How the Basal Ganglia Use Parallel Excitatory and Inhibitory Learning Pathways to Selectively Respond to Unexpected Rewarding Cues

Joshua Brown, Daniel Bullock, and Stephen Grossberg

Department of Cognitive and Neural Systems and Center for Adaptive Systems, Boston University, Boston, Massachusetts 02215

After classically conditioned learning, dopaminergic cells in the substantia nigra pars compacta (SNc) respond immediately to unexpected conditioned stimuli (CS) but omit formerly seen responses to expected unconditioned stimuli, notably rewards. These cells play an important role in reinforcement learning. A neural model explains the key neurophysiological properties of these cells before, during, and after conditioning, as well as related anatomical and neurophysiological data about the pedunculopontine tegmental nucleus (PPTN), lateral hypothalamus, ventral striatum, and striosomes. The model proposes how two parallel learning pathways from limbic cortex to the SNc, one devoted to excitatory conditioning (through the ventral striatum, ventral pallidum, and PPTN) and the other to adaptively timed inhibitory conditioning (through the striosomes), control SNc responses. The excitatory pathway generates CS-induced excitatory SNc dopamine bursts. The inhibitory pathway prevents dopamine bursts in response to predictable reward-related signals. When expected rewards are not received, striosomal inhibition of SNc that is unopposed by excitation results in a phasic drop in dopamine cell activity. The adaptively timed inhibitory learning uses an intracellular spectrum of timed responses that is proposed to be similar to adaptively timed cellular mechanisms in the hippocampus and cerebellum. These mechanisms are proposed to include metabotropic glutamate receptor-mediated Ca^{2+} spikes that occur with different delays in striosomal cells. A dopaminergic burst in concert with a Ca^{2+} spike is proposed to potentiate inhibitory learning. The model provides a biologically predictive alternative to temporal difference conditioning models and explains substantially more data than alternative models.

Key words: dopamine; substantia nigra; reward; basal ganglia; conditioning; pedunculopontine tegmental nucleus; lateral hypothalamus; striosomes; adaptive timing

Humans and animals can learn to predict both the amounts and times of expected rewards. The dopaminergic cells of the substantia nigra pars compacta (SNc) have unique firing patterns related to the predicted and actual times of reward (Ljungberg et al., 1992; Schultz et al., 1993; Mirenowicz and Schultz, 1994; Schultz et al., 1995; Hollerman and Schultz, 1998; Schultz, 1998). Figures 1 and 2 summarize some of their main properties, notably how learning enables the SNc cells to respond immediately to unexpected cues [conditioned stimulus (CS)] but to omit responses in an adaptively timed fashion to expected rewards [unconditioned stimulus (US)]. Because these firing patterns also act as learning signals in the striatum and elsewhere (Wickens and Kotter, 1995), they have been suggested to play a key role in both addictive behavior (Garris et al., 1999) and reinforcement learning. In particular, dopaminergic reward signals seem to strengthen the "incentive salience" or "wanting" of a certain reward, that is, the motivation to work for the reward in a given behavioral context, as distinct from the affective enjoyment or "liking" of a reward once consumed (Berridge and Robinson, 1998). The liking may be mediated by areas other than the basal ganglia (McDonald and White, 1993). Recent models (Houk et al., 1995; Montague et al., 1996; Contreras-Vidal & Schultz, 1997; Schultz et al., 1997; Berns and Sejnowski, 1998; Suri and Schultz, 1998) of the nigral dopamine cells have noted similarities between dopamine cell properties and well known learning algorithms, especially temporal difference (TD) models (Montague et al., 1996; Schultz et al., 1997; Suri and Schultz, 1998). Although providing a degree of insight into the information carried by the dopamine signal, the TD approach has not been able to answer the questions of what biological mechanisms actually compute the signal, and how. In particular, how does learning in the circuit that includes these cells enable them to produce a fast excitatory response to conditioned stimuli and a delayed, adaptively timed inhibition of response to rewarding unconditioned stimuli, in all of the experimental conditions summarized by Figures 1 and 2? We show here that the known anatomy and cell types in pathways afferent to dopamine cells lead to an explanation with significant advantages over previous models.

