

# Getting Formal with Dopamine and Reward

## Review

Wolfram Schultz<sup>1,2,3</sup>

<sup>1</sup>Institute of Physiology  
University of Fribourg  
CH-1700 Fribourg  
Switzerland

<sup>2</sup>Department of Anatomy  
University of Cambridge  
Cambridge CB2 3DY  
United Kingdom

Recent neurophysiological studies reveal that neurons in certain brain structures carry specific signals about past and future rewards. Dopamine neurons display a short-latency, phasic reward signal indicating the difference between actual and predicted rewards. The signal is useful for enhancing neuronal processing and learning behavioral reactions. It is distinctly different from dopamine's tonic enabling of numerous behavioral processes. Neurons in the striatum, frontal cortex, and amygdala also process reward information but provide more differentiated information for identifying and anticipating rewards and organizing goal-directed behavior. The different reward signals have complementary functions, and the optimal use of rewards in voluntary behavior would benefit from interactions between the signals. Addictive psychostimulant drugs may exert their action by amplifying the dopamine reward signal.

### Introduction

The discovery of neurons synthesizing and releasing the neurotransmitter dopamine in the brain has prompted a number of interesting proposals concerning their function. Following such far-fetched suggestions as an involvement in regulating blood pressure, more recent views favor roles in movements, goal-directed behavior, cognition, attention, and reward, to name only the most pertinent ones. On the one hand, the deficits in Parkinsonian patients, in which the nigrostriatal dopamine system has degenerated, impair the planning, initiation, and control of movements, learning and memory, and motivation and emotional reactions. On the other hand, several lines of evidence suggest a dopamine role in reward and approach behavior. Behavioral studies on patients and animals with experimentally impaired dopamine transmission have demonstrated a prime motivational role of dopamine projections to the nucleus accumbens and frontal cortex (Fibiger and Phillips, 1986; Canavan et al., 1989; Wise and Hoffman, 1992; Robinson and Berridge, 1993; Knowlton et al., 1996; Robbins and Everitt, 1996). These systems appear to be crucially involved in the use of reward information for learning and maintaining approach and consummatory behavior. Electrical self-stimulation studies on animals with electrodes implanted in their brains revealed a number of component structures of the brain's reward system (Wise, 2002

[this issue of *Neuron*]). The structures included the dopamine system, as many of the stimulation sites were in close proximity to axons of dopamine neurons or to axons presynaptic to them (Wise, 1996a). Finally, major drugs of abuse influence dopamine neurotransmission (Wise and Hoffman, 1992; Wise, 1996b; Wise, 2002 [this issue of *Neuron*]). Heroin and other opiates, cocaine, amphetamine, and nicotine lead to increases in dopamine concentration in the ventral striatum and frontal cortex, which appears to be a crucial mechanism of drug addiction. In view of these well-established results, several studies investigated neuronal mechanisms of reward by studying the impulse activity of single neurons in the dopamine system and other presumptive reward structures. In particular, we were interested to understand which specific information about rewards would be coded by the different neuronal systems.

The present review comprises seven sections which (1) summarize the basic neurophysiological results obtained from dopamine neurons; (2) interpret these results in relation to formal issues of animal learning theory; (3) make a case for different functions depending on the time courses of dopamine fluctuations and compare the reward processing with other functions of dopamine systems; (4) assess how the dopamine reward signal influences postsynaptic structures to mediate the known behavioral functions of reward; (5) compare the reward coding of dopamine neurons with other brain structures and discuss how neurons might use the reward information for controlling goal-directed behavior, decision-making, and intentional behavior; (6) suggest hypotheses on how addictive drugs might abuse the dopamine reward signal; and (7) provide an outlook on future experiments on dopamine and reward.

### Basic Results

Cell bodies of dopamine neurons are located in their majority in the ventroanterior midbrain (substantia nigra and ventral tegmental area), in groups numbered A8 to A10 from caudolateral to rostromedial (Andén et al., 1966). Their axons project differentially in a general topographic order to the striatum (caudate nucleus and putamen), ventral striatum including nucleus accumbens, and most areas of neocortex including, prominently, the prefrontal cortex. (An additional, smaller dopamine cell group is located in the hypothalamus but has different functions and is not the subject of the review.) Midbrain dopamine neurons show in their majority similar, phasic activations following rewards and reward-predicting stimuli (Schultz, 1998). There is a tendency for stronger responses in medial midbrain regions, such as the ventral tegmental area (group A10) and medial substantia nigra (medial group A9), as compared to more lateral regions (group A8 and lateral group A9). Response latencies (50–110 ms) and durations (<200 ms) are similar for rewards and reward-predicting stimuli. The dopamine reward response constitutes a relatively homogeneous population signal which is graded in magnitude by the responsiveness of individual neurons and by the fraction of responding neurons.

<sup>3</sup>Correspondence: ws234@cam.ac.uk

### **Activation by Primary Rewarding Stimuli**

About 75% of dopamine neurons show phasic activations when animals touch a small morsel of hidden food during exploratory movements in the absence of other phasic stimuli. Dopamine neurons are also activated by drops of liquid delivered to the mouth outside of any behavioral task or during the learning phase of different Pavlovian or instrumental tasks, such as visual or auditory reaction time tasks, spatial delayed response or alternation tasks, and visual discrimination (Ljungberg et al., 1991, 1992; Schultz et al., 1993; Mirenowicz and Schultz, 1994; Hollerman and Schultz, 1998). Dopamine neurons do not appear to discriminate between different food objects or different liquid rewards. However, their responses distinguish rewards from non-reward objects (Romo and Schultz, 1990). Dopamine neurons show no or only minor changes in activity prior to or during arm or eye movements (DeLong et al., 1983; Schultz et al., 1983b; Romo and Schultz, 1990), the few changes being unrelated to the spatial targets of movements (Schultz et al., 1983a).

Only 14% of dopamine neurons show the phasic activations when primary aversive stimuli are presented, such as an air puff to the hand or hypertonic saline to the mouth, and most of the activated neurons respond also to rewards (Mirenowicz and Schultz, 1996). Although being non-noxious, these stimuli are aversive in that they disrupt behavior and induce active avoidance reactions. However, dopamine neurons are not entirely insensitive to aversive stimuli, as they show depressions or activations with slower time courses following pain pinch in anesthetized monkeys (Schultz and Romo, 1987). Also, dopamine release is increased in the striatum following electric shock or tail pinch in awake rats (Louilot et al., 1986; Abercrombie et al., 1989; Doherty and Gratton, 1992; Young et al., 1992).

Dopamine neurons show phasic activations followed by depressions in response to novel or intense stimuli. These stimuli have both attentional and rewarding properties, as agents show orienting responses to these stimuli which they also find rewarding. These data might suggest that phasic dopamine activations reflect attention-inducing properties of stimuli, including rewards, rather than positive reinforcing components (Schultz, 1992; Redgrave et al., 1999; Horvitz, 2000). However, dopamine neurons are depressed rather than activated by the attention-generating omission of reward (Schultz et al., 1993), and they show only few activations to strong attention-generating events such as aversive stimuli (Mirenowicz and Schultz, 1996).

Taken together these findings suggest that the phasic activations of dopamine neurons report preferentially environmental events with rewarding value, whereas aversive events may be signaled primarily with a slower time course. The dopamine activations do not seem to code primarily attention, although coding of specific forms of attention associated with rewards cannot be excluded.

### **Activation by Conditioned, Reward-Predicting Stimuli**

About 55%–70% of dopamine neurons are activated by conditioned visual and auditory stimuli in various Pavlovian and instrumental conditioned tasks (Miller et al., 1981; Schultz, 1986; Schultz and Romo, 1990; Ljung-

berg et al., 1991, 1992; Mirenowicz and Schultz, 1994; Hollerman and Schultz, 1998). Dopamine responses occur close to behavioral reactions (Nishino et al., 1987). Conditioned stimuli are somewhat less effective than primary rewards in terms of response magnitude and fractions of neurons activated. Dopamine neurons respond only to the onset of conditioned stimuli and not to their offset, even if stimuli are used whose offsets, rather than their onsets, are valid predictors of reward (Schultz and Romo, 1990). Dopamine neurons do not distinguish between visual and auditory modalities of conditioned, reward-predicting stimuli. However, they discriminate between rewarding and neutral stimuli that are physically sufficiently dissimilar (Ljungberg et al., 1992) but show progressively more generalizing, activation-depression responses to unrewarded stimuli with increasing resemblance to reward-predicting stimuli (Mirenowicz and Schultz, 1996; Waelti et al., 2001). Only 11% of dopamine neurons, most of them with rewarding responses, show phasic activations in response to conditioned aversive visual or auditory stimuli in active avoidance tasks using air puffs or drops of hypertonic saline (Mirenowicz and Schultz, 1996). Thus, the phasic responses of dopamine neurons preferentially report environmental stimuli with positive motivational value, without discriminating between different sensory modalities.

### **Responses during Learning**

The dopamine activation undergoes systematic changes during the progress of learning. Primary rewards elicit neuronal activations during initial learning periods which decrease progressively and are transferred to the conditioned, reward-predicting stimuli with increasing learning, as shown in visual and auditory reaction time tasks (Ljungberg et al., 1992; Mirenowicz and Schultz, 1994) and spatial delayed response tasks (Schultz et al., 1993). During a transient learning period, both rewards and conditioned stimuli elicit an activation. After learning is complete, the activation switches instantaneously between unpredicted rewards and reward-predicting stimuli (which are tested in separate trials) (Romo and Schultz, 1990; Mirenowicz and Schultz, 1994).

One crucial difference between learning and fully acquired behavior is the degree of reward unpredictability. When monkeys learn repeatedly novel pictures in a learning task, they develop a learning habit by which they learn novel stimuli within a few trials (Harlow, 1949). These paradigms permit to study single neurons during a whole learning episode and compare the responses with familiar situations, while respecting the limited durations of neurophysiological recordings. In a well-established two-way discrimination learning habit, rewards are expected with a probability of 0.5 during the initial part of each learning episode (Figure 1A). A chance correct response leads to a reward and thus to an outcome that is better than expected (+50% prediction error). Dopamine neurons are activated by the reward in these correct learning trials (Figure 1B, top). By contrast, a chance erroneous response leads to no reward and thus to an outcome that is less than expected (−50% prediction error). Dopamine neurons show a depression at the habitual time of the reward in these error trials (Figure 1B, bottom). Thus dopamine neurons

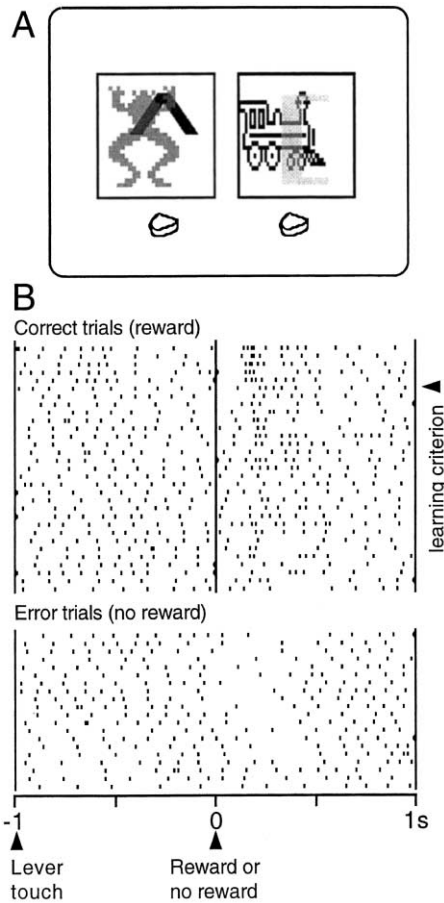


Figure 1. Dopamine Neurons Code a Reward Prediction Error during Learning

(A) Stimulus arrangement in the two-way visual discrimination task. The pictures are presented simultaneously at a computer monitor in front of the monkey, and touch of the rewarded pictures results in presentation of a drop of liquid reward 1.0 s after the touch. Touch of the other picture remains unrewarded but terminates the trial. Left and right stimulus positions alternated semirandomly. Sets of two new pictures are presented repeatedly in blocks of 20–60 trials, and the monkey finds out by trial and error which of the two pictures is rewarded. Thus, with novel pictures, the chance to choose a correct reward is 50%. (B) Dopamine activity during learning of novel picture pairs. (Top) Neuronal activation to liquid reward in correctly performed, rewarded trials (correct lever touch). Activation decreases rapidly after reaching the learning criterion of better than chance performance (arrow). The learning criterion was reached at the second correctly performed trial in the first series of four consecutive, correctly performed trials. (Bottom) Neuronal depression at the habitual time of reward in erroneously performed trials (wrong lever touched). Dots denote neuronal impulses, their horizontal distances representing real time intervals to the electric pulse opening the liquid reward valve, or the habitual time of valve-opening pulse (center arrow). Each line of dots shows one trial, the original sequence being from top to bottom. Reprinted from Hollerman and Schultz, 1998. Copyright (2001) Macmillan Magazines Ltd.

appear to signal the extent to which the rewarding outcome deviates from the prediction during learning. The activations to the reward during initial learning trials are gradually lost as behavioral performance ameliorates and the reward becomes increasingly predicted, whereas

the depression increases in error trials (Hollerman and Schultz, 1998).

### Relationships to Learning Theory *Unpredictability of Reward*

A crucial feature of dopamine responses is their dependency on event unpredictability. The activations following rewards do not occur when food or liquid rewards are preceded by phasic stimuli that have been conditioned to predict such rewards (Romo and Schultz, 1990; Ljungberg et al., 1992; Mirenowicz and Schultz, 1994). The loss of response is not due to a developing general insensitivity to rewards, as activations following rewards delivered outside of tasks do not decrement during several months of experimentation (Mirenowicz and Schultz, 1994). Importantly, the criterion of unpredictability includes the time of reward, as rewards elicit transient activations when they are delivered earlier or later than predicted, even though it is certain that the reward will eventually occur (Hollerman and Schultz, 1998).

Dopamine neurons are depressed exactly at the time of the usual occurrence of reward when a predicted reward is omitted. The depression occurs when animals fail to obtain reward because of erroneous behavior, when liquid delivery is blocked by the experimenter despite correct behavior, when a liquid valve opens audibly without delivering liquid, or when reward delivery is delayed for 0.5 or 1.0 s (Ljungberg et al., 1991; Schultz et al., 1993; Hollerman and Schultz, 1998). The depression occurs even in the absence of any stimuli at the time of the omitted reward which might trigger a response, indicating that the depression does not constitute a simple neuronal response but reflects an expectation process based on an internal clock tracking the precise time of predicted reward.

The reviewed data suggest that dopamine neurons are sensitive to the unpredictability of both the occurrence and the time of reward. They report rewards relative to their prediction, rather than signaling them unconditionally. This sensitivity can best be conceptualized and understood in the formal term of “prediction error.” The dopamine response is positive (activation) when rewards occur without being predicted, or are better than predicted. The response is nil when rewards occur as predicted. The response is negative (depression) when predicted rewards are omitted, or are worse than predicted. Thus, dopamine neurons report rewards according to the discrepancy between the occurrence and the prediction of reward, which can be termed an error in the prediction of reward (Schultz et al., 1997).

### *Dopamine Responses in a Formal Test for Prediction Errors*

The prediction error constitutes a central term of modern theories of conditioning (Rescorla and Wagner, 1972; Mackintosh, 1975; Pearce and Hall, 1980). These theories postulate that, in addition to being paired with a stimulus, a reinforcer must not be predicted in order to contribute to learning. Only stimuli can be learned which are associated with reinforcers that are at least to some degree unpredicted. Reinforcers occurring better than predicted induce learning, fully predicted reinforcers do not contribute to learning, and reinforcers that are worse than predicted, or omitted reinforcers, lead to extinction

of learned behavior. This description holds for a large variety of error-driven learning rules (Sutton and Barto, 1981) and complies with the intuitive concept of learning, according to which behavior changes as long as some outcome is different than predicted, whereas behavior does not change when all outcomes occur exactly as predicted.

The empirical evidence for the role of prediction errors in learning is based on the so-called blocking paradigm, in which a stimulus that is associated with a fully predicted reinforcer cannot be learned (Kamin, 1969). In this paradigm, a new stimulus is presented together with another stimulus that fully predicts the reinforcer. Thus, the reinforcer is expected and therefore generates a minimal prediction error, and the behavioral measures indicate that learning of the new stimulus is blocked. The blocking effect demonstrates that prediction errors, rather than simple stimulus-reinforcer pairings, play a crucial role in learning situations such as animal conditioning (Kamin, 1969), human conditioning (Martin and Levey, 1991), causal learning (Dickinson, 2001), and artificial network learning (Sutton and Barto, 1990).

