson and Gettner, 1995). This mechanism could facilitate quick orientations relative to established landmarks without computing a full spatial map of the environment. Relative information coding in general would allow subjects to reach quick decisions without going through the lengthy process of computing every aspect of the objects present in the environment.

Notes

We thank the Swiss National Science Foundation for their long-term support of this work. Further support was obtained from the European Community, McDonnell Foundation and Roche Research Foundation, and by postdoctoral fellowships from the FRS Quebec to L.T., NIMH to J.R.H., and Fyssen Foundation and FRM Paris to P. Apicella.

Address correspondence to W. Schultz, Institute of Physiology and Program in Neuroscience, University of Fribourg, CH-1700 Fribourg, Switzerland. Email: wolfram.schultz@unifr.ch.

References

- Aldridge JW, Anderson RJ, Murphy JT (1980) The role of the basal ganglia in controlling a movement initiated by a visually presented cue. *Brain Res* 192:3–16.
- Alexander GE, Crutcher MD (1990) Preparation for movement: neural representations of intended direction in three motor areas of the monkey. *J Neurophysiol* 64:133–150.
- Apicella P, Ljungberg T, Scarnati E, Schultz W (1991) Responses to reward in monkey dorsal and ventral striatum. *Exp Brain Res* 85:491–500.
- Apicella P, Scarnati E, Ljungberg T, Schultz W (1992) Neuronal activity in monkey striatum related to the expectation of predictable environmental events. *J Neurophysiol* 68:945–960.
- Barbas H (1988) Anatomic organization of basoventral and mediodorsal visual recipient prefrontal regions in the rhesus monkey. *J Comp Neurol* 276:313–342.
- Baylis LL, Gaffan D (1991) Amygdectomy and ventromedial prefrontal ablation produce similar deficits in food choice and in simple object discrimination learning for an unseen reward. *Exp Brain Res* 86:617–622.
- Bechara A, Damasio H, Tranel D, Anderson SW (1998) Dissociation of working memory from decision making within the human prefrontal cortex. *J Neurosci* 18:428–437.
- Carmichael ST, Price JL (1995) Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *J Comp Neurol* 363:615–641.
- Cavada C, Compañy T, Tejedor J, Cruz-Rizzolo RJ, Reinoso-Suárez F (2000) The anatomical connections of the macaque monkey orbitofrontal cortex: a review. *Cereb Cortex* 10:220–242.
- Critchley HG, Rolls ET (1996) Olfactory neuronal responses in the primate orbitofrontal cortex: analysis in an olfactory discrimination task. *J Neurophysiol* 75:1659–1672.
- Cromwell HC, Hassani OK, Schultz W (1998) Reward discrimination in primate striatum. *Soc Neurosci Abstr* 24:1652.
- Crutcher MD, Alexander GE (1990) Movement-related neuronal activity selectively coding either direction or muscle pattern in three motor areas of the monkey. *J Neurophysiol* 64:151–163.
- Crutcher MD, DeLong MR (1984) Single cell studies of the primate putamen. II. Relations to direction of movement and pattern of muscular activity. *Exp Brain Res* 53:244–258.
- Damasio AR (1994) *Descartes' error*. New York: Putnam.
- Dias R, Robbins TW, Roberts AC (1996) Dissociation in prefrontal cortex of affective and attentional shifts. *Nature* 380:69–72.
- Dickinson A, Balleine B (1994) Motivational control of goal-directed action. *Anim Learn Behav* 22:1–18.
- Divac I, Rosvold HE, Szwedbart MK (1967) Behavioral effects of selective ablation of the caudate nucleus. *J Comp Physiol Psychol* 63:184–190.
- Eblen F, Graybiel AM (1995) Highly restricted origin of prefrontal cortical inputs to striosomes in the macaque monkey. *J Neurosci* 15:5999–6013.
- Fibiger HC, Phillips AG (1986) Reward, motivation, cognition: psychobiology of mesotelencephalic dopamine systems. In: *Handbook of physiology*. Vol IV. The nervous system (Bloom FE, ed), pp 647–675. Baltimore, MD: Williams & Wilkins.
- Fujita K (1987) Species recognition by five macaque monkeys. *Primates* 28:353–366.
- Funahashi S, Bruce CJ, Goldman-Rakic PS (1990) Visuospatial coding in primate prefrontal neurons revealed by oculomotor paradigms. *J Neurophysiol* 63:814–831.
- Funahashi S, Chafee MV, Goldman-Rakic PS (1993) Prefrontal neuronal activity in rhesus monkeys performing a delayed anti-saccade task. *Nature* 365:753–756.
- Fuster JM (1973) Unit activity of prefrontal cortex during delayed-response performance: neuronal correlates of transient memory. *J Neurophysiol* 36:61–78.
- Fuster JM, Alexander GE (1971) Neuron activity related to short-term memory. *Science* 173:652–654.
- Gardiner TW, Nelson RJ (1992) Striatal neuronal activity during the initiation and execution of hand movements made in response to visual and vibratory cues. *Exp Brain Res* 92:15–26.
- Haber SN, Lynd E, Klein C, Groenewegen HJ (1990) Topographic organization of the ventral striatal efferent projections in the rhesus monkey: an autoradiographic tracing study. *J Comp Neurol* 293:282–298.
- Haber S, Kunishio K, Mizobuchi M, Lynd-Balta E (1995) The orbital and medial prefrontal circuit through the primate basal ganglia. *J Neurosci* 15:4851–4867.