We introduce a model in which the learned excitatory and inhibitory responses are subserved by different anatomical pathways, and the adaptively timed inhibitory learning is mediated by metabotropic glutamate receptor (mGluR)-driven Ca^{2+} spikes in striosomal cells. These Ca^{2+} spikes occur with a spectrum of
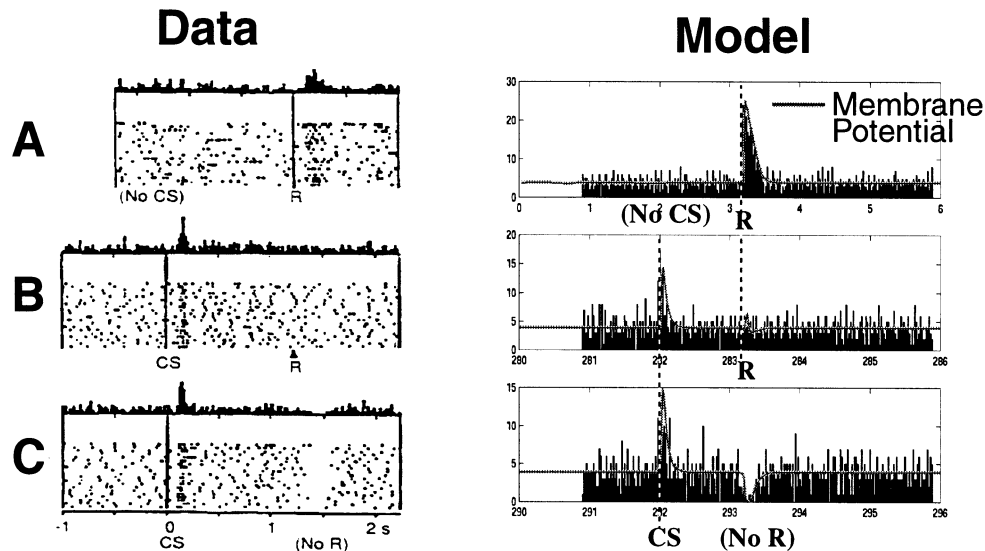
Data

Model

Figure 1. Dopamine cell firing patterns. *Left*, Data. *Right*, Model simulation, showing model spikes and underlying membrane potential. *A*, In naive monkeys, the dopamine cells fire a phasic burst when unpredicted primary reward R occurs (e.g., if the monkey receives a burst of apple juice unexpectedly). *B*, As the animal learns to expect the apple juice that reliably follows a sensory cue [conditioned stimulus (*CS*)] that precedes it by a fixed time interval, then the phasic dopamine burst disappears at the expected time of reward, and a new burst appears at the time of the reward-predicting CS. *C*, After learning, if the animal fails to receive reward at the expected time, a phasic depression in dopamine cell firing occurs. Thus, these cells reflect an adaptively timed expectation of reward that cancels the expected reward at the expected time. [The data in Figure 1 (*column 1*) are reprinted with permission from Schultz et al. (1997).]

temporal delays. When a Ca^{2+} spike and a dopamine burst occur at the same time, inhibitory learning is enhanced at the corresponding delays. To explicate these excitatory and inhibitory pathways, the model functionally explains and simulates the firing patterns of dopamine cells, striosomal cells of the striatum, pedunculopontine tegmental nucleus (PPTN) cells, ventral striatal cells, and lateral hypothalamic cells (see Figs. 1–3). Its mGluR-based spectral timing mechanism helps to explain more data than the temporal derivative operation that defines the class of TD models previously used to describe dopamine cell behavior. This model is shown schematically in Figure 4.

MATERIALS AND METHODS

Dopamine cell responses can be conditioned to phasic cues whose offsets occur long before the reward signals that they predict (Ljungberg et al., 1992). To bridge the temporal gap, a CS is assumed to activate a sustained working memory input to the model (Funahashi et al., 1989). A subsequent primary reward signal from a US is assumed to trigger a dopamine burst, which augments the weights between the working memory site and the ventral striatum (Wickens et al., 1996). This allows future CS presentations to elicit an immediate excitatory prediction of reward. The CS also activates a population of lagged inhibitory signals from the striosomes to the SNc. When a dopamine burst occurs at a sufficient lag after CS onset, it strengthens the subset of lagged inhibitory signals that are active at that time. These two types of learning enable a CS to generate an immediate, reward-predictive dopamine signal but also to cancel subsequent SNc excitation that would otherwise be caused by the predicted reward-related signals. When a response is made and reward is received, the working memory input is assumed to shut off (Funahashi et al., 1989).