If dopamine neurons were to code a reward prediction error, their responses should follow the basic assumptions formalized in the blocking paradigm. We investigated how dopamine neurons acquired responses to conditioned stimuli in relation to prediction errors and found considerable similarity with the acquisition of behavioral responses (Waelti et al., 2001). Monkeys failed to learn licking reactions to a novel stimulus when a liquid reward was predicted by a fully trained stimulus and failed to produce a prediction error (Figure 2). In parallel, dopamine neurons failed to acquire activations to the novel stimulus (Figure 3). Dopamine neurons also failed to be activated by the reward during this learning test (Figure 4). In control trials, we added a different novel stimulus to a non-reward-predicting stimulus and delivered a surprising reward which produced a prediction error (Figures 2–4). Monkeys learned readily to lick this novel control stimulus. Dopamine neurons acquired activations to this stimulus, and they were activated by the reward during the learning phase. These data suggest that the acquisition of dopamine responses, like behavioral reactions, is governed by prediction errors, and that both behavioral and neuronal learning is correlated with the activations of dopamine neurons by rewards.

#### Matching Phasic Dopamine Responses to a Learning Theory

The dopamine response reporting an error in the prediction of reward can be expressed as

$$\text{Dopamine Response} = \text{Reward Occurred} \\ - \text{Reward Predicted.}$$

This equation resembles the signed prediction error formalized in the Rescorla-Wagner learning rule, which describes the acquisition of associations between arbitrary stimuli and primary motivating events (reinforcers) in classical conditioning paradigms (Rescorla and Wagner, 1972; Dickinson, 1980). According to this rule, stimuli

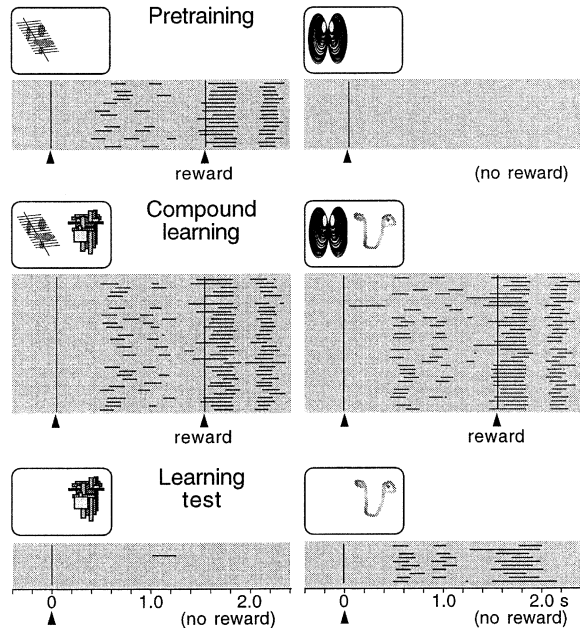


Figure 2. Blocking Paradigm Showing that Learning Depends on Prediction Error Rather than Stimulus-Reward Pairing Alone

Horizontal lines show licking behavior in the task, consecutive trials being shown from top to bottom in each panel. Stimuli were presented on a computer screen in front of the monkey. Three consecutive phases are employed. (Top) During pretraining, a stimulus is followed by a drop of liquid reward, but the control stimulus (right) goes unrewarded. (Middle) During compound learning, a stimulus is added to the established, reward-predicting stimulus without changing reward delivery (no prediction error). The control stimulus is paired with a different stimulus and followed by reward (right) (positive prediction error). (Bottom) In the learning test, the added stimuli are tested in occasional unrewarded trials in semirandom alternation. The results show that the added stimulus is not learned (failed to produce anticipatory licking), as the reward is predicted by the pretrained stimulus and produces no prediction error, although the stimulus is paired with the reward (left). By contrast, the added control stimulus is learned as the reward is not predicted and produces a prediction error (right). Reprinted by permission from *Nature* (Waelti et al., 2001). Copyright (2001) Macmillan Publishers Ltd.

gain associative strength over consecutive trials by being repeatedly paired with a primary motivating event,

$$\Delta V = \alpha\beta(\lambda - V).$$

$V$  is current associative strength of the stimulus,  $\lambda$  is maximum associative strength possibly sustained by the reinforcer,  $\alpha$  and  $\beta$  are learning constants. The unit of quantitative measure is arbitrary and can be derived, for example, from the intensity of the behavioral response or the magnitude of the neuronal response. The  $(\lambda - V)$  term indicates the degree to which the reinforcer occurs unpredictably and represents an error in the prediction of reinforcement. It determines the rate of learning, as associative strength increases when the error term is positive and the conditioned stimulus does not fully predict the reinforcement. When  $V = \lambda$ , the conditioned stimulus fully predicts the reinforcer, and  $V$  will not further increase. Thus, learning occurs only when

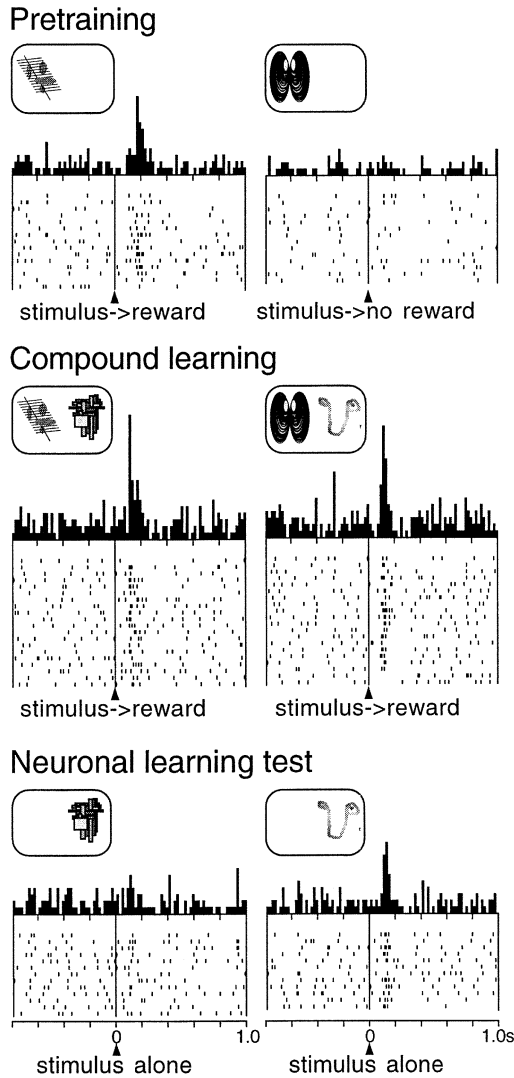


Figure 3. Learning of Dopamine Responses to Conditioned Stimuli in the Blocking Paradigm Depends on Prediction Error Rather than Stimulus-Reward Pairing Alone

(Top) During pretraining, differential activation follows reward-predicting stimulus but not unrewarded control stimulus (right). (Middle) After compound learning, activation to reward-predicting compound is maintained (no prediction error), and activation to control stimulus is learned (right) (positive prediction error). (Bottom) Test trials reveal absent (blocked) neuronal response to the added stimulus but learned response to the control stimulus. Dots denote neuronal impulses, referenced in time to the stimuli (arrows). Histograms contain the sums of raster dots. Reprinted with permission from *Nature* (Waelti et al., 2001). Copyright (2001) Macmillan Publishers Ltd.

the reinforcer is not fully predicted by a conditioned stimulus. The  $(\lambda - V)$  error term becomes negative when a predicted reinforcer fails to occur, leading to a loss of associative strength of the conditioned stimulus (extinction). Note that the Rescorla-Wagner rule uses  $\lambda$  as the maximal possible associative strength of the reinforcer, whereas the equation describing the dopamine response uses the actual occurrence of reward.

The comparison between the phasic dopamine re-

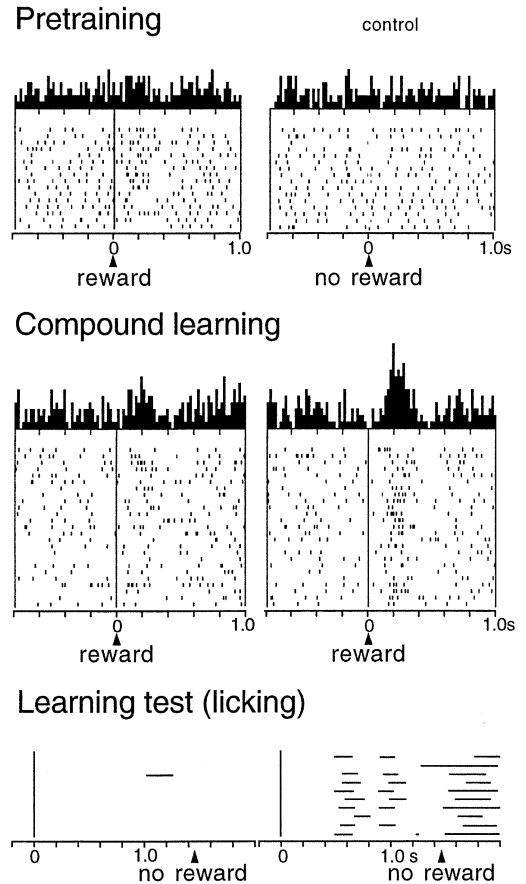


Figure 4. Behavioral Learning Correlated with Dopamine Response at the Time of the Reward in the Blocking Paradigm

(Top) After pretraining, absence of reward responses in fully predicted trials. (Middle) During compound learning, there is no substantial response to predicted reward when stimulus is added (no prediction error), but an activation to surprising reward occurs in control trials in which the added control stimulus produces a surprising reward (right) (prediction error). (Bottom) Behavioral learning test reveals blocked learning of added stimulus, but learning of control stimulus (right). Thus behavioral learning occurs only when dopamine neurons are activated by reward (right). Reprinted with permission from *Nature* (Waelti et al., 2001). Copyright (2001) Macmillan Publishers Ltd.

sponse and the error term of a major learning theory suggests that the dopamine response constitutes an explicit prediction error signal that might be used for modifying synaptic processing. This potential use of prediction error signals has been explored in temporal difference (TD) reinforcement models which implement the Rescorla-Wagner rule and were developed on purely theoretical grounds (Sutton and Barto, 1981). The TD teaching signal resembles the dopamine reward signal in most aspects, and networks using an explicit TD teaching signal similar to dopamine neurons learn even complicated behavioral tasks, such as foraging behavior (Montague et al., 1995), decision making (Montague et al., 1996), sequential movements (Suri and Schultz, 1998), and delayed responding (Suri and Schultz, 1999).

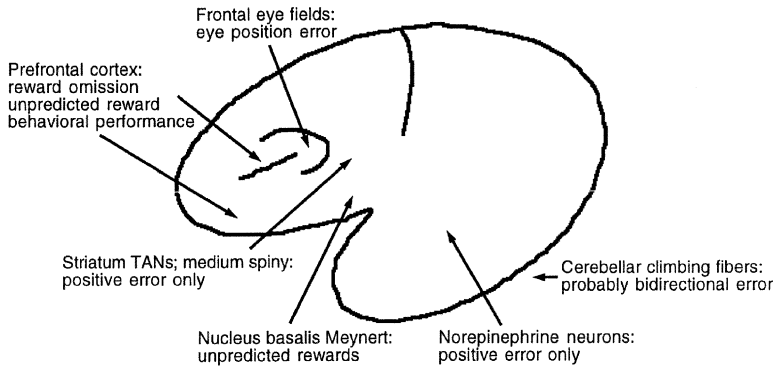


Figure 5. Neural Structures Sensitive to Event Predictability and Errors  
See text for references.

### Neural Coding of Prediction Errors in Other Systems

The cited research has demonstrated that error-driven learning advances only when prediction errors occur and that single dopamine neurons code reward prediction errors in a signed, bidirectional fashion and in compliance with major learning rules. The question arises as to which extent other neuronal learning systems may code explicitly prediction errors. Prime suspects are (1) divergently projecting, ascending systems, such as the monoamine and cholinergic systems, which are anatomically separated from the dopamine systems, (2) classical error-coding structures, such as the cerebellum, which has virtually no connection with the dopamine systems, and (3) other systems with known plasticity, such as the striatum and various cortical areas, which are to a good extent postsynaptic to dopamine neurons. It should be emphasized that the term “prediction error” is a formalism that refers only to the discrepancy between the outcome and its prediction and does not indicate whether the prediction concerns rewards, stimuli, or movements. The following paragraphs evaluate candidate systems in terms of both formalism and content of prediction errors (Figure 5).

Norepinephrine neurons in locus coeruleus show biphasic activating-depressant responses following a wide range of events which elicit attentional orienting reactions. These events include visual, auditory, and somatosensory stimuli (Foote et al., 1980; Aston-Jones and Bloom, 1981; Rasmussen et al., 1986), food objects and liquids delivered outside of behavioral tasks (Foote et al., 1980; Vankov et al., 1995), conditioned rewarding, or aversive stimuli (Rasmussen et al., 1986; Sara and Segal, 1991; Aston-Jones et al., 1994), and infrequent visual stimuli (Aston-Jones et al., 1994). The responses are transient and occur only for a few trials until behavioral situations become stable. They may reflect changes in stimulus occurrence or meaning during learning, reversal, and extinction (Sara and Segal, 1991; Aston-Jones et al., 1997). Thus norepinephrine neurons are driven primarily by the arousing and attention-grabbing components of large varieties of stimuli. The responses are undoubtedly related to event unpredictability, but they probably do not code a full prediction error.

Cholinergic neurons of the nucleus basalis Meynert in the basal forebrain are phasically activated by unfamiliar stimuli (Wilson and Rolls, 1990a), and by rewarding and

aversive stimuli (Wilson and Rolls, 1990b). Some of these neurons respond selectively during reversal (Wilson and Rolls, 1990c) or to unpredicted rewards (Richardson and DeLong, 1990). However, many cholinergic neurons respond to fully predicted rewards in well-established situations (Richardson and DeLong, 1986, 1990; Mitchell et al., 1987), and their responses to reinforcer omission have not been tested. Thus it is unclear whether they might code prediction errors.

The climbing fiber projection from the inferior olive to Purkinje cells of the cerebellar cortex constitutes probably the first known candidate for a neuronal prediction error signal in the brain (Ito, 1989; Kawato and Gomi, 1992; Llinas and Welsh, 1993; Houk et al., 1996; Kim et al., 1998). A climbing fiber, error-driven teaching signal may influence heterosynaptic plasticity by modifying the efficacy of the parallel fiber–Purkinje cell synapses (Marr, 1969). The prediction error carried by climbing fibers may concern, separately, movements and aversive events. Errors in motor performance lead to an increase in climbing fiber activity. The error signal occurs when movements adapt to changes in load or visuo-motor gain and reflects magnitudes of errors in visual reaching (Gilbert and Thach, 1977; Ojakangas and Ebner, 1992; Kitazawa et al., 1998). The aversive prediction error may involve inhibitory connections from nucleus interpositus to the inferior olive and the climbing fiber projection to the cerebellar cortex. Climbing fibers are activated by unpredicted aversive events, do not respond to fully predicted aversive events, and are depressed during extinction, thus complying with a bidirectional error signal according to the formalism of the Rescorla-Wagner rule (Thompson and Gluck, 1991; Kim et al., 1998; Medina et al., 2002). Inactivation of the interpositus-inferior olive projection by lesions or GABA antagonist modifies the climbing fiber response and leads to deficits in learning and produces unblocking (Sears and Steinmetz, 1991; Thompson and Gluck, 1991; Kim et al., 1998).

The striatum (caudate nucleus, putamen, ventral striatum) is composed of slowly firing, medium size spiny neurons, which represent about 95% of the striatal neuronal population, and tonically active interneurons (TANs). TANs respond to primary rewards and conditioned stimuli, sometimes depending on event unpredictability (Apicella et al., 1997; Sardo et al., 2000). Slowly firing striatal neurons respond to primary rewards

with activations which depend occasionally on reward unpredictability, but negative prediction errors were not tested (O.K. Hassani, H.C. Cromwell, and W.S., unpublished data).

The prefrontal cortex shows a considerable spectrum of error coding in humans and animals. Changes in blood flow and brain activity occur in the anterior cingulate, dorsolateral prefrontal, and orbitofrontal parts when agents make behavioral errors or receive instructions that differ from predictions (Gemba et al., 1986; Falkenstein et al., 1991; Berns et al., 1997; Nobre et al., 1999; Fletcher et al., 2001). Some neurons in the dorsolateral prefrontal cortex are particularly activated when monkeys make errors in their behavioral reactions or when reward occurs differently than predicted (Niki and Watanabe, 1979; Watanabe, 1989). Some orbitofrontal neurons are activated when rewards occur unpredictably (Tremblay and Schultz, 2000) or when reward is omitted with incorrect behavior or during reversal (Rosenkilde et al., 1981; Thorpe et al., 1983). Some frontal eye field neurons are sensitive to the difference between current and future eye positions (Umeno and Goldberg, 1997).