- Hikosaka K, Watanabe M (2000) Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. *Cereb Cortex* 10:263–271.
- Hikosaka O, Sakamoto M, Usui S (1989a) Functional properties of monkey caudate neurons. I. Activities related to saccadic eye movements. *J Neurophysiol* 61:780–798.
- Hikosaka O, Sakamoto M, Usui S (1989b) Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. *J Neurophysiol* 61:814–832.
- Hollerman JR, Schultz W (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neurosci* 1:304–309.
- Hollerman JR, Tremblay L, Schultz W (1998) Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J Neurophysiol* 80:947–963.
- Iversen SD, Mishkin M (1970) Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. *Exp Brain Res* 11:376–386.
- Jacobsen CF, Nissen HW (1937) Studies of cerebral function in primates. IV. The effects of frontal lobe lesions on the delayed alternation habit in monkeys. *J Comp Physiol Psychol* 23:101–112.
- Johnstone S, Rolls ET (1990) Delay, discriminatory, and modality specific neurons in the striatum and pallidum during short-term memory tasks. *Brain Res* 522:147–151.
- Kawagoe R, Takikawa Y, Hikosaka O (1998) Expectation of reward modulates cognitive signals in the basal ganglia. *Nature Neurosci* 1:411–416.
- Kubota K, Niki H (1971) Prefrontal cortical unit activity and delayed alternation performance in monkeys. *J Neurophysiol* 34:337–347.
- Kubota K, Iwamoto T, Suzuki H (1974) Visuokinetic activities of primate prefrontal neurons during delayed-response performance. *J Neurophysiol* 37:1197–1212.
- Ljungberg T, Apicella P, Schultz W (1992) Responses of monkey dopamine neurons during learning of behavioral reactions. *J Neurophysiol* 67:145–163.
- Mirenowicz J, Schultz W (1994) Importance of unpredictability for reward responses in primate dopamine neurons. *J Neurophysiol* 72:1024–1027.
- Mirenowicz J, Schultz W (1996) Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* 379:449–451.
- Montgomery EB Jr, Buchholz SR (1991) The striatum and motor cortex in motor initiation and execution. *Brain Res* 549:222–229.
- Nader MA, Woolverton WL (1991) Effects of increasing the magnitude of an alternative reinforcer on drug choice in a discrete trials choice procedure. *Psychopharmacology* 105:169–174.
- Nishijo H, Ono T, Nishino H (1988) Single neuron responses in amygdala of alert monkey during complex sensory stimulation with affective significance. *J Neurosci* 8:3570–3583.
- Olson CR, Gettner SN (1995) Object-centered direction selectivity in the macaque supplementary eye field. *Science* 269:985–988.
- Öngür D, Price JL (2000) The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb Cortex* 10:206–219.
- Owen AM, Stern CE, Look RB, Tracey I, Rosen BR, Petrides M (1998) Functional organization of spatial and nonspatial working memory processing within the human lateral prefrontal cortex. *Proc Natl Acad Sci USA* 95:7721–7726.
- Pandya DN, Yeterian EH (1996) Comparison of prefrontal architectures and connections. *Phil Trans R Soc Lond B* 351:1423–1432.
- Petrides M, Pandya DN (1988) Association fiber pathways to the frontal cortex from the superior temporal region in the rhesus monkey. *J Comp Neurol* 273:52–66.
- Rao SC, Rainer G, Miller EK (1997) Integration of what and where in the primate prefrontal cortex. *Science* 276:821–824.
- Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: *Classical conditioning. Vol II. Current research and theory* (Black AH, Prokasy WF eds), pp 64–99 New York: Appleton Century Crofts.
- Robbins TW, Everitt BJ (1996) Neurobehavioural mechanisms of reward and motivation. *Curr Opin Neurobiol* 6:228–236.
- Rolls ET (1996) The orbitofrontal cortex. *Phil Trans R Soc Lond B* 351:1433–1444.
- Rolls ET (2000) The orbitofrontal cortex and reward. *Cereb Cortex* 10:284–294.
- Rolls ET, Yaxley S, Sienkiewicz ZJ (1990) Gustatory responses of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey. *J Neurophysiol* 64:1055–1066.