We propose that the PPTN is responsible for the phasic bursts of activity in SNc dopamine cells (Figs. 1 and 2) and thus plays a key role in the learning and maintenance of instrumental tasks. Experiments showing monosynaptic glutamatergic and cholinergic PPTN-to-SNc projections (Scarnati et al., 1988; Conde, 1992; Futami et al., 1995) support this hypothesis. Conde (1992) has suggested that the PPTN provides the main source of excitation to the SNc, and PPTN cells have been found to fire phasically in response to primary reward or reward-predicting conditioned stimuli, or both, leaving them well situated to provide this kind of SNc input (Dormont et al., 1998) (Fig. 3*A*). The phasic nature of

PPTN signaling is attributable to habituation, or accommodation, in SNc-projecting PPTN cells (Takakusaki et al., 1997). Lesions of the PPTN produced hemiparkinsonian symptoms, as if the SNc itself had been lesioned (Kojima et al., 1997), and reversible PPTN inactivation mimics extinction in an instrumental task, even while rewards, if provided, are readily consumed (Conde et al., 1998).

PPTN afferents. From where does the PPTN receive these response-motivating reward and reward-predicting signals? We propose that the primary reward signals come from the lateral hypothalamus, whereas the excitatory reward-prediction signals (which generate a CS-induced dopamine burst) travel via the ventral striatum–ventral pallidum pathway, which receives input mainly from limbic cortex (Schultz et al., 1992) (Fig. 4). Lateral hypothalamic neurons are known to play a role in feeding behavior and to fire phasically in response to primary reward (Nakamura and Ono, 1986), as in Figure 3*D*. A strong lateral hypothalamus–PPTN projection has been found and confirmed by both anterograde and retrograde labeling (Semba and Fibiger, 1992), and the primary reward signal explains the similar phasic reward response in the PPTN. Thus, the lateral hypothalamus seems to be a principal source of excitation to the PPTN.

Likewise, more than one-fourth of the ventral pallidum projects collaterals to the PPTN (Mogenson and Wu, 1986). The ventral pallidum receives projections from the matrisomes of the ventral striatum (Yang and Mogenson, 1987), which responds to both predicted and primary reward (Schultz et al., 1992), as in Figure 3*B*. The double inhibition from ventral striatum to ventral pallidum to PPTN results in net excitation from ventral striatum to PPTN. We predict that the sustained, CS-induced striatal activation that is shown in Figure 3*B* is attributable to receipt of a working memory trace of the CS from limbic cortex, which is enhanced by learning of CS-reward contingencies (Dias et al., 1996). The transient component in Figure 3*B* results from a phasic primary reward signal from the lateral hypothalamus (Nakamura and Ono, 1986; Brog et al., 1993). We suggest that the ventral striatum is a main pathway of excitatory reward predictions.

Other PPTN afferents are possible candidates for generating phasic PPTN responses. Some other possible sources, found by retrograde labeling from the PPTN, include the central nucleus of the amygdala (CNA) and the subthalamic nucleus (STN) (Semba and Fibiger, 1992). The amygdala does not appear to provide the main source of excitation, despite its processing of emotional valence information. In particular, it has been shown that rats with amygdala lesions could still learn operant tasks (McDonald and White, 1993). After CNA damage, rats can learn
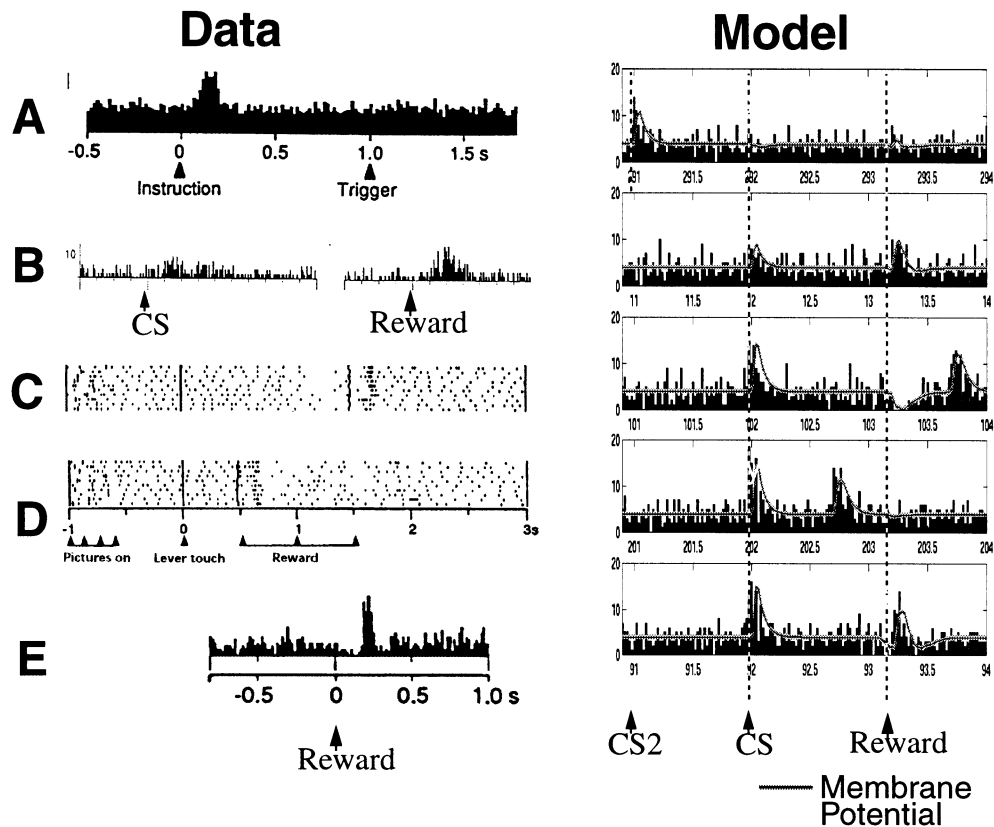
Figure 2. Dopamine cell firing patterns. *Left*, Data. *Right*, Model simulation, showing model spikes and underlying membrane potential. *A*, The dopamine cells learn to fire in response to the earliest consistent predictor of reward. When CS2 (*Instruction*) consistently precedes the original CS (*Trigger*) by a fixed interval, the dopamine cells learn to fire only in response to CS2. [Data reprinted with permission from Schultz et al. (1993).] *B*, During training, the cell fires weakly in response to both the CS and reward. [Data reprinted with permission from Ljungberg et al. (1992).] *C*, Temporal variability in reward occurrence. When reward is received later than predicted, a depression occurs at the time of predicted reward, followed by a phasic burst at the time of actual reward. *D*, Likewise, if reward occurs earlier than predicted, a phasic burst occurs at the time of actual reward. No depression follows because the CS is released from working memory. [Data in *C* and *D* reprinted with permission from Hollerman and Schultz (1998).] *E*, When there is random variability in the timing of primary reward across trials (e.g., when the reward depends on an operant response to the CS), the striosomal cells produce a "Mexican hat" depression on either side of the dopamine spike. [Data reprinted with permission from Schultz et al. (1993).]