#### **Different Dopamine Functions with Different Time Courses**

Before discussing the potential functions of the phasic dopamine signal, we should relate the reward function to other functions of dopamine systems assessed in various clinical and experimental situations. Much original knowledge of the functions of the dopamine systems has been derived from pathological or experimental alterations of dopamine neurotransmission by lesions and administration of receptor antagonists.

Degeneration of the nigrostriatal dopamine system in Parkinson's disease leads to deficits in movement (akinesia, tremor, rigidity), cognition (attention, bradyphrenia, planning, learning), and motivation (reduced emotional responses, depression). Similar deficits are induced by experimental lesions of the nigrostriatal dopamine system or blockade of striatal dopamine receptors by neuroleptics (Poirier, 1960; Burns et al., 1983). Experimental impairments of dopamine neurotransmission in the nucleus accumbens lead to motivational deficits in approach behavior, reward-directed learning, and attentional responses (Fibiger and Phillips, 1986; Wise and Hoffman, 1992; Robinson and Berridge, 1993; Robbins and Everitt, 1996). Reduced dopamine receptor stimulation following lesions of dopamine afferents and local administration of dopamine antagonists in prefrontal cortex lead to impairments in behavioral performance (Brozoski et al., 1979; Simon et al., 1980) and sustained neuronal activity in spatial delayed response tasks (Sawaguchi and Goldman-Rakic, 1991).

These alterations modify both tonic and phasic aspects of dopamine neurotransmission. Dopamine exerts a tonic influence through its sustained, low extracellular concentration in the striatum (5–10 nM) and other dopamine-innervated areas. Dopamine in these concentrations stimulates tonically the extrasynaptic, mostly D2-type dopamine receptors in their high affinity state (Richfield et al., 1989). Interestingly, increases of prefrontal dopamine turnover, rather than decreases, also

induce impairments (Murphy et al., 1996; Elliott et al., 1997). Dopamine concentration is regulated locally within a narrow range by synaptic overflow, extrasynaptic release, reuptake transport, negative feedback control on synthesis, and release via autoreceptors, and presynaptic influences of other neurotransmitters (Chesselet, 1984). Apparently, the tonic stimulation of dopamine receptors should be neither too low nor too high to assure an optimal function of a dopamine-sensitive brain region. These results argue strongly in favor of a role of the ambient, sustained dopamine concentration and receptor stimulation in a large number of behavioral situations.

Many pharmacologically and lesion-induced deficits are considerably ameliorated by systemic dopamine precursor or receptor agonist drug therapy. However, some deficits are not fully restored, such as discrimination deficits (Ahlenius, 1974) and learning, appetitive, and reward deficits (Canavan et al., 1989; Linden et al., 1990; Vriezen and Moscovitch, 1990; Sprengelmeyer et al., 1995; Knowlton et al., 1996). Dopamine agonist treatment cannot in any simple manner reconstitute phasic information transmitted by neuronal impulses, and the deficits remaining after dopamine replacement therapy point to a role of phasic, impulse-related functions. On the other hand, the recovery of many deficits by dopamine agonist therapy suggests that ambient dopamine receptor stimulation has a decisive influence on postsynaptic neurons which is largely independent of phasic changes of dopamine impulse activity.

The separation of dopamine functions into phasic and tonic aspects may be too simplistic, and the available data suggest that at least a third intermediate component should be considered. Studies of dopamine release with voltammetry and microdialysis revealed a number of behavioral relationships with timescales of minutes, including the processing of reward, feeding, drinking, punishment, stress, and social behavior (Loulot et al., 1986; Abercrombie et al., 1989; Young et al., 1992).

Whereas the phasic reward responses of dopamine neurons play a rather focused role in behavior, the deficits arising from reductions in the sustained stimulation of dopamine receptors show a rather wide involvement in behavioral processes. These deficits are not explained by reductions in reward function. It thus appears that dopamine acts at several different timescales, from fast impulse responses related to reward processing via slower fluctuations resembling a local hormone in a wider range of behavioral processes to the tonic function of enabling postsynaptic neuronal systems (Figure 6). These considerations suggest that dopamine is involved in a range of functions dependent on the rate at which its concentrations fluctuate.

#### **Influence on Postsynaptic Structures *Global Reinforcement Signal***

The function of the dopamine reward prediction error signal depends not only on the characteristics of the dopamine signal itself but also on the influence of this signal on postsynaptic structures. The midbrain dopamine systems consist of relatively small numbers of neurons which project in a divergent and widespread manner to much larger numbers of postsynaptic neurons in

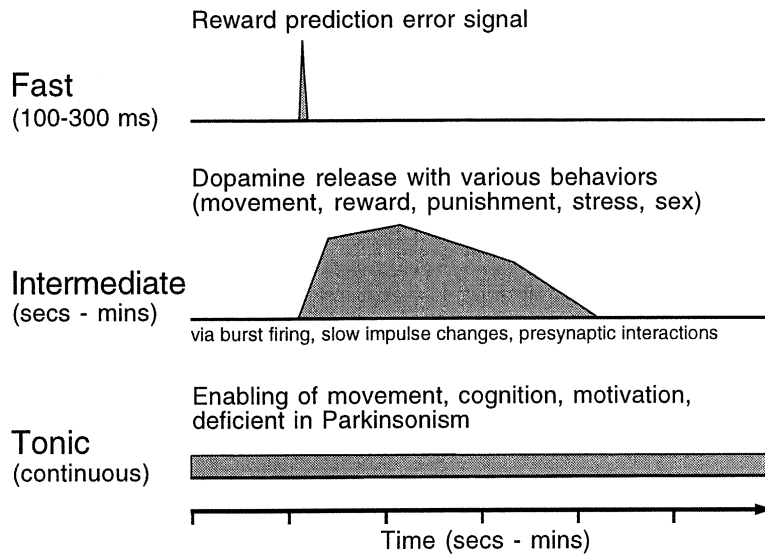


Figure 6. Different Temporal Operating Modes for Different Dopamine Functions

the striatum, including the ventral striatum with nucleus accumbens, the frontal cortex, and a few other structures. There are about 80,000–116,000 dopamine neurons in each substantia nigra in macaque monkeys (German et al., 1988; Percheron et al., 1989). Each macaque striatum contains about 31 million neurons, resulting in a nigrostriatal divergence factor of 250–400. Each dopamine axon ramifies abundantly in a limited terminal area in the striatum and, in the rat, has about 500,000 axonal varicosities from which dopamine is released (Andén et al., 1966). The dopamine innervation reaches nearly every neuron in the striatum through the moderately topographic nigrostriatal projection (Groves et al., 1995; Lynd-Balta and Haber, 1994) and a considerable proportion of neurons in superficial and deep layers of frontal cortex (Berger et al., 1988; Williams and Goldman-Rakic, 1993). In the striatum, every medium-sized spiny neuron receives at its dendritic spines an average of 1,000 dopaminergic synapses and about 5,000–10,000 cortical synapses. The dopamine varicosities often contact the same dendritic spines that are also contacted by cortical inputs (Freund et al., 1984; Smith et al., 1994). A similar dendritic convergence between dopamine and other afferents, including those originating from the hippocampus, is known for cortical pyramidal cells (Goldman-Rakic et al., 1989; Carr and Sesack, 1996). Thus, the rather homogeneous dopamine response, which advances as a simultaneous, parallel wave of activity, is broadcast as a kind of global reinforcement signal along divergent anatomical projections to the much larger populations of neurons in the striatum and frontal cortex (Figure 7, left).

The phasic dopamine reward activation involves 70%–80% of dopamine neurons responding with latencies of 50–100 ms and durations of <200 ms. The impulse response would lead to increases of extracellular dopamine concentrations in the striatum for 200–400 ms, rising from a baseline of 5–10 nM to peaks of 150–400 nM (see Schultz, 1998, based on data by Chergui et al., 1994; Dugast et al., 1994). These concentrations

would be sufficient to transiently activate D1 receptors in their low affinity state (Richfield et al., 1989). Similar effects occur probably in the cortex, concentrations would be slightly lower and less homogeneous with fewer varicosities. Total increases of extracellular dopamine in the striatum last 200 ms after a single impulse and 500–600 ms after multiple impulses of 20–100 ms intervals during 100–200 ms (Chergui et al., 1994; Dugast et al., 1994). Concentrations are homogeneous during most of this period within a sphere of 4  $\mu\text{m}$  diameter (Gonon, 1997), which is the average distance between varicosities (Doucet et al., 1986; Groves et al., 1995). The extrasynaptic dopamine reuptake transporter restricts maximal diffusion to 12  $\mu\text{m}$  and subsequently brings concentrations back to their baseline of 5–10 nM (Herrera-Marschitz et al., 1996). Thus, synaptically released dopamine diffuses rapidly into the immediate juxtasympaptic area and reaches short peaks of regionally homogeneous extracellular concentrations.

Some of the other neural systems carrying prediction errors may exert their influences in a similarly global, divergent way, such as the norepinephrine neurons of locus ceruleus and the cholinergic neurons of the basal forebrain carrying potential attentional prediction errors. By contrast, most of the other mentioned error-coding systems show selective relationships to particular physical aspects of stimuli or parameters of movements and may send their prediction error in an anatomically specific manner to selected, distributed groups of postsynaptic neurons. This would particularly hold for the cerebellar climbing fibers and the different areas of frontal cortex.

#### **Selectivity of Influence on Postsynaptic Neurons**

The question arises how a global reinforcement signal can have selective action on specific postsynaptic neurons. The potential action of the dopamine prediction error signal may be illustrated with an anatomically based model of synaptic inputs to medium-sized striatal spiny neurons (Freund et al., 1984; Goldman-Rakic et al., 1989; Smith et al., 1994) (Figure 7, right). Cortical



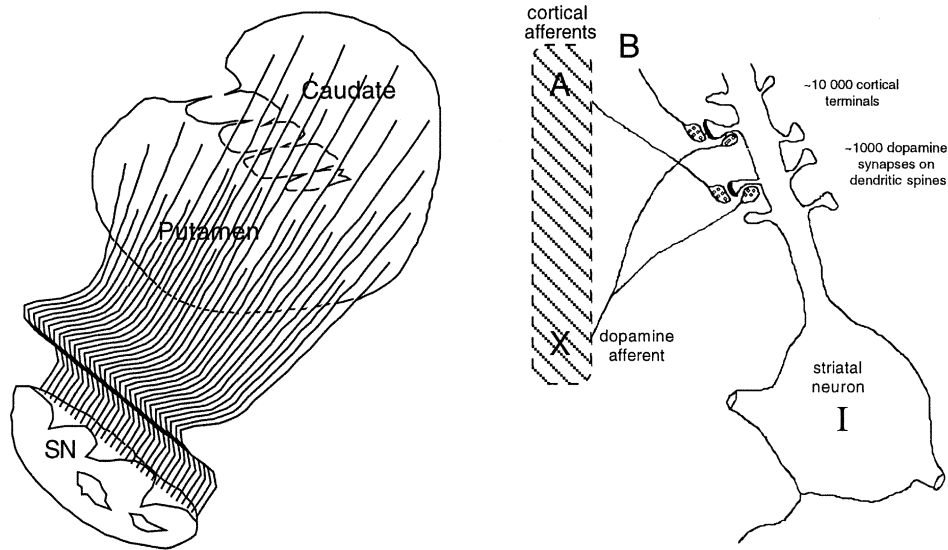


Figure 7. Potential Influences of Phasic Dopamine Reward Signal on Postsynaptic Structures, Based on Known Morphology

(Left) The rather homogeneous population response of impulses of dopamine neurons to reward-related stimuli and its progression from the substantia nigra to the striatum can be schematically displayed as a wave of activity. A similar activity from the dorsomedial substantia nigra and adjoining group A10 to the frontal cortex is omitted for reasons of simplicity. (Right) Synaptic arrangement of inputs from cortex and dopamine neurons to medium size spiny striatal neurons. The dendritic spines are contacted at their tips by cortical axons and at their stems by a dopamine axon. In this example, cortical neurons A and B converge at the tip of dendritic spines of a single striatal neuron I. The efficacy of these connections is modifiable by increased use, e.g., by short-term dopamine-mediated enhancement or by long-term posttetanic potentiation. The modification occurs only when dopamine input X, coming indiscriminately to the stems of the same dendritic spines, is active at about the same time as the cortical input, or following it to be consistent with a reinforcement signal. In the present example, cortical input A, but not B, is active simultaneously with, or a few seconds earlier than, dopamine neuron X when a reward-related event occurs. In the case of the dopamine input arriving after the cortical input, the synapses activated by input A should be marked for later modification (eligibility trace). The mechanism leads to a modification of the A→ I transmission, but leaves the B→ I transmission unaltered. Anatomical data from Freund et al. (1984), drawing modified from Smith and Bolam (1990).

inputs from different origins contact different dendritic spines of striatal or cortical neurons. The same spines are also unselectively contacted by a common dopamine input. Let us assume that activity on the dopamine input signals the occurrence of a reward prediction error in the environment. At the same time, a cortical input is active that codes a specific aspect of the same reward-related event, such as its sensory modality, body side, color, texture, or position, or a specific parameter of a movement. By contrast, other cortical inputs related to events that do not occur at this moment are inactive. The occurrence of the positive or negative dopamine prediction error signal would lead to a global, spatially unselective increase or reduction of dopamine release at most varicosities. However, only those synapses that are activated by cortical inputs at the same postsynaptic spines would be influenced by the dopamine signal, whereas synapses not activated by coincident cortical inputs would remain unchanged. Synaptic transmission would change every time an error signal occurs. By contrast, synapses would not undergo further changes when the behavioral outcome is fully predicted and no neuronal error signal occurs. A directional error signal, such as is exhibited by dopamine neurons, supports not only increments in synaptic transmission under the influence of a positive error signal but may also mediate decrements when a negative error signal occurs in the form of a depression of the baseline rate of activity. Thus

the heterosynaptic influence of the dopamine signal on postsynaptic structures would derive its selectivity from coincidence with activity in the cortical inputs to the same postsynaptic spines.

In the model outlined, any signal occurring on dopamine axons would lead to heterosynaptic influences, irrespective of origin and content. Our research indicates that the natural signal during behavior reflects primarily a reward prediction error, and the induced changes at postsynaptic sites would be related to rewards. However, any activity on the “labeled line” of dopamine axons induced by other means, for example by electrical currents in self-stimulation or other experiments, or by drugs acting on dopamine phasic concentrations, may produce comparable heterosynaptic influences at postsynaptic spines of striatal neurons.

The nature of the influence of the dopamine signal on postsynaptic structures is entirely dependent on the influence of dopamine on the membranes of the neurons innervated by the dopamine axonal varicosities. Recent research has shown various effects of dopamine on postsynaptic membranes in a number of in vitro and in vivo settings that may contribute to mechanisms for reward-directed approach behavior and learning sustained by a phasic dopamine signal. Dopamine has immediate effects on signal processing in postsynaptic neurons, irrespective of longer lasting changes in synaptic transmission. The immediate effects may be com-

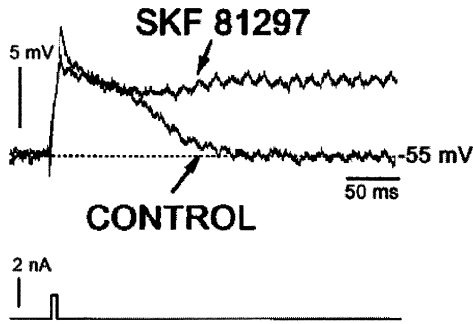


Figure 8. Dopamine D1 Receptor Agonist Enhances Slow Depolarization in Medium Size Spiny Neuron in Striatum

Subthreshold depolarization is induced in vitro by intracellular current injection (calibration of 2 nA), D1 agonist is applied to bath solution (1  $\mu$ M SKF 81297). The D1 receptor antagonist SCH 23390 prevents the enhancement (not shown). The D2 receptor agonist quinpirole (5  $\mu$ M) does not produce the enhancement (data not shown). Reprinted from Hernandez-Lopez et al., 1997. Copyright (1997) Society for Neuroscience.

bined with long-term changes in the dopamine neurons themselves or in neurons further downstream. Studies have found also that dopamine enables or produces long-term depression and long-term potentiation in immediate postsynaptic neurons. The different possibilities will be discussed in the following paragraphs.

#### **Immediate Effects of Dopamine Signal**

In the striatum, dopamine enhances cortically evoked excitations in postsynaptic neurons while reducing spontaneous excitation and increasing hyperpolarized membrane down states, mainly via D1 receptors (Cepeda et al., 1993; Hernandez-Lopez et al., 1997; Ding and Perkel, 2002) (Figure 8). In the cortex, D1 activation increases the efficacy of local inputs (Yang and Seaman, 1996). These effects of dopamine increases the signal-to-noise ratio of active inputs to striatal and cortical neurons.