- Romo R, Schultz W (1990) Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. *J Neurophysiol* 63:592–606.
- Romo R, Schultz W (1992) Role of primate basal ganglia and frontal cortex in the internal generation of movements. III. Neuronal activity in the supplementary motor area. *Exp Brain Res* 91:396–407.
- Rosenkilde CE, Bauer RH, Fuster JM (1981) Single cell activity in ventral prefrontal cortex of behaving monkeys. *Brain Res* 209:375–394.
- Schoenbaum G, Chiba AA, Gallagher M (1998) Orbitofrontal cortex and basolateral amygdala encode expected outcome during learning. *Nature Neurosci* 1:155–159.
- Schultz W (1998) Predictive reward signal of dopamine neurons. *J Neurophysiol* 80:1–27.
- Schultz W, Romo R (1992) Role of primate basal ganglia and frontal cortex in the internal generation of movements. I. Preparatory activity in the anterior striatum. *Exp Brain Res* 91:363–384.
- Schultz W, Apicella P, Scarnati E, Ljungberg T (1992) Neuronal activity in monkey ventral striatum related to the expectation of reward. *J Neurosci* 12:4595–4610.
- Schultz W, Apicella P, Ljungberg T (1993) Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J Neurosci* 13:900–913.
- Schultz W, Apicella P, Romo R, Scarnati E (1995) Context-dependent activity in primate striatum reflecting past and future behavioral events. In: *Models of information processing in the basal ganglia* (Houk JC, Davis JL, Beiser DG, eds), pp 11–28. Cambridge, MA: MIT Press.
- Schultz W, Dayan P, Montague RR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599.
- Selemon LD, Goldman-Rakic PS (1985) Longitudinal topography and interdigitation of corticostriatal projections in the rhesus monkey. *J Neurosci* 5:776–794.
- Shidara M, Aigner TG, Richmond BJ (1998) Neuronal signals in the monkey ventral striatum related to progress through a predictable series of trials. *J Neurosci* 18:2613–2625.
- Shima K, Tanji J (1998) Role for cingulate motor area cells in voluntary movement selection based on reward. *Science* 282:1335–1338.
- Sutton RS, Barto AG (1981) Toward a modern theory of adaptive networks: expectation and prediction. *Psychol Rev* 88:135–170.
- Thorpe SJ, Rolls ET, Maddison S (1983) The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp Brain Res* 49:93–115.
- Tremblay L, Schultz W (1999) Relative reward preference in primate the orbitofrontal cortex. *Nature* 398:704–708.
- Tremblay L, Schultz W (2000a) Reward-related neuronal activity during go-nogo task performance in primate orbitofrontal cortex. *J Neurophysiol* (in press).
- Tremblay L, Schultz W (2000b) Modifications of reward expectation-related neuronal activity during learning in primate orbitofrontal cortex. *J Neurophysiol* (in press).
- Tremblay L, Hollerman JR, Schultz W (1998) Modifications of reward expectation-related neuronal activity during learning in primate striatum. *J Neurophysiol* 80:964–977.
- Ungerleider LG, Courtney SM, Haxby JV (1998) A neural system for human visual working memory. *Proc Natl Acad Sci USA* 95:883–890.
- Washburn DA, Hopkins WD, Rumbaugh DM (1991) Perceived control in rhesus monkeys (*Macaca mulatta*): enhanced video-task performance. *J Exp Psychol: Anim Behav Proc* 17:123–129.
- Watanabe M (1992a) Frontal units coding the associative significance of visual and auditory stimuli. *Exp Brain Res* 89:233–247.
- Watanabe M (1992b) Prefrontal unit activity during associative learning in the monkey. *Exp Brain Res* 80:296–309.
- Watanabe M (1996) Reward expectancy in primate prefrontal neurons. *Nature* 382:629–632.
- Wilson FAW, O'Scalaidhe SP, Goldman-Rakic PS (1993) Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science* 260:1955–1958.
- Wise RA (1996) Neurobiology of addiction. *Curr Opin Neurobiol* 6:243–251.
- Wise RA, Rompre PP (1989) Brain dopamine and reward. *Annu Rev Psychol* 40:191–225.