second-order conditioning although they fail to learn a conditioned orienting response (Gallagher and Chiba, 1996). Similarly, some studies suggest a modulatory rather than an excitatory role of the STN-to-SNc projection (Smith and Grace, 1992), and cell recording studies have not yet shown reward-predicting activity in the STN.

Striosomes. What suppresses the dopamine burst response to primary reward after conditioning has occurred, and what causes the transient activity drop when expected reward is not received (Fig. 1)? The striosomal cells provide a significant source of GABAergic inhibition to the SNc (Gerfen, 1992), which could account for both of these phenomena. In turn, striosomal cells receive dopaminergic projections from the SNc (Gerfen, 1992). We propose that an intracellular spectral timing mechanism (Grossberg and Schmajuk, 1989; Grossberg and Merrill 1992, 1996; Fiala et al., 1996) provides the function needed. Specifically, the striosomal cells briefly inhibit SNc dopamine cells, after a learned delay period, to provide an inhibitory expectation of reward. The model incorporates striosomal cells in both the dorsal and ventral aspects of the striatum. Likewise, model dopamine cells correspond to both dorsal and ventral SNc cells, which despite certain differences have similar inputs and response properties. Gerfen (1992) has noted the distinction between the dorsal and ventral tiers of the SNc: dorsal tier SNc cells project to the matrisomes of the striatum (including the model ventral striatal cells), whereas ventral tier SNc cells project to the striosomes. The model lumps together the ventral and dorsal tiers of the SNc on the basis of their similarities.

It has been suggested that striosomal cells provide adaptively timed inhibition to the dopamine cells (Contreras-Vidal and Schultz, 1997), much as cerebellar Purkinje cells provide adaptively timed inhibition of interpositus nucleus cells (Fiala et al., 1996), but this general hypothesis

must be coupled to a biologically supported local mechanism. Given evidence that striatal learning is suppressed by mGluR blockers (Calabresi et al., 1992a) and Ca^{2+}-chelators (Calabresi et al., 1994), we suggest the following striosomal cell model: conditioned stimuli excite a glutamatergic corticostriatal pathway that activates mGluRs on striosomal neurons. These in turn cause a delayed transient rise in intracellular Ca^{2+}, at least partly via NMDA channels (Calabresi et al., 1992b), which are known to be potentiated by mGluR1 receptor activation (Pisani et al., 1996). This Ca^{2+} response is proposed to be a basis for both learning and generating an adaptively timed inhibitory striosomal–SNc signal. The model uses a population of striosomal cells with a range of delayed responses (Fig. 5), which, taken together, constitute the "spectrum" of possible learned delays.