At the systems level, dopamine exerts a focusing effect whereby only the strongest inputs pass through the striatum to external and internal pallidum, whereas weaker activity is lost (Brown and Arbutnott, 1983; Yim and Mogenson, 1982; Toan and Schultz, 1985; Fillion et al., 1988). In this model, dopamine globally reduces all cortical influences. Only the strongest inputs pass to striatal neurons, whereas other, weaker inputs become ineffective. This function benefits from a nonlinear, contrast-enhancing mechanism, such as the threshold for generating action potentials. A comparable enhancement of strongest inputs would occur in neurons that are predominantly excited by dopamine. In a similar contrast-enhancing manner, signals from norepinephrine neurons are known to increase the signal-to-noise ratio of responses in cerebellar Purkinje cells (Freedman et al., 1977) and potentiate excitatory and inhibitory influences in cerebral cortex (Waterhouse and Woodward, 1980; Sessler et al., 1995).

Through the enhancing and focusing effect, the dopamine signal may serve as an immediate instruction, biasing, gating, or enabling signal, and modify the ways in which other, coincident inputs influence the postsynaptic neurons. This function could select or prioritize cer-

tain simultaneously active inputs to striatal neurons over less active inputs. With competing synaptic inputs, neuronal activities occurring simultaneously with the dopamine signal would be processed with higher priority, as certain inputs would be selected over others depending on their temporal coincidence with the error signal (Schultz, 1998). As a consequence, the dopamine signal could produce a rapid switch of attentional and behavioral processing to reward-predicting, error-generating external events (Redgrave et al., 1999). Behaviorally, the agent would show an orienting and approach response toward the error-generating event, and the attention induced by the reward prediction error could increase the associability of the stimulus. Taken together, the global error message of dopamine neurons could be used for dynamically and instantaneously selecting which external stimuli and behavioral reactions are processed within the limited channel capacity of neuronal transmission.

#### **Immediate Dopamine Effect Coupled with Upstream or Downstream Plasticity**

The enhancing and focusing effect of dopamine may be combined with neuronal plasticity to provide a mechanism for reward-driven learning, irrespective of a direct influence of dopamine on postsynaptic plasticity. According to one mechanism, the immediate influence of the dopamine prediction error signal may benefit from the demonstrated plasticity of the response of the dopamine neurons themselves. The response shift during learning from the primary reward to the stimulus predicting the reward produces a reward-predicting signal. Behavioral reactions would benefit from advance information provided by reward-predicting stimuli and become more frequent, rapid, and precise. For example, the learning of lever pressing is facilitated by pairing the behavioral reaction with a previously conditioned stimulus (Lovibond, 1983). Thus the transfer of response to conditioned reinforcers may constitute a mechanism for behavioral learning even in the absence of dopamine-mediated synaptic plasticity in target areas. In this case, only the shift to predictive coding would occur during learning, and the advance information provided by this earlier signal would have a focusing effect in the striatum and cortex and enhance neurotransmission at the time of the prediction signal.

According to another mechanism, the immediate focusing and enhancing effect of the dopamine signal may lead to Hebbian type plasticity at synapses further downstream of the dopamine terminals in the striatum and cortex. Long-term potentiation has been described for excitatory synapses in the prefrontal cortex (Hirsch and Crepel, 1990) and may operate also further downstream of dopamine influences. The synaptic effects of striatal neurons and the subsequent basal ganglia neurons are mostly inhibitory, but synaptic plasticity at inhibitory synapses have only been described rarely. Nevertheless, an enhancement followed by downstream long-term potentiation could be a way in which the dopamine error message could contribute to learning even without mediating synaptic plasticity directly.

#### **Plasticity Effect of Dopamine Signal**

The direct influences of the dopamine prediction error signal may not only concern immediate membrane reactions but could also lead to synaptic changes in immediate

postsynaptic neurons which may provide a mechanism for reward-driven learning. Dopamine neurotransmission may play a role in postsynaptic plasticity in at least two ways. First, dopamine may have a permissive effect allowing synaptic plasticity to take place. This effect can be investigated by modifying the tonic stimulation of dopamine receptors. Second, phasic increases of dopamine activity could induce synaptic plasticity as a kind of teaching signal.

Recent experiments demonstrate that dopamine agonists augment synaptic plasticity, and dopamine-depleting lesions or dopamine receptor antagonists disrupt plasticity in the striatum and cortex. Several manipulations impair long-lasting posttetanic depression in the striatum and cortex, such as lesions of the nigrostriatal dopamine system, application of D1 or D2 dopamine receptor antagonists, and knockout of D2 receptors (Calabresi et al., 1992, 1997; Otani et al., 1998). Application of dopamine restores the posttetanic depression. Dopamine D1 receptor antagonists impair the induction of long-term potentiation induced by tetanic corticostriatal stimulation, whereas blockade of D2 receptors has no effect (Kerr and Wickens, 2001). Lesions of the nigrostriatal dopamine system impair posttetanic potentiation in the striatum, and dopamine application restores the potentiation (Kerr and Wickens, 2001). Dopamine lesions and receptor antagonists impair striatal synaptic plasticity also during development (Tang et al., 2001). In the prefrontal cortex, lesions of the mesocortical dopamine projection and D1 receptor antagonist administration impair long-term potentiation at hippocampal-prefrontal synapses, whereas D1 receptor agonist administration or adenylyl cyclase activation enhance the potentiation (Gurden et al., 1999, 2000) (Figure 9A). In the hippocampus, D1 dopamine receptor agonist administration enhances long-term potentiation and reduces depotentiation (Otmakhova and Lisman, 1996, 1998). In the amygdala, the dopamine receptor antagonist haloperidol reduces enhanced postsynaptic responses following aversive Pavlovian conditioning (Rosenkranz and Grace, 2002). In *Aplysia*, a dopamine receptor antagonist blocks operant conditioning of a motor pattern, in doses below those reducing synaptic transmission (Nargeot et al., 1999). These data suggest that dopamine neurotransmission influences the expression of changes in postsynaptic plasticity or may even induce such changes. This function may reflect a permissive, enabling effect of dopamine receptor stimulation induced by the tonic extracellular concentration of dopamine, or it may be due to dopamine being released phasically by impulses of dopamine neurons or via presynaptic interactions.

Recent experiments have electrically stimulated the ascending dopamine neurons with single shocks and thus induced impulses which mimic the dopamine response to unpredicted rewards. Some of these studies have followed the basic notions of behavioral reinforcement learning according to which the unconditioned stimulus (US) should follow, rather than precede, the conditioned stimulus (CS) in order to be an effective reinforcer. If the dopamine response were to act as a teaching signal in reinforcement learning and constitute a neuronal US, it should effectively occur a few seconds after the event to be conditioned, rather than before the event (no backward conditioning).

In testing these possibilities, electrical stimulation of the substantia nigra at positions supporting intracranial self-stimulation transforms corticostriatal depression into potentiation, presumably via activation of axons of dopamine neurons (Reynolds et al., 2001) (Figure 9B). The induction of potentiation is correlated with the speed of self-stimulation learning in the same animals and blocked with D1 receptor antagonist. The application of short pulses of dopamine (5–20 ms) together with depolarization of the postsynaptic striatal neuron induces similar corticostriatal potentiation (Wickens et al., 1996), and phasic application of dopamine in the prefrontal cortex facilitates long-term potentiation (Blond et al., 2002). In the auditory cortex, acoustic stimulation with a specific frequency selectively enlarges receptive fields when the stimulation is immediately followed by electrical stimulation in the area of dopamine neurons in the ventral tegmental area (Bao et al., 2001) (Figure 9C). The effect is blocked by D1 and D2 antagonists and does not occur when the electrical stimulation is applied before the sound (no backward conditioning). Although the electrical stimulation does not necessarily reflect a behavioral situation in which reward occurs, the results demonstrate that activity on the “labeled line” of dopamine axons may induce long-lasting changes in synaptic transmission. In *Aplysia*, contingent, but not unpaired, brief iontophoretic puffs of dopamine induce operant conditioning in an in vitro cellular model (Brembs et al., 2002). Taken together, these data suggest that the phasic activation of dopamine neurons may induce long-term synaptic changes. Some of the effects depend on the exact timing of the dopamine activation relative to the event to be conditioned, in analogy to behavioral learning, and are thus compatible with the notion of dopamine responses acting as a neurochemical reinforcement or teaching signal.

The action of a teaching signal can be formalized by applying the Rescorla-Wagner learning rule to changes in synaptic weight (Sutton and Barto, 1981). A dopamine teaching signal could modify the weights of corticostriatal or corticocortical synapses according to the three factor Hebbian learning rule  $\Delta\omega = \epsilon r i o$ , with  $\omega$  as synaptic weight,  $\epsilon$  as learning constant, and the three factors  $r$  (dopamine prediction error signal),  $i$  (presynaptic input activation), and  $o$  (activation of postsynaptic striatal neuron). The dopamine error signal would modify neuronal transmission at synapses with Hebbian type plasticity based on coincident pre- and postsynaptic activity (Figure 7, right).

The full use of a dopamine teaching signal may involve the response transfer from the primary reward to the reward-predicting stimulus. The biological mechanisms for the backward shift in time are unknown and can be modeled with eligibility traces (Suri and Schultz, 1999). Subsequently, the dopamine signal from the reward-predicting stimulus may serve as teaching signal. The response transfer may mediate the phenomenon of conditioned reinforcement, as predictors of primary reinforcers acquire reinforcing properties themselves. Synaptic weights would be selectively modified on the basis of stimulus- and behavior-related activity coincident with the occurrence of the reward-predicting stimulus and compatible with the three factor learning rule ( $\Delta\omega = \epsilon r i o$ ), in the same manner as with a dopamine signal

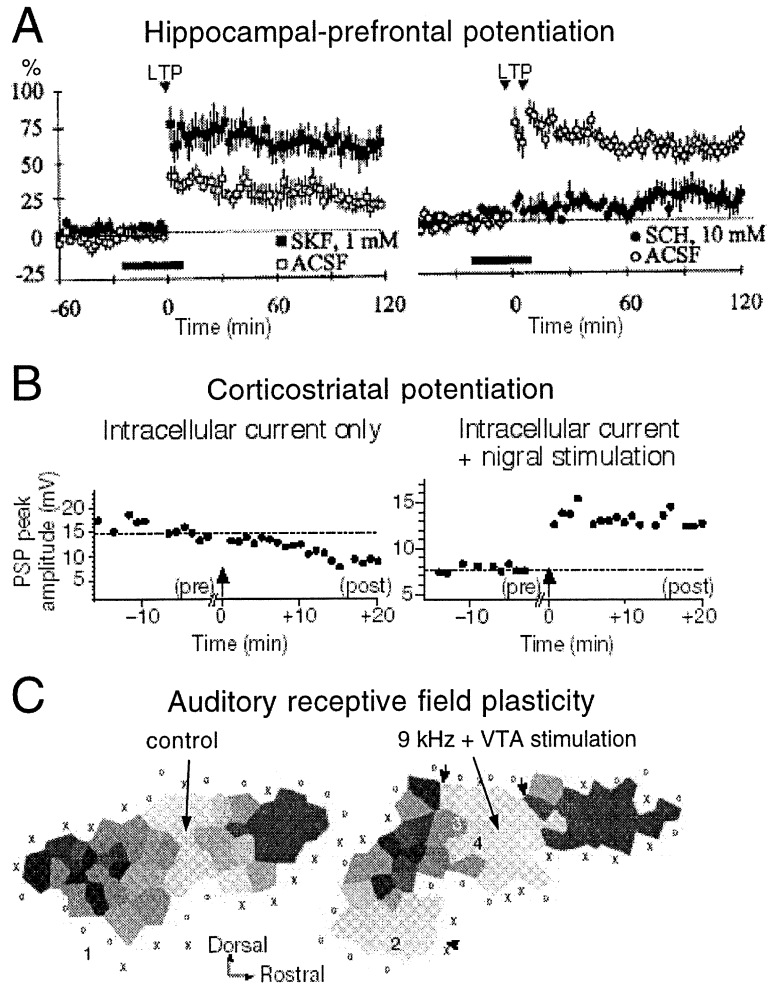


Figure 9. Plasticity Induced by Dopamine Activity

(A) Dopamine D1 receptor agonist enhances hippocampal-prefrontal long-term potentiation *in vitro*. Tetanic stimulation is indicated by arrowheads. The D1 agonist SKF 81297 is applied to the bath solution and produces the effect at 0.1 and 1.0 mM (left). The D1 receptor antagonist SCH 23390 prevents the long-term potentiation at 2.0 and 5.0 mM (right). The D2 receptor antagonist sulpiride (5 and 10 mM) does not prevent the long-term potentiation (data not shown). ACSF, artificial cerebro spinal fluid. Reprinted from Gurden et al., 2000. Copyright (2000) Society for Neuroscience. (B) Electrical stimulation in rat substantia nigra induces potentiation of corticostriatal synaptic transmission *in vivo*. Stimulation is performed through implanted electrodes at positions and with parameters that support intracranial self-stimulation in the same animals before recordings (arrows). The degree of potentiation is correlated with training time for self-stimulation behavior across animals (one electrode per animal). Corticostriatal activation is provided by spontaneous impulse activity. The striatal neuron under study is depolarized by current injections, together with nigral stimulation, to levels inducing action potentials. The potentiation following electrical stimulation is blocked by the D1 receptor antagonist SCH 23390. PSP, post-synaptic potential. Reprinted with permission from *Nature* (Reynolds et al., 2001). Copyright (2001) Macmillan Publishers Ltd. (C) Electrical stimulation in rat ventral tegmental area (VTA) increases cortical auditory receptive fields of frequencies of immediately preceding sounds (9 kHz). The effect is blocked by D1 and D2 antagonists and does not occur with electrical stimulation applied before the sound (no backward conditioning). Reprinted with permission from *Nature* (Bao et al., 2001). Copyright (2001) Macmillan Publishers Ltd.

elicited by primary reward. Modeling studies demonstrate that conditioned reinforcement signals are more efficient for acquiring chains or sequences of predictive stimuli than signals from primary reinforcers occurring only after the behavioral reaction. The progressively earlier anticipation of the primary reward allows to assign the credit for reward more easily (Sutton and Barto, 1981; Friston et al., 1994; Suri and Schultz, 1998).

Plasticity based on three factor learning rules may also play a role with cerebellar climbing fibers which influence the efficacy of parallel fiber synapses on Purkinje cells during motor and aversive learning (Eccles et al., 1967; Marr, 1969; Ito, 1989; Kawato and Gomi, 1992). Synchronous coactivation of parallel and climbing fiber inputs leads to synaptic plasticity in the form of long-term depression at parallel fiber–Purkinje cell synapses, which otherwise undergo a mild, baseline long-term potentiation through the tonically active parallel fiber input alone or the asynchronous activation of the two input systems (Sakurai, 1987; Linden and Connor, 1993). In analogy, receptive fields of cerebellar neurons can be increased or decreased depending on the degree of coactivation of parallel and climbing fibers (Jörntell

and Ekerot, 2002). The effects of norepinephrine neurons are also thought to comply with a three factor learning rule, by inducing or facilitating the induction of long-term depression and potentiation in the hippocampus (Dahl and Sarvey, 1989; Katsuki et al., 1997).

### Use of Rewards for Directing Behavior *The Dopamine Reward Alert Signal*

The main characteristics of the dopamine response may provide indications about its potential function in behavior. The response transmits a reward prediction error, has a rather short latency, fails to discriminate between different rewards, and is broadcast as a global, diverging signal to postsynaptic neurons in the striatum and frontal cortex where it may have focusing, excitation-enhancing, and plasticity-inducing effects. The reward prediction error signal indicates the appetitive value of environmental events relative to prediction but does not discriminate between different foods, liquids, and reward-predicting stimuli, nor between visual, auditory, and somatosensory modalities. The short latency of the response assures an early information about the surprising occurrence or omission of a rewarding event, but

the trade-off is a lack of time for more differentiated evaluation of the specific nature of the particular rewarding event. Thus, the dopamine signal may constitute an alert message about a reward prediction error without detailed information about the nature of the reward.

The dopamine response may constitute a bottom-up signal that rapidly informs postsynaptic structures about surprising rewards, or reward omissions, but the signal needs to be supplemented by additional information to be useful for a full appreciation of the reward. The advantage for behavior provided by such a reward alert signal may consist in allowing rapid behavioral reactions toward rewarding events, whereas the exact nature of the rewards would be evaluated by slower systems during the approach behavior to the reward. The dopamine response to reward-predicting stimuli would increase the temporal advantage further, as behavioral reactions can be initiated even before the reward itself becomes available. The evolutionary advantage of such a rapid, bottom-up, reward-detecting system becomes apparent in situations in which the speed of approach behavior determines the competition for limited food and liquid resources.

Despite the limited reward information provided by dopamine responses, their use for approach behavior and learning could be substantial, as shown by the recently developed temporal difference reinforcement models using dopamine-like teaching signals and architectures similar to the diverging dopamine projections (Sutton and Barto, 1981; Montague et al., 1996). Without additional explicit reward information, such reinforcement models replicate foraging behavior of honeybees (Montague et al., 1995), simulate human decision making (Montague et al., 1996), learn spatial delayed response tasks typical for frontal cognitive functions (Suri and Schultz, 1999), balance a pole on a cart wheel (Barto et al., 1983), learn to play world class backgammon (Tesauro, 1994), move robots about two-dimensional space and avoid obstacles (Fagg, 1993), insert pegs into holes (Gullapalli et al., 1994), and learn a number of simple behavioral reactions, such as eye movements (Friston et al., 1994), sequential movements (Suri and Schultz, 1998), and orienting reactions (Contreras-Vidal and Schultz, 1999). Although some of these tasks occur in reduced experimental situations and comprise rather basic behavioral reactions without much event discrimination, the results demonstrate the computational power of simple bottom-up reward systems in a range of behavioral tasks. All of the mentioned reinforcement models use the reward message as a straightforward teaching signal that directly influences modifiable synapses in executive structures. However, the learning mechanisms could involve a two stage process in which the dopamine reward message would first focus neuronal processing onto the events surrounding and leading to the reward, and the actual learning and plasticity mechanisms occur downstream from the incoming and enhanced reward message. Although neuronal plasticity downstream of dopamine neurons has rarely been investigated, it might be interesting to test the utility of such more differentiated learning models.

### **Reward Discrimination**

Often more than one reward occurs in natural settings, and agents need to select among different alternatives. As the full detection and discrimination of rewards lies beyond the bottom-up, error-reporting capacity of dopamine neurons, additional reward processing mechanisms are necessary to employ the full potential of rewards for approach behavior and learning. Neurons in a number of brain structures discriminate between different food and liquid rewards, such as the dorsal and ventral striatum, orbitofrontal cortex, and amygdala. These structures show transient responses following the presentation of rewards and reward-predicting cues (Thorpe et al., 1983; Nishijo et al., 1988; Bowman et al., 1996; Hassani et al., 2001). Orbitofrontal neurons discriminate between rewards but not between the spatial positions or visual aspects of reward-predicting visual cues (Tremblay and Schultz, 1999). Some orbitofrontal neurons respond selectively to the taste of fat (Rolls et al., 1999). The selectivity of reward-discriminating orbitofrontal neurons can change when the agent is satiated on specific foods, demonstrating a relationship to the motivational value of the food objects rather than their physical appearance (Critchley and Rolls, 1996).

Some orbitofrontal neurons discriminate reward on the basis of relative preferences for the different rewards rather than physical properties. As an example, a reward that is more preferred than another reward may elicit a stronger neuronal response than the less preferred reward. However, the initially more preferred reward may not produce a response when an even more preferred reward becomes available, which then draws the response (Tremblay and Schultz, 1999). These orbitofrontal neurons may subserve the perception of rewards and produce signals for executive structures involved in selecting rewards according to their values. These neuronal mechanisms may underlie the well-known reward contrast effects (Flaherty, 1996) and relate to the more general way of coding information in the brain relative to other available information.

Although the properties of the dopamine response comply well with theoretical notions about learning systems, neurons in the other reward systems, with their selective relationships to specific aspects of rewards, are likely to have also the capacity for synaptic plasticity. Neuronal transmission in the frontal cortex and striatum may undergo long-term potentiation and depression (Hirsch and Crepel, 1990; Calabresi et al., 1992), and this mechanism probably also holds for the reward-processing neurons in these structures. Moreover, some neurons in these structures are sensitive to event unpredictability (see above) and may thus provide teaching signals for other neurons. Due to different anatomic architectures, these projections would not exert the global, divergent influences of dopamine neurons but affect rather selected groups of neurons in a more specific and selective manner. The different learning systems would contribute different aspects to the acquisition of reward-directed behavior, although some overlap may occur. A dysfunction of dopamine neurons would lead to a deficient prediction error and result in slower and less efficient learning, whereas dysfunctions of the other reward systems may produce deficits in learning to select appropriate rewards.

### **Rewards as Explicit Goals of Behavior**

Besides properly identifying the available rewards, agents would benefit from anticipating rewards, evaluate the prospective gains and losses associated with each reward, make decisions about which reward to pursue, and initiate and control approach and consummatory behavior. A simple neuronal mechanism for reward prediction may consist of phasic responses to reward-predicting cues which have been associated with rewards through Pavlovian conditioning. Such reward-predicting responses are found in dopamine neurons and the other reward structures described above.

More sophisticated neuronal mechanisms of reward expectation consist of sustained activations which immediately precede an individual reward for several seconds and last until the reward is delivered. These activations may permit behavior-related neurons to access stored representations of future rewards. Sustained activations during reward expectation are found in classical reward structures, such as the striatum (Hikosaka et al., 1989; Apicella et al., 1992), subthalamic nucleus (Matsumura et al., 1992), and orbitofrontal cortex (Tremblay and Schultz, 1999), but also in more motor structures, such as the cortical supplementary eye field (Amador et al., 2000; Stuphorn et al., 2000). They discriminate between rewarded and unrewarded outcomes (Hollerman et al., 1998) and between different food and liquid rewards (Tremblay and Schultz, 1999; Hassani et al., 2001). The expectation-related activations discriminate in their temporal relationships between rewards and other predictable task events, such as movement-eliciting stimuli or instruction cues.

Expected rewards may serve as goals for voluntary behavior if information about the reward is present while behavioral reactions toward the reward are being prepared and executed (Dickinson and Balleine, 1994). Recent data suggest that neurons may integrate reward information into neuronal processes before the behavior leading to the reward occurs. Neurons in the striatum and dorsolateral and orbital frontal cortex show sustained activations during the preparation of arm or eye movements in delayed response tasks. The activations discriminate between rewarded and unrewarded outcomes (Hollerman et al., 1998; Kawagoe et al., 1998) and between different kinds and amounts of food and liquid rewards (Watanabe, 1996; Leon and Shadlen, 1999; Tremblay and Schultz, 1999; Hassani et al., 2001) (Figure 10, top). Similar reward-discriminating activations occur also during the execution of the movement. These data suggest that a predicted reward can influence neuronal activity related to the preparation and execution of the movement toward the reward and provide evidence on how neuronal mechanisms may guide behavior toward rewarding goals.

Reward expectations change systematically with experience during learning. The change can be demonstrated in tasks using rewarded and unrewarded outcomes. Behavioral reactions indicate that animals expect reward initially when new instruction stimuli are presented and later differentiate their expectations. Neurons in the striatum and orbitofrontal cortex show activity reflecting the expectation of reward during initial trials with novel stimuli, before any of them have been learned. Reward expectation activity becomes progres-

sively restricted to rewarded trials as animals learn the significance of the instructions (Tremblay et al., 1998; Tremblay and Schultz, 2000). Similar learning-related activity changes occur also during the preparation of the movement leading to the reward (Figure 10, bottom), as well as during its execution. Thus the reward expectation activity is not directly evoked by the explicit reward prediction of a novel instruction cue, which has not yet been established as a reward predictor. Rather the activity seems to reflect an internal expectation based on a representation which is retrieved in time by the novel, unspecific instruction cue and in content by the known learning context.

Many behavioral situations involve choices between different rewards, and agents make decisions that optimize the values of the rewards they seek. Greater effort is extended to rewards with higher values. Neurons in the parietal cortex show stronger task-related activity when monkeys chose larger, or more frequent, or smaller, or less frequent, rewards (Platt and Glimcher, 1999). Some neurons in the orbitofrontal cortex code reward preferences relative to other available rewards and thus may provide information to neuronal mechanisms that direct behavior toward rewards with higher value and away from rewards with lower value (Tremblay and Schultz, 1999). Together, these cortical neurons appear to code basic parameters of motivational value contributing to choice behavior. Changes in motivational value may lead to changes in reward-seeking behavior. Such changes involves neurons in the cingulate motor area of the frontal cortex which become selectively active when monkeys switch to a different movement after the current movement has produced less reward (Shima and Tanji, 1998). These activities do not occur when movements switch after an explicit trigger nor with reward reduction without movement switch. These data suggest a role in internal decision processes for movements based on the assessment of reward value. Although these examples are only a beginning in the study of brain mechanisms underlying reward-related decisions, they show nicely how information about future rewards could be used by the brain to make decisions about behavioral reactions leading to future reward.

### **Interaction between Different Reward Systems**

Given the central importance of rewards for survival, reproduction, and competitive gains, it may not be surprising that several specialized and only partly overlapping brain mechanisms have developed during evolution. These mechanisms may subserve the different needs for rewards in the large variety of behavioral situations in which individual agents live. The phasic dopamine reward alert signal provides only partial information about rewards but seems appropriate for inducing synaptic changes to prioritize the processing of reward information and learn even relatively complex behavioral reactions. In restricted behavioral situations, the dopamine signal alone may be sufficient for directing sequential behavior (Talwar et al., 2002) and for learning operant lever pressing for electrical brain stimulation (Wise, 1996a) and more complicated tasks in reinforcement models (Sutton and Barto, 1981).

However, many behavioral situations provide a larger variety of rewards, and the bottom-up dopamine reward signal would need to be supplemented by more specific

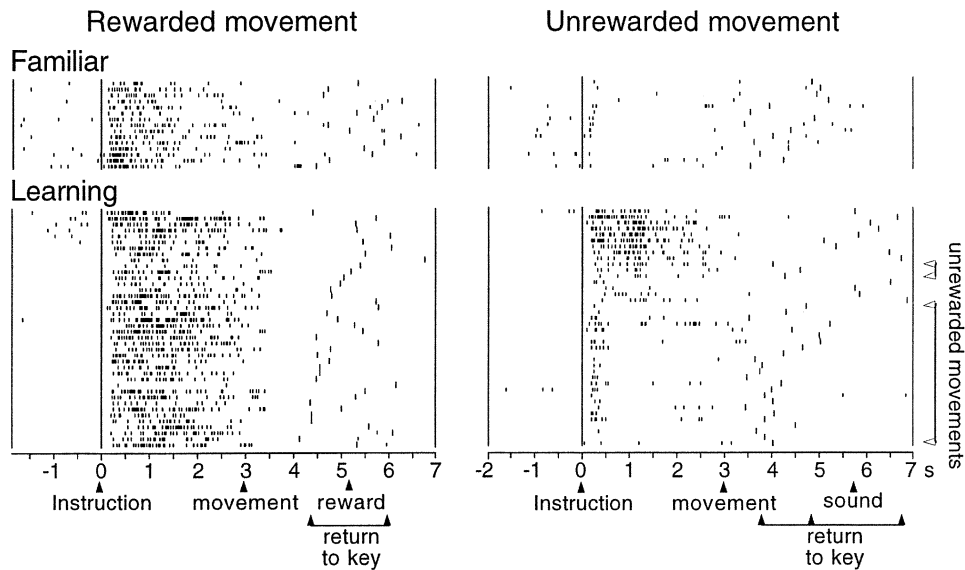


Figure 10. Sustained Activation in Caudate Neuron during the Preparation of Arm Movement and Its Adaptation during Learning (Top) During familiar performance, this caudate neuron shows a sustained response in rewarded movement trials, only a transient response in nonmovement trials (not shown), and no response in unrewarded trials. The presence or absence of reward and the required movement or nonmovement reaction are indicated by the initial instruction. Typically, the hand returns later to the resting key in rewarded as compared to unrewarded movement trials. (Bottom) During learning, the sustained response occurs initially also in unrewarded movement trials, which are performed with parameters of rewarded movements. The response disappears when movement parameters become typical for unrewarded movements (arrows to the right). Rewarded movement, nonmovement, and unrewarded movement trials alternated semirandomly during the experiment and are separated for analysis. Familiar and learning trials were performed in separate blocks. The sequence of trials is plotted chronologically from top to bottom, learning rasters beginning with the first presentations of new instructions. Reprinted from Tremblay et al., 1998. Copyright (1998) American Physiological Society.

reward information from the other reward systems. The striatum, frontal cortex, and amygdala process various components of the reward event that are useful for appropriately directing complex reward-directed behavior and making informed choices. This information may enter into cognitive representations of rewards and include the specific nature of individual rewards, the motivational value relative to other rewards, the compatibility of the reward with the requirements of the agent, and explicit representations about future rewards necessary for planning goal-directed behavior. None of this information is reflected in the dopamine reward responses. The information about the reward should be processed in close association with information concerning the approach and consumption of the reward. The mentioned conjoint behavior- and reward-related activities in the striatum (Hollerman et al., 1998; Kawagoe et al., 1998; Hassani et al., 2001) and frontal cortex (Watanabe, 1996) may constitute such a mechanism.

Thus situations with limited demands may engage only some of the neuronal reward processes, whereas more challenging situations may require to combine the fast bottom-up dopamine reward alert and teaching signal with slower top-down signals processing differential informations about the expected reward and about the ways to obtain it. The distinction between bottom-up dopamine and top-down striatal, cortical, and amygdalar reward signals may be too schematic to be biologically practical, but it illustrates how reward systems with different characteristics could work together to fulfill the needs of individual agents to use rewards as reinforcers for sophisticated behavior.

### Analogy to Theories of Emotion

The idea of a bottom-up dopamine reward alert signal and top-down reward discrimination and representation systems resembles to some extent one of the earlier theories on emotion, the Schachter-Singer Cognitive Arousal Theory (Schachter and Singer, 1962; see also LeDoux, 1998). According to this theory, emotion-inducing stimuli elicit states of arousal (via feedback from vegetative reactions) which do not contain specific informations about the eliciting event. It is the cognitive, well-differentiated assessment of the situation that allows agents to identify the event and generate a distinctive emotional feeling. An elaboration of this theory suggests that the arousal state should also enter into cognitive representations to be effective for generating emotions (Valins, 1966). According to this theory, emotions are generated by the combined action of unspecific arousal and specific cognitive interpretations.

Although the comparison is superficial at this stage and should not suggest a simple relationship between reward and emotional processes, it describes how the interactions of reward systems may correspond to this particular theory of emotion. The rapid dopamine reward alert signal may resemble the emotional, unspecific bottom-up arousal, although the dopamine signal does not seem to engage vegetative feedback, and the specific reward informations provided by striatal, cortical, and amygdalar systems resemble the cognitive emotional interpretations. The reward discrimination of responses in the orbitofrontal cortex, and possibly the amygdala, may correspond to the entry of arousal into cognitive representations (Valins, 1966). Although aspects of the

Schachter-Singer theory and the Valins amendment have been superseded by more recent variations (LeDoux, 1998), the general comparisons suggest that reward and emotions may share some basic underlying physiological mechanisms.

### Possible Neurophysiological Mechanisms of Psychostimulant Drugs

The increasing knowledge about neurophysiological reward mechanisms in the brain may help to provide us with a better understanding of the mechanisms of action of addictive drugs, in particular psychostimulants (cocaine and amphetamine). Psychopharmacological studies have used local, intracerebral drug injections to identify the dopamine system and its postsynaptic targets in the ventral striatum and frontal cortex as the critical structures (Wise and Hoffman, 1992; Robinson and Berridge, 1993; Wise, 2002 [this issue of *Neuron*]). This section describes several mechanisms by which addictive drugs may influence neuronal excitability and interact with the reward signal of dopamine neurons investigated in behaving animals. The description focuses on psychostimulants which have been most closely investigated.

#### Enhancement of Existing Reward Signal

Drugs of abuse may amplify the existing dopamine responses to natural rewards and reward-related environmental events. Cocaine and amphetamine appear to exert their addictive effects by blocking the reuptake of dopamine in the ventral striatum and frontal cortex (Giros et al., 1996). Blockade of dopamine reuptake leads to sustained increases in extracellular concentrations of dopamine. Within seconds and minutes after concentration changes, negative feedback and metabolism reduce release, synthesis, and concentration to a large extent (Church et al., 1987; Suaud-Chagny et al., 1995; Gonon, 1997).

Phasic increases in dopamine following neuronal responses to rewards would occur in fractions of seconds (see above) and thus are particularly amenable for enhancement by reuptake blockade, as they are faster and reach their peaks before most negative feedback becomes effective. This mechanism would lead to a massively enhanced chemical dopamine signal following primary rewards and reward-predicting stimuli. It would also increase dopamine concentrations after neuronal responses that for some reason are weak. The enhanced dopamine concentration could override depressant neuronal responses following stimuli resembling rewards, novel stimuli, and particularly salient stimuli that are frequent in everyday life.