Fiala et al. (1996) proposed a model of adaptively timed conditioning in which cerebellar Purkinje cells generate a spectrum of differently delayed Ca^{2+} spikes after excitation of mGluR1 receptors. A Ca^{2+} spike by itself activates a Ca^{2+}-dependent K^+ conductance, which is hyperpolarizing. In addition, when a climbing fiber signal is received at the same time as a delayed Ca^{2+} spike, it causes a long-term increase in the Ca^{2+}-dependent K^+ channel conductance. Thus, in the cerebellar model, the Ca^{2+} spike is a basis for both immediate hyperpolarization and learned long-term depression (LTD).

We propose that a related but distinct mechanism operates in striosomal cells, which, unlike Purkinje cells (Crepel et al., 1996), possess NMDA receptors. In this context, a mGluR1-mediated delayed Ca^{2+} spike can be amplified and thus serve to transiently increase rather than decrease striosomal cell activity. A class of recently discovered Ca-inhibited K^+ channels (Joiner et al., 1998) may also contribute to a Ca-dependent depolarization. A Ca^{2+} spike combined with a phasic
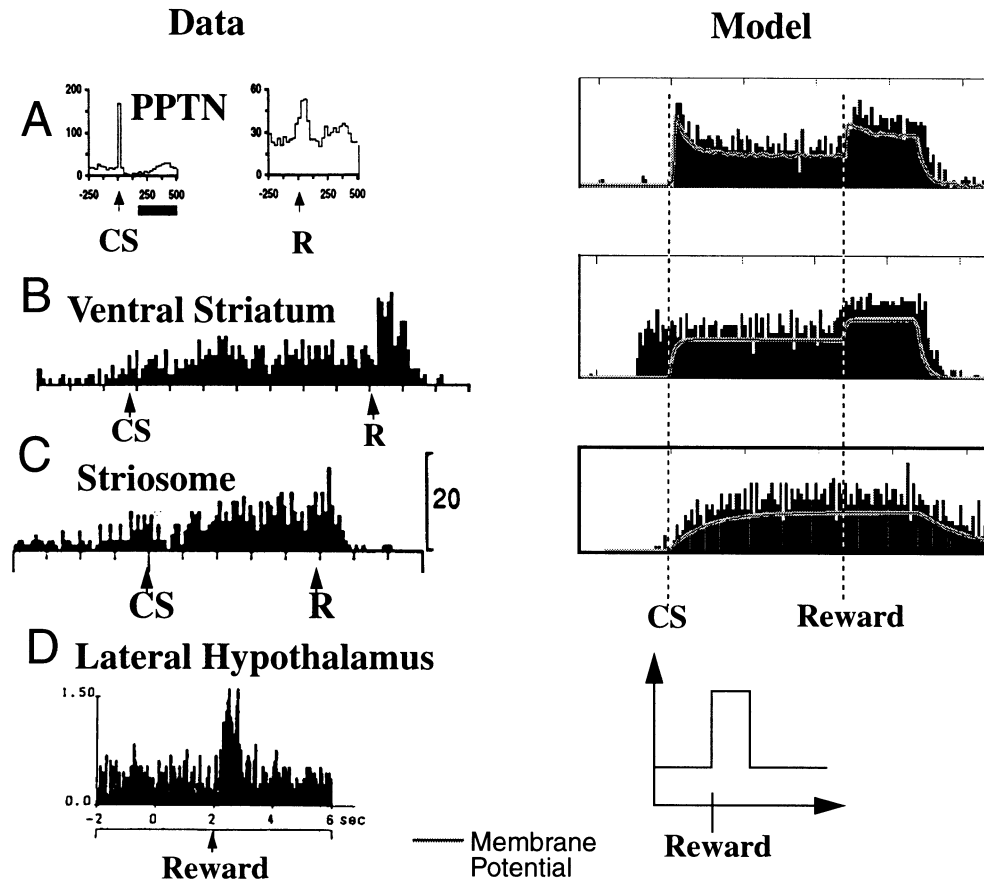
Figure 3. Trained firing patterns in PPTN, ventral striatum, striosomes, and lateral hypothalamus. *Left*, Data. *Right*, Model simulations, showing model spikes and underlying membrane potential. *A*, PPTN cell (cat), showing phasic responses to both CS and primary reward. [Data reprinted with permission from Dormont et al. (1998).] In the model, phasic signaling is caused by accommodation or habituation (Takakusaki et al., 1997), which causes the cell to fire in response to the earliest reward-predicting CS and US reward, but not to subsequent CSs before reward. *B*, Ventral striatal cells show sustained working memory-like response between trigger and a US reward, and a phasic response to the US reward. [Data reprinted with permission from Schultz et al. (1992).] *C*, A ventral striatal cell, predicted here to be a striosomal cell, shows buildup to phasic primary reward response. For the model cell, $j = 39$. [Data reprinted with permission from Schultz et al. (1992).] *D*, A lateral hypothalamic neuron with a strong, phasic response to glucose reward. [Data reprinted with permission from Nakamura and Ono (1986).] The majority of these neurons fired in response to primary reward but not to a reward-predicting CS. The model lateral hypothalamic input is a rectangular pulse.