Already before the dopamine message reaches the striatum and frontal cortex, addictive drugs may amplify the responses of dopamine neurons to incoming excitations. Amphetamine potentiates excitatory inputs to dopamine neurons (Tong et al., 1995) and removes long-term depression, leading to increased excitatory drive (Jones et al., 2000; Figure 11A). Cocaine increases long-term potentiation for a few days in dopamine neurons, but not in GABA neurons of the ventral midbrain (Ungless et al., 2001). Nicotine enhances glutamate-evoked long-term potentiation at dopamine neurons (Mansvelder and McGehee, 2000) and shifts dopamine neurons toward

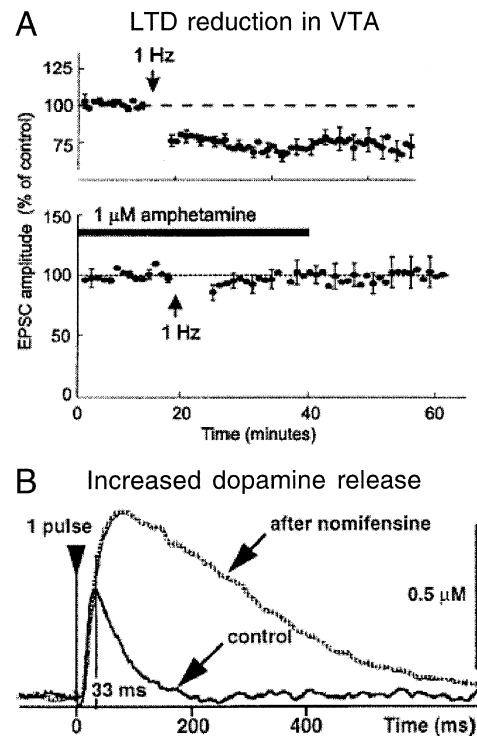


Figure 11. Enhancement of Phasic Dopamine Signal by Psychostimulant Drugs

(A) Amphetamine increases excitatory drive of dopamine neurons by reducing long-term depression (LTD) at ventral tegmental area (VTA) dopamine neurons in vitro. By contrast, the same dose of amphetamine fails to block LTD at hippocampal CA3 to CA1 synapses, demonstrating synapse specificity. EPSC, excitatory post-synaptic current. Reprinted from Jones et al., 2000. Copyright (2000) Society for Neuroscience. (B) The dopamine reuptake blocker nomifensine increases and prolongs dopamine release in the striatum induced by single action potentials on dopamine axons. Similar effects are obtained with cocaine and other dopamine reuptake blockers in the striatum and nucleus accumbens. Reprinted from Gonon, 1997. Copyright (1997) Society for Neuroscience.

higher levels of excitability (Mansvelder et al., 2002). These results suggest that dopamine neurons react stronger and more readily to excitatory inputs under the influence of these addictive drugs. Dopamine neurons would react stronger to rewards and release more dopamine in the target areas, such as the ventral striatum and frontal cortex.

The drug effects on dopamine perikarya and terminals may combine in a two-stage process to produce a particularly enhanced reward signal. Psychostimulant addictive drugs enhance the dopamine response to excitatory inputs at the level of cell bodies and amplify the enhanced response further in the target structures by blocking dopamine reuptake. The exaggerated reward prediction error message would constitute a very powerful focusing and teaching signal and produce modifications in synaptic transmission leading to substantial behavioral changes. One of the resulting effects may be an increase of movement-related activity in the striatum, as seen with the administration of amphetamine (West et al., 1997).



### ***Induction of Illusory Reward Signal***

A more hypothetical drug action may consist of producing an artificial neuronal reward signal in the absence of a real reward. It is long known that systemic injections of nicotine lead to activation of dopamine neurons (Lichtensteiger et al., 1976) and increase their burst firing (Grenhoff et al., 1986), which in turn induces a disproportionately strong dopamine release (Gonon, 1988; Figure 11B). If similar mechanisms hold for psychostimulants, the resulting artificial and illusory reward signal may resemble a true dopamine reward prediction error message and lead to focusing and plasticity effects on post-synaptic sites.

### ***Activation of Reward Detection Mechanisms***

It is possible that addictive drugs constitute rewards in their own right, activate the same neuronal mechanisms that detect and process natural rewards, and engage the existing reward mechanisms for influencing behavior. Drug rewards do not exert their influence on the brain through peripheral sensory receptors, as natural rewards do, and there is no need to extract the reward information from primary sensory inputs. Rather, drugs enter the brain directly via blood vessels and may exert their influence on behavior through more direct influences on reward-processing brain structures. The magnitude of their influences on the brain is not limited by the physiology of peripheral sensory receptors, as with natural rewards, but by the toxicity of the substances. This mechanism may permit much stronger drug influences on brain mechanisms than natural rewards.

Neurons in the ventral striatum show phasic responses to the injection of drugs and display activity during the expectation of drugs, similar to natural rewards (Carelli et al., 1993; Chang et al., 1994; Bowman et al., 1996; Peoples et al., 1998). Movement-related striatal activity appears to be influenced by drugs in a similar manner as by natural rewards, and some of the striatal neurons distinguish between drugs and natural rewards (West et al., 1997). These data may suggest that drug rewards are treated by these neurons as rewards in their own right. However, it is difficult to estimate how directly these responses in the striatum are related to the crucial role of dopamine in drug action.

### ***Outlook on Issues of Dopamine and Reward Matching between Dopamine Lesion Deficits and Neural Mechanisms***

The often close matches between lesion deficits and neuronal activity in the different areas of visual cortex do not seem to be paralleled in the basal ganglia and the dopamine system. Strikingly, Parkinsonian symptoms, which are due to degeneration of the nigrostriatal dopamine system, include movement deficits, but the activity of dopamine neurons in intact brains does not show clear covariations with movements (DeLong et al., 1983; Schultz et al., 1983a, 1983b). Although Parkinsonian patients and experimentally dopamine-depleted animals show some deficits in approach behavior and affective reactions (Canavan et al., 1989; Linden et al., 1990; Vriezen and Moscovitch, 1990; Sprengelmeyer et al., 1995; Knowlton et al., 1996), which might be due to the absence of a phasic reward signal, very few studies have used the particular stimulus characteristics that would

elicit a dopamine reward response in the intact animal. Only when such studies have been done can a more precise assessment be made of the use of the dopamine reward signal for approach behavior and different forms of learning. It is possible that reward-directed learning occurs also without a dopamine teaching signal, as rewards are also processed in structures outside of the dopamine system, but dopamine lesions should produce some deficits, in particular regarding the efficacy or speed of learning that might reflect the coding of prediction errors.

Another related issue concerns the relationship between the immediate, enhancing, and focusing effects of dopamine and the dopamine-mediated synaptic plasticity. Although the last years have seen a large increase in experimental evidence for both mechanisms at the membrane level, it is unclear which mechanisms would be operational in physiological situations of approach behavior and learning, and how these mechanisms may explain the lesion deficits in behavior. The goal would be to obtain a circuit understanding of the different synaptic functions of dopamine in the striatum and frontal cortex, in analogy to the current advances in cerebellar physiology (Sakurai, 1987; Linden and Connor, 1993; Kim et al., 1998; Medina et al., 2002; Attwell et al., 2002). Dopamine research may also profit from current investigations on relationships between long-term potentiation and behavioral learning in the hippocampus.

### ***Relationships to Learning Theory***

The evaluation of dopamine responses within the concepts and constraints of formal learning theories has proven to be interesting by revealing striking similarities between the activity of single neurons and the behavior of a whole organism (Waelti et al., 2001). Similarly interesting results were obtained by testing cerebellar learning functions in the same context (Kim et al., 1998). However, more tests need to be done to interpret, and to explicitly investigate, dopamine activity in this direction, and the current tests provide arguments that dopamine responses comply better with the Rescorla-Wagner rule, as compared to attentional learning rules such as the Pearce-Hall and Mackintosh rules. It would be necessary to test a much wider spectrum of behavioral situations, to see how far the match with the Rescorla-Wagner rule can go, whether some of the other learning rules may apply in certain situations, in particular in those not covered by the Rescorla-Wagner rule, and whether other neuronal activities than the phasic responses occur in these situations.

Having gained experience in interpreting dopamine activity and cerebellar function with formal learning theories, it would be interesting to test other neuronal systems with these rules. A candidate system is the nor-epinephrine neurons, which are probably the most investigated neurons with divergent projections besides the dopamine system and show a number of similarities with dopamine responses (Foote et al., 1980; Aston-Jones and Bloom, 1981; Rasmussen et al., 1986; Sara and Segal, 1991; Aston-Jones et al., 1994, 1997; Vankov et al., 1995). Other systems in which predictions from learning theories can be used to advance their understanding include the serotonin neurons and the cholinergic neurons of the nucleus basalis Meynert. This line of experimentation could be extended to the cerebral

cortex, the striatum, and the amygdala, as these structures process predictions of future outcomes and should, in some ways, be sensitive to prediction errors (see above) and therefore be understandable in terms of learning theories. A step in this direction has already been made in human imaging, the tested prediction being between drugs and diseases (Fletcher et al., 2001).

#### **More Extensive Models of Reinforcement Learning**

Current reinforcement models can be used to simulate biological learning. In particular, the temporal difference model (Sutton and Barto, 1981) can be mapped onto the anatomical organization of the basal ganglia and frontal cortex (Barto, 1995; Houk et al., 1995), and the characteristics of its teaching signal correspond in the important aspects to the dopamine prediction error signal (Montague et al., 1996; Suri and Schultz, 1999). Neural networks using such teaching signals can learn quite complex tasks, including backgammon and delayed responding (Tesauro, 1994; Suri and Schultz, 1999). However, these tasks are learned in the restricted environment defined by the person designing the model, and it would be interesting to see how these models can behave in real world situations containing many irrelevant events and abrupt state transitions. It might be necessary to include other reward systems, with their different functions, to model a more realistic agent that discriminates between different rewards, shows sensitivity to outcome devaluation, and uses reward representations to approach explicit goals of behavior (Dickinson and Balleine, 1994). It would be interesting to assess the location of different forms of synaptic plasticity in these models, to understand how important a dopamine-related plasticity would be, compared to plasticity further downstream. It would be interesting also to teach robots some behavioral tasks with dopamine reinforcement signals and see how far the actions can become sophisticated and general.

#### **From Reward Anticipation to Choice Behavior and Decision Making**

After understanding how some brain structures react to the delivery of rewards, it would be interesting to see how the reward information can be used for organizing behavioral output and making decisions about behavioral reactions between competing future outcomes. Recent theories have delineated the requirements for considering rewards as goals for voluntary behavior (Dickinson and Balleine, 1994). Recent experiments suggested that rewards are represented in neuronal activity during behavioral acts directed at obtaining these outcomes (Hollerman et al., 1998; Kawagoe et al., 1998; Hassani et al., 2001), and it would be interesting to see if neurons could actually code the contingencies between actions and outcomes and thus fulfill a basic criterion of goal-directed action. It is not surprising that rewards are coded in the cerebral cortex, and we would expect to see more reward coding in association with behavior-related activity in the different areas of frontal cortex. Only when the role of cortical reward processing has been more investigated can the question be evaluated to which extent reward systems may function in bottom-up and top-down modes similar to earlier ideas about emotions (Schachter and Singer, 1962; Valins, 1966;

LeDoux, 1998), or whether the reward systems are organized in a more refined way.

#### **Addictive Drugs**

The neurophysiological study of drug addiction in behaving animals has largely focused on psychostimulant drugs (cocaine, amphetamine), and it would be important to understand more about the processing of other rewarding drugs, such as heroin and nicotine. Do these drugs simply activate some dopamine neurons in a tonic way, or could they change the threshold or gain for inputs to dopamine neurons and thus produce an imaginary reward signal, similar to cocaine? How do opiates act on neurons with opiate receptors? What are the behavioral relationships of these opiate receptive neurons in the first place? Only when we have such information can we make an estimate on how important the dopamine system really is in opiate addiction.

#### **Acknowledgments**

My sincere thanks are due to the collaborators in the cited experiments and to Dr. Susan Jones for discussions on drug effects. The experimental work was conducted at the University of Fribourg and supported by the Swiss NSF, the Human Capital and Mobility and the Biomed 2 programs of the European Community via the Swiss Office of Education and Science, the James S. McDonnell Foundation, the Roche Research Foundation, the United Parkinson Foundation (Chicago), and the British Council. The writing of this review was supported by the Wellcome Trust.

#### **References**

- Abercrombie, E.D., Keefe, K.A., DiFrischia, D.S., and Zigmond, M.J. (1989). Differential effect of stress on in vivo dopamine release in striatum, nucleus accumbens, and medial frontal cortex. *J. Neurochem.* 52, 1655–1658.
- Ahlenius, S. (1974). Effects of low and high doses of L-dopa on the tetrabenazine or  $\alpha$ -methyltyrosine-induced suppression of behaviour in a successive discrimination task. *Psychopharmacologia* 39, 199–212.
- Amador, N., Schlag-Rey, M., and Schlag, J. (2000). Reward-predicting and reward-detecting neuronal activity in the primate supplementary eye field. *J. Neurophysiol.* 84, 2166–2170.
- Andén, N.E., Fuxe, K., Hamberger, B., and Hökfelt, T. (1966). A quantitative study on the nigro-neostriatal dopamine neurons. *Acta physiol. scand.* 67, 306–312.
- Apicella, P., Scarnati, E., Ljungberg, T., and Schultz, W. (1992). Neuronal activity in monkey striatum related to the expectation of predictable environmental events. *J. Neurophysiol.* 68, 945–960.
- Apicella, P., Legallet, E., and Trouche, E. (1997). Responses of tonically discharging neurons in the monkey striatum to primary rewards delivered during different behavioral states. *Exp. Brain Res.* 116, 456–466.
- Aston-Jones, G., and Bloom, F.E. (1981). Norepinephrine-containing locus coeruleus neurons in behaving rats exhibit pronounced responses to nonnoxious environmental stimuli. *J. Neurosci.* 1, 887–900.
- Aston-Jones, G., Rajkowski, J., Kubiak, P., and Alexinsky, T. (1994). Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task. *J. Neurosci.* 14, 4467–4480.
- Aston-Jones, G., Rajkowski, J., and Kubiak, P. (1997). Conditioned responses of monkey locus coeruleus neurons anticipate acquisition of discriminative behavior in a vigilance task. *Neuroscience* 80, 697–716.
- Attwell, P.J., Cooke, S.F., and Yeo, C.H. (2002). Cerebellar function in consolidation of a motor memory. *Neuron* 34, 1011–1020.
- Bao, S., Chan, V.T., and Merzenich, M.M. (2001). Cortical remodelling