burst of dopamine acting on striosomal D1 receptors would also allow long-term potentiation (LTP) in striosomal cells. It has been suggested that increased Ca^{2+} combined with a dopamine burst could result in a potentiation of glutamate receptors (LTP) (Houk et al., 1995), and dopamine bursts have been shown to reverse corticostriatal LTD and instead cause LTP (Wickens et al., 1996). Thus, a delayed Ca^{2+} spike in the striosomal cells could serve as both a signaling gate and one component of a learning gate.

Recent work on the cerebellum (Finch and Augustine, 1998; Takechi et al., 1998) has supported the Fiala et al. (1996) cerebellar model and demonstrated the feasibility of direct calcium imaging in local regions of a dendritic arbor using high-speed confocal microscopy. We suggest that the same technique could be used in neostriatal cells to investigate the predictions regarding striosomal Ca dynamics. Pharmacological inactivation of mGluR1 and IP_3 might also verify whether they are essential components of the Ca spike cascade, as in the cerebellum.

Functionally, the striosomal cells of the model need to receive a sustained input that is activated when a CS first occurs, as a reference point for the delayed inhibitory signal. Striosomal cells receive excitatory signals from deep layer V of limbic cortex (Gerfen, 1992). The sustained working memory signal initiates a steady rise of the intracellular calcium level, e.g., via an mGluR1-IP_3-Ca cascade (as in the cerebellum) (Finch and Augustine, 1998; Takechi et al., 1998), which causes a calcium spike on reaching a threshold. The sustained input hereby leads to a delayed, phasic response within the striosomal cell. A related property of the model is that if the sustained input strength is proportional to the CS intensity, then a weaker CS causes an increase in the rise time to

threshold, resulting in a slower perceived rate of time passage. This property agrees with behavioral data (Wilkie, 1987), although because of the complexity of cortical processing, the striosomal inputs may not be directly proportional to external stimulus intensity. The model simulations assume a simple two-state working memory input that is either on or off and could be generated by passing a gradually rising input through a sharp sigmoidal signal function. The maximum delay that a single spectrum can adaptively time is still unknown and needs to be investigated biochemically (cf. Fiala et al., 1996). Spectral timing of a single event also needs to be supplemented by inter-event timing mechanisms that involve network interactions, including prefrontal cortex and cerebellum (Buonomano and Mauk, 1994; Grossberg and Merrill, 1996).

RESULTS

Given the above background, the model mechanisms can now be summarized as follows (Fig. 4).

First, a primary reward signal is generated in the lateral hypothalamus (Nakamura and Ono, 1986) (Fig. 3*D*). This directly excites the PPTN (Semba and Fibiger, 1992), which fires a brief burst and then accommodates or habituates (Takakusaki et al., 1997; Dormont et al., 1998). This brief burst directly excites the SNc by cholinergic and/or glutamatergic projections (Conde, 1992) and thereby causes a phasic dopamine burst to the striatum (Gerfen, 1992) at the time of primary reward.
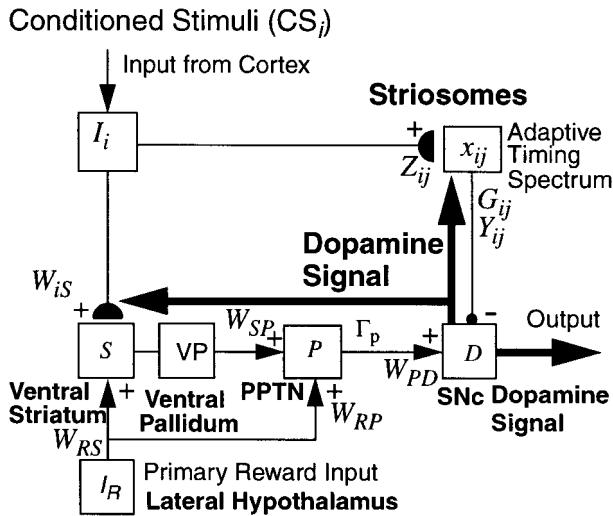
Figure 4. Model circuit. Cortical inputs (I_i) excited by conditioned stimuli learn to excite the SNc (D) via the ventral striatal (S)-to-ventral pallidal-to-PPTN (P)-to-SNc path. The inputs I_i excite the ventral striatum via adaptive weights W_{iS}, and the ventral striatum excites the PPTN, via double inhibition through the ventral pallidum, with strength W_{SP}. When the PPTN activity exceeds a threshold Γ_P, it excites the dopamine cell with strength W_{PD}. The striosomes, which contain an adaptive spectral timing mechanism (x_{ij}, G_{ij}, Y_{ij}, Z_{ij}), learn to generate lagged, adaptively timed signals that inhibit reward-related activation of SNc. Primary reward signals (I_R) from the lateral hypothalamus both excite the PPTN directly (with strength W_{RP}) and act as training signals to the ventral striatum S (with strength W_{RS}). *Arrowheads* denote excitatory pathways, *circles* denote inhibitory pathways, and *hemidisks* denote synapses at which learning occurs. *Thick pathways* denote dopaminergic signals.

Suppose that a CS is received and stored in prefrontal working memory at some time τ before the actual reward. This CS trace generates output signals along adaptive pathways to both the ventral striatum and the striosomes. When primary reward occurs, a dopamine burst facilitates LTP in the limbic cortical–ventral striatal path (Brog et al., 1993). Thus, the CS representation in limbic prefrontal cortex learns to excite the dopamine cells via the limbic cortical–ventral striatal–ventral pallidum-PPTN-SNc pathway (Yang and Mogenson, 1987). In the model, the ventral striatum and ventral pallidum are lumped for simplicity into a single ventral basal ganglia node, which causes net excitation of the PPTN.

The limbic cortical projection to the striosomes (Gerfen, 1992; Eblen and Graybiel, 1995) activates a spectrum of delayed Ca^{2+} spikes in the striosomal cells via metabotropic glutamate receptors. When a dopamine burst arrives from the SNc, it strengthens the CS-activated limbic cortical connections to any currently spiking components of the striosomal timing spectrum. The striosomal cells hereby learn to inhibit the dopamine burst at its expected time via the inhibitory striosomal–SNc path (Gerfen, 1992).

On a later trial in the trained model, when the CS is received at the expected time before an actual reward, its working memory trace tonically activates the ventral striatal model cell, which in turn excites the PPTN, causing an immediate dopamine burst in the SNc. The adaptively timed inhibition via the striosomal cells then inhibits the SNc so that the subsequent primary reward signal does not elicit a dopamine burst in the SNc. If the primary reward signal is absent on a trial, then the striosomal inhibition causes a phasic dip in the dopamine signal. These three properties explain the dopamine cell data of Figure 1.
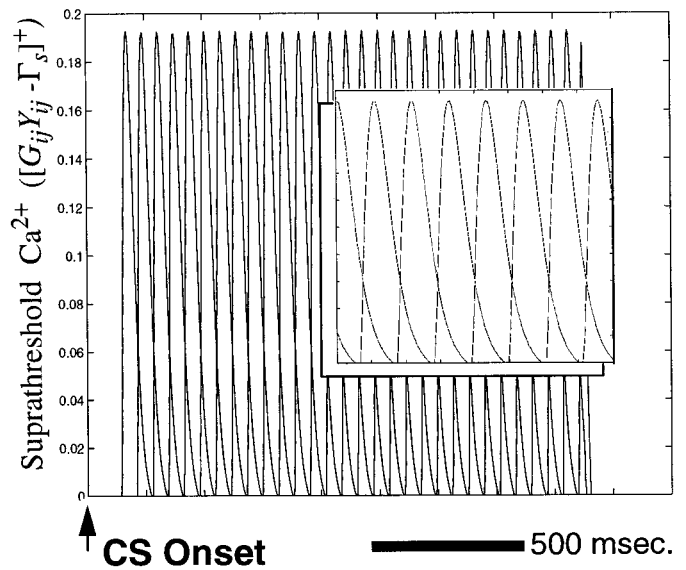

Figure 5. Striosomal spectral timing model and closeup (*inset*), showing individual timing pulses. Each curve represents the suprathreshold intracellular Ca^{2+} concentration $[G_{ij}Y_{ij} - \Gamma_s]^+$ of one striosomal cell. The peaks are spread out in time so that reward can be predicted at various times after CS onset, by strengthening the inhibitory effect of the striosomal cell with the appropriate delay. The model uses 40 peaks, spanning ~2 sec and beginning 100 msec after the CSs (Grossberg and Schmajuk, 1989). Model properties are robust when different numbers of peaks are used. It is important that the peaks be sufficiently narrow and tightly spaced to permit fine temporal resolution in the reward-canceling signal. However, a trade-off ensues in that more timed signals must be used as the time between peaks is reduced. The timed signals must not begin too early after the CS, or they will erroneously cancel the CS-induced dopamine burst. The 100 msec post-CS onset delay prevents this from happening.