- induced by activity of ventral tegmental dopamine neurons. *Nature* 412, 79–83.
- Barto, A.G. (1995). Adaptive critics and the basal ganglia. In *Models of Information Processing in the Basal Ganglia*, J.C.Houk, J.L.Davis, and D.G.Beiser, eds. (Cambridge, MA: MIT Press), pp. 215–232.
- Barto, A.G., Sutton, R.S., and Anderson, C.W. (1983). Neuronlike adaptive elements that can solve difficult learning problems. *IEEE Trans. Syst. Man and Cybern. SMC-13*, 834–846.
- Berger, B., Trotter, S., Verney, C., Gaspar, P., and Alvarez, C. (1988). Regional and laminar distribution of the dopamine and serotonin innervation in the macaque cerebral cortex: a radioautographic study. *J. Comp. Neurol.* 273, 99–119.
- Berns, G.S., Cohen, J.D., and Mintun, M.A. (1997). Brain regions responsive to novelty in the absence of awareness. *Science* 276, 1272–1275.
- Blond, O., Crepel, F., and Otani, S. (2002). Long-term potentiation in rat prefrontal slices facilitated by phased application of dopamine. *Eur. J. Pharmacol.* 438, 115–116.
- Bowman, E.M., Aigner, T.G., and Richmond, B.J. (1996). Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. *J. Neurophysiol.* 75, 1061–1073.
- Brembs, B., Lorenzetti, F.D., Reyes, F.D., Baxter, D.A., and Byrne, J.H. (2002). Operant learning in aplysia: neuronal correlates and mechanisms. *Science* 296, 1706–1709.
- Brown, J.R., and Arbuthnot, G.W. (1983). The electrophysiology of dopamine (D<sub>2</sub>) receptors: A study of the actions of dopamine on corticostriatal transmission. *Neuroscience* 10, 349–355.
- Brozoski, T.J., Brown, R.M., Rosvold, H.E., and Goldman, P.S. (1979). Cognitive deficit caused by regional depletion of dopamine in prefrontal cortex of rhesus monkey. *Science* 205, 929–932.
- Burns, R.S., Chieh, C.C., Markey, S.P., Ebert, M.H., Jacobowitz, D.M., and Kopin, I.J. (1983). A primate model of parkinsonism, selective destruction of dopaminergic neurons in the pars compacta of the substantia nigra by N-methyl-4-phenyl-1,2,3,6-tetrahydropyridine. *Proc. Natl. Acad. Sci. USA* 80, 4546–4550.
- Calabresi, P., Maj, R., Pisani, A., Mercuri, N.B., and Bernardi, G. (1992). Long-term synaptic depression in the striatum: physiological and pharmacological characterization. *J. Neurosci.* 12, 4224–4233.
- Calabresi, P., Saiardi, A., Pisani, A., Baik, J.H., Centonze, D., Mercuri, N.B., Bernardi, G., and Borelli, E. (1997). Abnormal synaptic plasticity in the striatum of mice lacking dopamine D2 receptors. *J. Neurosci.* 17, 4536–4544.
- Canavan, A.G.M., Passingham, R.E., Marsden, C.D., Quinn, N., Wyke, M., and Polkey, C.E. (1989). The performance on learning tasks of patients in the early stages of Parkinson's disease. *Neuropsychologia* 27, 141–156.
- Carelli, R.M., King, V.C., Hampson, R.E., and Deadwyler, S.A. (1993). Firing patterns of nucleus accumbens neurons during cocaine self-administration. *Brain Res.* 626, 14–22.
- Carr, D.B., and Sesack, S.R. (1996). Hippocampal afferents to the rat prefrontal cortex: synaptic targets and relation to dopamine terminals. *J. Comp. Neurol.* 369, 1–15.
- Cepeda, C., Buchwald, N.A., and Levine, M.S. (1993). Neuromodulatory actions of dopamine in the neostriatum are dependent upon the excitatory amino acid receptor subtypes activated. *Proc. Natl. Acad. Sci. USA* 90, 9576–9580.
- Chang, J.Y., Sawyer, S.F., Lee, R.S., and Woodward, D.J. (1994). Electrophysiological and pharmacological evidence for the role of the nucleus accumbens in cocaine self-administration in freely moving rats. *J. Neurosci.* 14, 1224–1244.
- Chergui, K., Suaud-Chagny, M.F., and Gonon, F. (1994). Nonlinear relationship between impulse flow, dopamine release and dopamine elimination in the rat brain in vivo. *Neuroscience* 62, 641–645.
- Chesselet, M.F. (1984). Presynaptic regulation of neurotransmitter release in the brain: facts and hypothesis. *Neuroscience* 12, 347–375.
- Church, W.H., Justice, J.B., Jr., and Byrd, L.D. (1987). Extracellular dopamine in rat striatum following uptake inhibition by cocaine, nomifensine and bntropine. *Eur. J. Pharmacol.* 139, 345–348.
- Contreras-Vidal, J.L., and Schultz, W. (1999). A predictive reinforcement model of dopamine neurons for learning approach behavior. *J. Comput. Neurosci.* 6, 191–214.
- Critchley, H.G., and Rolls, E.T. (1996). Hunger and satiety modify the responses of olfactory and visual neurons in the primate orbitofrontal cortex. *J. Neurophysiol.* 75, 1673–1686.
- Dahl, D., and Sarvey, J.M. (1989). Norepinephrine induces pathway-specific long-lasting potentiation and depression in the hippocampal dentate gyrus. *Proc. Natl. Acad. Sci.* 86, 4776–4780.
- DeLong, M.R., Crutcher, M.D., and Georgopoulos, A.P. (1983). Relations between movement and single cell discharge in the substantia nigra of the behaving monkey. *J. Neurosci.* 3, 1599–1606.
- Dickinson, A. (1980). *Contemporary Animal Learning Theory* (Cambridge, United Kingdom: Cambridge University Press).
- Dickinson, A. (2001). Causal learning: an associative analysis. *Q. J. Exp. Psychol.* 54B, 3–25.
- Dickinson, A., and Balleine, B. (1994). Motivational control of goal-directed action. *Anim. Learn. Behav.* 22, 1–18.
- Ding, L., and Perkel, D.J. (2002). Dopamine modulates excitability of spiny neurons in the avian basal ganglia. *J. Neurosci.* 22, 5210–5218.
- Doherty, M.D., and Gratton, A. (1992). High-speed chronoamperometric measurements of mesolimbic and nigrostriatal dopamine release associated with repeated daily stress. *Brain Res.* 586, 295–302.
- Doucet, G., Descarries, L., and Garcia, S. (1986). Quantification of the dopamine innervation in adult rat neostriatum. *Neuroscience* 19, 427–445.
- Dugast, C., Suaud-Chagny, M.F., and Gonon, F. (1994). Continuous in vivo monitoring of evoked dopamine release in the rat nucleus accumbens by amperometry. *Neuroscience* 62, 647–654.
- Eccles, J.C., Ito, M., and Szentagothai, J. (1967). *The Cerebellum as a Neuronal Machine* (Berlin: Springer).
- Elliott, R., Sahakian, B.J., Matthews, K., Bannerjee, A., Rimmer, J., and Robbins, T.W. (1997). Effects of methylphenidate on spatial working memory and planning in healthy young adults. *Psychopharmacology (Berl.)* 131, 196–206.
- Fagg, A.H. (1993). Reinforcement learning for robotic reaching and grasping. In *New Perspectives in the Control of the Reach to Grasp Movement*, K.M.B. Bennet and U.Castiello, eds. (North Holland), pp. 281–308.
- Falkenstein, M., Hohnsbein, J., Hoormann, J., and Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalogr. Clin. Neurophysiol.* 78, 447–455.
- Fibiger, H.C., and Phillips, A.G. (1986). Reward, motivation, cognition, psychobiology of mesotelencephalic dopamine systems. In *Handbook of Physiology—The Nervous System IV*, F.E. Bloom, ed. (Baltimore, MD: Williams and Wilkins), pp. 647–675.
- Filion, M., Tremblay, L., and Bédard, P.J. (1988). Abnormal influences of passive limb movement on the activity of globus pallidus neurons in parkinsonian monkey. *Brain Res.* 444, 165–176.
- Flaherty, C.F. (1996). *Incentive Relativity* (Cambridge, United Kingdom: Cambridge University Press).
- Fletcher, P.C., Anderson, J.M., Shanks, D.R., Honey, R., Carpenter, T.A., Donovan, T., Papadakis, N., and Bullmore, E.T. (2001). Responses of human frontal cortex to surprising events are predicted by formal associative learning theory. *Nat. Neurosci.* 4, 1043–1048.
- Foote, S.L., Aston-Jones, G., and Bloom, F.E. (1980). Impulse activity of locus coeruleus neurons in awake rats and monkeys is a function of sensory stimulation and arousal. *Proc. Natl. Acad. Sci. USA* 77, 3033–3037.
- Freedman, R., Hoffer, B.J., Woodward, D.J., and Puro, D. (1977). Interaction of norepinephrine with cerebellar activity evoked by mossy and climbing fibers. *Exp. Neurol.* 55, 269–288.
- Freund, T.F., Powell, J.F., and Smith, A.D. (1984). Tyrosine hydroxylase-immunoreactive boutons in synaptic contact with identified striatonigral neurons, with particular reference to dendritic spines. *Neuroscience* 13, 1189–1215.
- Friston, K.J., Tononi, G., Reeke, G.N., Jr., Sporns, O., and Edelman,

- G.M. (1994). Value-dependent selection in the brain: simulation in a synthetic neural model. *Neuroscience* 59, 229–243.
- Gemba, H., Sakai, K., and Brooks, V.B. (1986). “Error” potentials in limbic cortex (anterior cingulate area 24) of monkeys during motor learning. *Neurosci. Lett.* 70, 223–227.
- German, D.C., Dubach, M., Askari, S., Speciale, S.G., and Bowden, D.M. (1988). 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)-induced parkinsonian syndrome in macaca fascicularis: which mid-brain dopaminergic neurons are lost? *Neuroscience* 24, 161–174.
- Gilbert, P.F.C., and Thach, W.T. (1977). Purkinje cell activity during motor learning. *Brain Res.* 128, 309–328.
- Giros, B., Jaber, M., Jones, S.R., Wightman, R.M., and Caron, M.G. (1996). Hyperlocomotion and indifference to cocaine and amphetamine in mice lacking the dopamine transporter. *Nature* 379, 606–612.
- Goldman-Rakic, P.S., Leranth, C., Williams, M.S., Mons, N., and Geffard, M. (1989). Dopamine synaptic complex with pyramidal neurons in primate cerebral cortex. *Proc. Natl. Acad. Sci. USA* 86, 9015–9019.
- Gonon, F. (1988). Nonlinear relationship between impulse flow and dopamine released by rat midbrain dopaminergic neurons as studied by *in vivo* electrochemistry. *Neuroscience* 24, 19–28.
- Gonon, F. (1997). Prolonged and extrasynaptic excitatory action of dopamine mediated by D1 receptors in the rat striatum *in vivo*. *J. Neurosci.* 17, 5972–5978.
- Grenhoff, J., Aston-Jones, G., and Svensson, T.H. (1986). Nicotinic effects on the firing pattern of midbrain dopamine neurons. *Acta Physiol. Scand.* 128, 351–358.
- Groves, P.M., Garcia-Munoz, M., Linder, J.C., Manley, M.S., Martone, M.E., and Young, S.J. (1995). Elements of the intrinsic organization and information processing in the neostriatum. In *Models of Information Processing in the Basal Ganglia*, J.C. Houk, J.L. Davis, and D.G. Beiser, eds. (Cambridge, MA: MIT Press), pp. 51–96.
- Gullapalli, V., Barto, A.G., and Grupen, R.A. (1994). Learning admittance mapping for force-guided assembly. In *Proceedings of the 1994 International Conference on Robotics and Automation* (Los Alamitos, CA: Computer Society Press), pp. 2633–2638.
- Gurden, H., Tassin, J.P., and Jay, T.M. (1999). Integrity of the mesocortical DA system is necessary for complete expression of *in vivo* hippocampal-prefrontal cortex long-term potentiation. *Neuroscience* 94, 1019–1027.
- Gurden, H., Takita, M., and Jay, T.M. (2000). Essential role of D1 but not D2 receptors in the NMDA receptor-dependent long-term potentiation at hippocampal-prefrontal cortex synapses *in vivo*. *J. Neurosci.* 20, RC106.
- Harlow, H.F. (1949). The formation of learning sets. *Psychol. Rev.* 56, 51–65.
- Hassani, O.K., Cromwell, H.C., and Schultz, W. (2001). Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *J. Neurophysiol.* 85, 2477–2489.
- Hernandez-Lopez, S., Bargas, J., Surmeier, D.J., Reyes, A., and Galarraga, E. (1997). D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type  $Ca^{2+}$  conductance. *J. Neurosci.* 17, 3334–3342.
- Herrera-Marschitz, M., You, Z.B., Gojny, M., Meana, J.J., Silveira, R., Godukhin, O.V., Chen, Y., Espinoza, S., Pettersson, E., Loidl, C.F., et al. (1996). On the origin of extracellular glutamate levels monitored in the basal ganglia of the rat by *in vivo* microdialysis. *J. Neurochem.* 66, 1726–1735.
- Hikosaka, O., Sakamoto, M., and Usui, S. (1989). Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. *J. Neurophysiol.* 61, 814–832.
- Hirsch, J.C., and Crepel, F. (1990). Use-dependent changes in synaptic efficacy in rat prefrontal neurons *in vitro*. *J. Physiol.* 427, 31–49.
- Hollerman, J.R., and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* 1, 304–309.
- Hollerman, J.R., Tremblay, L., and Schultz, W. (1998). Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J. Neurophysiol.* 80, 947–963.
- Horvitz, J.C. (2000). Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* 96, 651–656.
- Houk, J.C., Adams, J.L., and Barto, A.G. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In *Models of Information Processing in the Basal Ganglia*, J.C. Houk, J.L. Davis, and D.G. Beiser, eds. (Cambridge, MA: MIT Press), pp. 249–270.
- Houk, J.C., Buckingham, J.T., and Barto, A.G. (1996). Models of the cerebellum and motor learning. *Behav. Brain Sci.* 19, 368–383.
- Ito, M. (1989). Long-term depression. *Annu. Rev. Neurosci.* 12, 85–102.
- Jones, S., Kornblum, J.L., and Kauer, J.A. (2000). Amphetamine blocks long-term synaptic depression in the ventral tegmental area. *J. Neurosci.* 20, 5575–5580.
- Jörntell, H., and Ekerot, C.-F. (2002). Reciprocal bidirectional plasticity of parallel fiber receptive fields in cerebellar Purkinje cells and their afferent interneurons. *Neuron* 34, 797–806.
- Kamin, L.J. (1969). Selective association and conditioning. In *Fundamental Issues in Instrumental Learning*, N.J. Mackintosh and W.K. Honig, eds. (Halifax, Canada: Dalhousie University Press), pp. 42–64.
- Katsuki, H., Izumi, Y., and Zorumski, C.F. (1997). Noradrenergic regulation of synaptic plasticity in the hippocampal CA1 region. *J. Neurophysiol.* 77, 3013–3020.
- Kawagoe, R., Takikawa, Y., and Hikosaka, O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. *Nat. Neurosci.* 1, 411–416.
- Kawato, M., and Gomi, H. (1992). The cerebellum and VOR/OKR learning models. *Trends Neurosci.* 15, 445–453.
- Kerr, J.N., and Wickens, J.R. (2001). Dopamine D-1/D-5 receptor activation is required for long-term potentiation in the rat neostriatum *in vitro*. *J. Neurophysiol.* 85, 117–124.
- Kim, J.J., Krupa, D.J., and Thompson, R.F. (1998). Inhibitory cerebello-olivary projections and blocking effect in classical conditioning. *Science* 279, 570–573.
- Kitazawa, S., Kimura, T., and Yin, P.B. (1998). Cerebellar complex spikes encode both destinations and errors in arm movement. *Nature* 392, 494–497.
- Knowlton, B.J., Mangels, J.A., and Squire, L.R. (1996). A neostriatal habit learning system in humans. *Science* 273, 1399–1402.
- LeDoux, J. (1998). *The Emotional Brain* (New York: Simon & Schuster).
- Leon, M.I., and Shadlen, M.N. (1999). Effect of expected reward magnitude on the responses of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron* 24, 415–425.
- Lichtensteiger, W., Felix, D., Lienhart, R., and Hefti, F. (1976). A quantitative correlation between single unit activity and fluorescence intensity of dopamine neurones in zona compacta of substantia nigra, as demonstrated under the influence of nicotine and physostigmine. *Brain Res.* 117, 85–103.
- Linden, A., Bracke-Tollmicht, R., Lutzenberger, W., Canavan, A.G.M., Scholz, E., Diener, H.C., and Birbaumer, N. (1990). Slow cortical potentials in Parkinsonian patients during the course of an associative learning task. *J. Psychophysiol.* 4, 145–162.
- Linden, D.J., and Connor, J.A. (1993). Cellular mechanisms of long-term depression in the cerebellum. *Curr. Opin. Neurobiol.* 3, 401–406.
- Ljungberg, T., Apicella, P., and Schultz, W. (1991). Responses of monkey midbrain dopamine neurons during delayed alternation performance. *Brain Res.* 586, 337–341.
- Ljungberg, T., Apicella, P., and Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol.* 67, 145–163.
- Llinas, R., and Welsh, J.P. (1993). On the cerebellum and motor learning. *Curr. Opin. Neurobiol.* 3, 958–965.
- Lovibond, P.F. (1983). Facilitation of instrumental behavior by a Pav-

- lovian appetitive conditioned stimulus. *J. Exp. Psychol. Anim. Behav. Process.* 9, 225–247.
- Louilot, A., LeMoal, M., and Simon, H. (1986). Differential reactivity of dopaminergic neurons in the nucleus accumbens in response to different behavioral situations. An in vivo voltammetric study in free moving rats. *Brain Res.* 397, 395–400.
- Lynd-Balta, E., and Haber, S.N. (1994). Primate striatonigral projections: a comparison of the sensorimotor-related striatum and the ventral striatum. *J. Comp. Neurol.* 345, 562–578.
- Mackintosh, N.J. (1975). A theory of attention: variations in the associability of stimulus with reinforcement. *Psychol. Rev.* 82, 276–298.
- Mansvelder, H.D., and McGehee, D.S. (2000). Long-term potentiation of excitatory inputs to brain reward areas by nicotine. *Neuron* 27, 349–357.
- Mansvelder, H.D., Keith, J.R., and McGehee, D.S. (2002). Synaptic mechanisms underlie nicotine-induced excitability of brain areas. *Neuron* 33, 905–919.
- Marr, D. (1969). A theory of cerebellar cortex. *J. Physiol.* 202, 437–470.
- Martin, I., and Levey, A.B. (1991). Blocking observed in human eyelid conditioning. *Q. J. Exp. Psychol.* 43B, 233–255.
- Matsumura, M., Kojima, J., Gardiner, T.W., and Hikosaka, O. (1992). Visual and oculomotor functions of monkey subthalamic nucleus. *J. Neurophysiol.* 67, 1615–1632.
- Medina, J.F., Noes, W.L., and Mauk, M.D. (2002). Inhibition of climbing fibers is a signal for the extinction of conditioned eyelid responses. *Nature* 416, 330–333.
- Miller, J.D., Sanghera, M.K., and German, D.C. (1981). Mesencephalic dopaminergic unit activity in the behaviorally conditioned rat. *Life Sci.* 29, 1255–1263.
- Mirenowicz, J., and Schultz, W. (1994). Importance of unpredictability for reward responses in primate dopamine neurons. *J. Neurophysiol.* 72, 1024–1027.
- Mirenowicz, J., and Schultz, W. (1996). Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* 379, 449–451.
- Mitchell, S.J., Richardson, R.T., Baker, F.H., and DeLong, M.R. (1987). The primate globus pallidus: neuronal activity related to direction of movement. *Exp. Brain Res.* 68, 491–505.
- Montague, P.R., Dayan, P., Person, C., and Sejnowski, T.J. (1995). Bee foraging in uncertain environments using predictive hebbian learning. *Nature* 377, 725–728.
- Montague, P.R., Dayan, P., and Sejnowski, T.J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947.
- Murphy, B.L., Arnsten, A.F., Goldman-Rakic, P.S., and Roth, R.H. (1996). Increased dopamine turnover in the prefrontal cortex impairs spatial working memory performance in rats and monkeys. *Proc. Natl. Acad. Sci. USA* 93, 1325–1329.
- Nargeot, R., Baxter, D.A., Patterson, G.W., and Byrne, J.H. (1999). Dopaminergic synapses mediate neuronal changes in an analogue of operant conditioning. *J. Neurophysiol.* 81, 1983–1987.
- Niki, H., and Watanabe, M. (1979). Prefrontal and cingulate unit activity during timing behavior in the monkey. *Brain Res.* 171, 213–224.
- Nishijo, H., Ono, T., and Nishino, H. (1988). Topographic distribution of modality-specific amygdalar neurons in alert monkey. *J. Neurosci.* 8, 3556–3569.
- Nishino, H., Ono, T., Muramoto, K.I., Fukuda, M., and Sasaki, K. (1987). Neuronal activity in the ventral tegmental area (VTA) during motivated bar press feeding behavior in the monkey. *Brain Res.* 413, 302–313.
- Nobre, A.C., Coull, J.T., Frith, C.D., and Mesulam, M.M. (1999). Orbitofrontal cortex is activated during breaches of expectation in tasks of visual attention. *Nat. Neurosci.* 2, 11–12.
- Ojakangas, C.L., and Ebner, T.J. (1992). Purkinje cell complex and simple spike changes during a voluntary arm movement learning task in the monkey. *J. Neurophysiol.* 68, 2222–2236.
- Otani, S., Blond, O., Desce, J.M., and Crepel, F. (1998). Dopamine facilitates long-term depression of glutamatergic transmission in rat prefrontal cortex. *Neuroscience* 85, 669–676.
- Otmakhova, N.A., and Lisman, J.E. (1996). D1/D5 dopamine receptor activation increases the magnitude of early long-term potentiation at CA1 hippocampal synapses. *J. Neurosci.* 16, 7478–7486.
- Otmakhova, N.A., and Lisman, J.E. (1998). D1/D5 dopamine receptors inhibit depotentiation at CA1 synapses via cAMP-dependent mechanism. *J. Neurosci.* 18, 1273–1279.
- Pearce, J.M., and Hall, G. (1980). A model for Pavlovian conditioning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* 87, 532–551.
- Peoples, L.L., Gee, F., Bibi, R., and West, M.O. (1998). Phasic firing time locked to cocaine self-infusion and locomotion: dissociable firing patterns of single nucleus accumbens neurons in the rat. *J. Neurosci.* 18, 7588–7598.
- Percheron, G., Francois, C., Yelnik, J., and Fenelon, G. (1989). The primate nigro-striato-pallido-nigral system. Not a mere loop. In *Neural Mechanisms in Disorders of Movement*, A.R. Crossman and M.A. Sambrook, eds. (London: John Libbey), pp. 103–109.
- Platt, M.L., and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.
- Poirier, L.J. (1960). Experimental and histological study of midbrain dyskinesias. *J. Neurophysiol.* 23, 534–551.
- Rasmussen, K., Morilak, D.A., and Jacobs, B.L. (1986). Single unit activity of locus coeruleus neurons in the freely moving cat. I. During naturalistic behaviors and in response to simple and complex stimuli. *Brain Res.* 371, 324–334.
- Redgrave, P., Prescott, T.J., and Gurney, K. (1999). Is the short-latency dopamine response too short to signal reward? *Trends Neurosci.* 22, 146–151.
- Rescorla, R.A., and Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical Conditioning II: Current Research and Theory*, A.H. Black and W.F. Prokasy, eds. (New York: Appleton Century Crofts), pp. 64–99.
- Reynolds, J.N.J., Hyland, B.I., and Wickens, J.R. (2001). A cellular mechanism of reward-related learning. *Nature* 413, 67–70.
- Richardson, R.T., and DeLong, M.R. (1986). Nucleus basalis of Meynert neuronal activity during a delayed response task in monkey. *Brain Res.* 399, 364–368.
- Richardson, R.T., and DeLong, M.R. (1990). Context-dependent responses of primate nucleus basalis neurons in a go/no-go task. *J. Neurosci.* 10, 2528–2540.
- Richfield, E.K., Pennney, J.B., and Young, A.B. (1989). Anatomical and affinity state comparisons between dopamine D1 and D2 receptors in the rat central nervous system. *Neuroscience* 30, 767–777.
- Robbins, T.W., and Everitt, B.J. (1996). Neurobehavioural mechanisms of reward and motivation. *Curr. Opin. Neurobiol.* 6, 228–236.
- Robinson, T.E., and Berridge, K.C. (1993). The neural basis for drug craving: an incentive-sensitization theory of addiction. *Brain Res. Brain Res. Rev.* 18, 247–291.
- Rolls, E.T., Critchley, H.D., Browning, A.S., Hernadi, I., and Lenard, L. (1999). Responses to the sensory properties of fat of neurons in the primate orbitofrontal cortex. *J. Neurosci.* 19, 1532–1540.
- Romo, R., and 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.
- Rosenkilde, C.E., Bauer, R.H., and Fuster, J.M. (1981). Single cell activity in ventral prefrontal cortex of behaving monkeys. *Brain Res.* 209, 375–394.
- Rosenkranz, J.A., and Grace, A.A. (2002). Dopamine-mediated modulation of odour-evoked amygdala potentials during pavlovian conditioning. *Nature* 417, 282–287.
- Sakurai, M. (1987). Synaptic modification of parallel fibre-Purkinje

- cell transmission in in vitro guinea-pig cerebellar slices. *J. Physiol.* 394, 463–480.
- Sara, S.J., and Segal, M. (1991). Plasticity of sensory responses of locus coeruleus neurons in the behaving rat: implications for cognition. *Prog. Brain Res.* 88, 571–585.
- Sardo, P., Ravel, S., Legallet, E., and Apicella, P. (2000). Influence of the predicted time of stimuli eliciting movements on responses to tonically active neurons in the monkey striatum. *Eur. J. Neurosci.* 12, 1801–1816.
- Sawaguchi, T., and Goldman-Rakic, P.S. (1991). D1 dopamine receptors in prefrontal cortex: involvement in working memory. *Science* 251, 947–950.
- Schachter, S., and Singer, J.E. (1962). Cognitive, social and physiological determinants of emotional state. *Psychol. Rev.* 69, 379–399.
- Schultz, W. (1986). Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey. *J. Neurophysiol.* 56, 1439–1462.
- Schultz, W. (1992). Activity of dopamine neurons in the behaving primate. *Semin. Neurosci.* 4, 129–138.
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27.
- Schultz, W., and Romo, R. (1987). Responses of nigrostriatal dopamine neurons to high intensity somatosensory stimulation in the anesthetized monkey. *J. Neurophysiol.* 57, 201–217.
- Schultz, W., and Romo, R. (1990). Dopamine neurons of the monkey midbrain: Contingencies of responses to stimuli eliciting immediate behavioral reactions. *J. Neurophysiol.* 63, 607–624.
- Schultz, W., Aebischer, P., and Ruffieux, A. (1983a). The encoding of motor acts by the substantia nigra. *Exp. Brain Res. Suppl.* 7, 171–180.
- Schultz, W., Ruffieux, A., and Aebischer, P. (1983b). The activity of pars compacta neurons of the monkey substantia nigra in relation to motor activation. *Exp. Brain Res.* 51, 377–387.
- Schultz, W., Apicella, P., and 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., Dayan, P., and Montague, R.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.
- Sears, L.L., and Steinmetz, J.E. (1991). Dorsal accessory inferior olive activity diminishes during acquisition of the rabbit classically conditioned eyelid response. *Brain Res.* 545, 114–122.
- Sessler, F.M., Liu, W., Kirifides, M.L., Mouradian, R.D., Lin, R.C., and Waterhouse, B.D. (1995). Noradrenergic enhancement of GABA-induced input resistance changes in layer V regular spiking pyramidal neurons of rat somatosensory cortex. *Brain Res.* 675, 171–182.
- Shima, K., and Tanji, J. (1998). Role for cingulate motor area cells in voluntary movement selection based on reward. *Science* 282, 1335–1338.
- Simon, H., Scatton, B., and LeMoal, M. (1980). Dopaminergic A10 neurons are involved in cognitive functions. *Nature* 286, 150–151.
- Smith, A.D., and Bolam, J.P. (1990). The neural network of the basal ganglia as revealed by the study of synaptic connections of identified neurones. *Trends Neurosci.* 13, 259–265.
- Smith, Y., Bennett, B.D., Bolam, J.P., Parent, A., and Sadikot, A.F. (1994). Synaptic relationships between dopaminergic afferents and cortical or thalamic input in the sensorimotor territory of the striatum in monkey. *J. Comp. Neurol.* 344, 1–19.
- Sprengelmeyer, R., Canavan, A.G.M., Lange, H.W., and Hömberg, V. (1995). Associative learning in degenerative neostriatal disorders: contrasts in explicit and implicit remembering between Parkinson's and Huntington's disease patients. *Mov. Disord.* 10, 85–91.
- Stuphorn, V., Taylor, T.L., and Schall, J.D. (2000). Performance monitoring by the supplementary eye field. *Nature* 408, 857–860.
- Suaud-Chagny, M.F., Dugast, C., Chergui, K., Msghina, M., and Gonon, F. (1995). Uptake of dopamine released by impulse flow in the rat mesolimbic and striatal system in vivo. *J. Neurochem.* 65, 2603–2611.
- Suri, R.E., and Schultz, W. (1998). Learning of sequential movements by neural network model with dopamine-like reinforcement signal. *Exp. Brain Res.* 121, 350–354.
- Suri, R.E., and Schultz, W. (1999). A neural network with dopamine-like reinforcement signal that learns a spatial delayed response task. *Neuroscience* 91, 871–890.
- Sutton, R.S., and Barto, A.G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychol. Rev.* 88, 135–170.
- Sutton, R.S., and Barto, A.G. (1990). Time-derivative models of Pavlovian reinforcement. In *Learning and Computational Neuroscience: Foundations of Adaptive Networks*, M. Gabriel and J. Moore, eds. (Cambridge, MA: MIT Press), pp. 497–537.
- Talwar, S.K., Xu, S., Hawley, E.S., Weiss, S.A., Moxon, K.A., and Chapin, J.K. (2002). Rat navigation guided by remote control. *Nature* 417, 37–38.
- Tang, K.C., Low, M.J., Grandy, D.K., and Lovinger, D.M. (2001). Dopamine-dependent synaptic plasticity in striatum during in vivo development. *Proc. Natl. Acad. Sci. USA* 98, 1255–1260.
- Tesauro, G. (1994). TD-Gammon, a self-teaching backgammon program, achieves master-level play. *Neural Comput.* 6, 215–219.
- Thompson, R.F., and Gluck, M.A. (1991). Brain substrates of basic associative learning and memory. In *Perspectives on Cognitive Neuroscience*, R.G. Lister and H.J. Weingartner, eds. (New York: Oxford University Press), pp. 25–45.
- Thorpe, S.J., Rolls, E.T., and Maddison, S. (1983). The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp. Brain Res.* 49, 93–115.
- Toan, D.L., and Schultz, W. (1985). Responses of rat pallidum cells to cortex stimulation and effects of altered dopaminergic activity. *Neuroscience* 15, 683–694.
- Tong, Z.-Y., Overton, P.G., and Clark, D. (1995). Chronic administration of (+) amphetamine alters the reactivity of midbrain dopaminergic neurons to prefrontal cortex stimulation in the rat. *Brain Res.* 674, 63–74.
- Tremblay, L., and Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature* 398, 704–708.
- Tremblay, L., and Schultz, W. (2000). Modifications of reward expectation-related neuronal activity during learning in primate orbitofrontal cortex. *J. Neurophysiol.* 83, 1877–1885.
- Tremblay, L., Hollerman, J.R., and Schultz, W. (1998). Modifications of reward expectation-related neuronal activity during learning in primate striatum. *J. Neurophysiol.* 80, 964–977.
- Umeno, M.M., and Goldberg, M.E. (1997). Spatial processing in the monkey frontal eye field. I. Predictive visual responses. *J. Neurophysiol.* 78, 1373–1383.
- Ungless, M.A., Whistler, J.L., Malenka, R.C., and Bonci, A. (2001). Single cocaine exposure in vivo induces long-term potentiation in dopamine neurons. *Nature* 411, 583–587.
- Valins, S. (1966). Cognitive effects of false heart-rate feedback. *J. Pers. Soc. Psychol.* 4, 400–408.
- Vankov, A., Hervé-Minvielle, A., and Sara, S.J. (1995). Response to novelty and its rapid habituation in locus coeruleus neurons of the freely exploring rat. *Eur. J. Neurosci.* 7, 1180–1187.
- Vriezen, E.R., and Moscovitch, M. (1990). Memory for temporal order and conditional associative-learning in patients with Parkinson's disease. *Neuropsychologia* 28, 1283–1293.
- Yang, C.R., and Seaman, J.K. (1996). Dopamine D1 receptor actions in layer V–VI rat prefrontal cortex neurons in vitro: Modulation of dendritic-somatic signal integration. *J. Neurosci.* 16, 1922–1935.
- Yim, C.Y., and Mogenson, G.J. (1982). Response of nucleus accumbens neurons to amygdala stimulation and its modification by dopamine. *Brain Res.* 239, 401–415.
- Young, A.M.J., Joseph, M.H., and Gray, J.A. (1992). Increased dopamine release in vivo in nucleus accumbens and caudate nucleus of the rat during drinking: a microdialysis study. *Neuroscience* 48, 871–876.
- Waelti, P., Dickinson, A., and Schultz, W. (2001). Dopamine re-

sponses comply with basic assumptions of formal learning theory. *Nature* 412, 43–48.

Watanabe, M. (1989). The appropriateness of behavioral responses coded in post-trial activity of primate prefrontal units. *Neurosci. Lett.* 101, 113–117.

Watanabe, M. (1996). Reward expectancy in primate prefrontal neurons. *Nature* 382, 629–632.

Waterhouse, B.D., and Woodward, D.J. (1980). Interaction of norepinephrine with cerebrocortical activity evoked by stimulation of somatosensory afferent pathways in the rat. *Exp. Neurol.* 67, 11–34.

West, M.O., Peoples, L.L., Michael, A.J., Chapin, J.K., and Woodward, D.J. (1997). Low-dose amphetamine elevates movement-related firing of rat striatal neurons. *Brain Res.* 745, 331–335.

Wickens, J.R., Begg, A.J., and Arbuthnott, G.W. (1996). Dopamine reverses the depression of rat corticostriatal synapses which normally follows high-frequency stimulation of cortex in vitro. *Neuroscience* 70, 1–5.

Williams, S.M., and Goldman-Rakic, P.S. (1993). Characterization of the dopaminergic innervation of the primate frontal cortex using a dopamine-specific antibody. *Cereb. Cortex* 3, 199–222.

Wilson, F.A.W., and Rolls, E.T. (1990a). Neuronal responses related to the novelty and familiarity of visual stimuli in the substantia innominata: diagonal band of Broca and periventricular region of the primate forebrain. *Exp. Brain Res.* 80, 104–120.

Wilson, F.A.W., and Rolls, E.T. (1990b). Neuronal responses related to reinforcement in the primate basal forebrain. *Brain Res.* 509, 213–231.

Wilson, F.A.W., and Rolls, E.T. (1990c). Learning and memory is reflected in the responses of reinforcement-related neurons in the primate basal forebrain. *J. Neurosci.* 10, 1254–1267.

Wise, R.A. (1996a). Addictive drugs and brain stimulation reward. *Annu. Rev. Neurosci.* 19, 319–340.

Wise, R.A. (1996b). Neurobiology of addiction. *Curr. Opin. Neurobiol.* 6, 243–251.

Wise, R.A. (2002). Brain reward circuitry: insights from unsensed incentives. *Neuron* 36, this issue, 229–240.

Wise, R.A., and Hoffman, D.C. (1992). Localization of drug reward mechanisms by intracranial injections. *Synapse* 10, 247–263.