The model was also used to simulate various other task situations for which dopamine cell responses are known. It successfully reproduced all the key SNc dopamine cell data (Figs. 1, 2) as well as firing patterns of known cell types in the PPTN (Fig. 3*A*) and ventral striatum (Fig. 3*B*), which are afferent to the nigral dopamine cells. In particular, dopamine cell responses were simulated in eight task situations (Figs. 1, 2). First, the model received primary reward (R) only and showed a strong response to the reward (Fig. 1*A*). We then trained the model with a CS preceding R. During training, the model fired weakly in response to both the CS and R (Fig. 2*B*). As training neared completion, the model SNc responded strongly and only to the CS (Fig. 1*B*). In the trained model, we examined the effect of omitting R and found a transient depression at the predicted time of reward (Fig. 1*C*). To test the effects of higher-order conditioning, we first trained the model with the CS–R association. Then we introduced an additional conditioned stimulus (*CS2*), which consistently occurred 1 sec before the CS. With training, the model dopaminergic cells learned to respond only to CS2 (Fig. 2*A*).

Recent work has examined dopamine cell responses under conditions of variable reward timing (Hollerman and Schultz, 1998). The model successfully simulated these data as well. When the reward R was delayed (Fig. 2*C*), model dopamine cells responded with the characteristic depression at the expected time of R and then showed a burst later when R did occur. Similarly, if R occurred before the expected time, model dopamine cells again showed a burst in response to R. They did not, however, show a dip at the expected time of R (Fig. 2*D*), in agreement with

the data, because the working memory trace shut off when R was received. In some cases, the timing of primary reward may vary from trial to trial because of its dependence on an operant response. The model dopamine response was simulated when the timing of R varied randomly on an interval spanning 200 msec before and after the expected (mean) time of R, with a uniform random distribution. This caused model striosomal cells to learn to inhibit the dopamine signal during the entire interval in which the dopamine bursts occurred. Because this interval of inhibition is wider than the dopamine burst, model striosomal cells produced tails of depressed firing on either side of the dopamine burst (Fig. 2*E*), generating a kind of temporal Mexican hat function, as in the data (Schultz et al., 1993).

The PPTN model responses also agree with the cell recording data from conditioning tasks (Dormont et al., 1998), which show transient bursts in response to both CS and R (Fig. 3*A*). In addition, when a CS2 preceded the CS, the model PPTN response to the later CS disappeared. This lack of response to subsequent CSs agrees with the data of Dormont et al. (1998), which show a similar disappearance of the CS-induced PPTN response in that delay task.

Model ventral striatal cells also simulated known cell firing patterns (Fig. 3*B*). After the model learned the CS–R association, CS onset produced tonic activity, followed by a phasic burst in response to the R signal from the hypothalamus (Fig. 3*D*).

DISCUSSION

The present model explains and predicts significantly more data than previous models through its use of parallel learning pathways. Several models have attempted to describe the dopamine cell behavior by a TD algorithm (Montague et al., 1996; Schultz et al., 1997; Suri and Schultz, 1998). These models suggest that the dopaminergic SNc cells compute a temporal derivative of predicted reward. In other words, they fire in response to the sum of the time-derivative of reward prediction and the actual reward received. These models have not been linked with structures in the brain that might compute the required signals. The Suri and Schultz (1998) model has simulated much of the known dopamine cell data. However, their model can only learn a single fixed interstimulus interval (ISI) that corresponds to the longest-duration timed signal $[x_{1m}(t)]$ in their model. If the ISI is shorter than this, dopamine bursts will strengthen all of the active stimulus representations predicting reward at the time of the dopamine burst or later. Thus, their model generates inhibitory reward predictions beyond the primary reward time and predicts a lasting depression of dopamine firing subsequent to primary reward, which is not found in the data.

In contrast to TD models that compute time derivatives immediately before dopamine cells, our spectral timing model uses two distinct pathways: the ventral striatum and PPTN for initial excitatory reward prediction and the striosomal cells for timed, inhibitory reward prediction. The fast excitation and delayed inhibition are hereby computed by separate structures within the brain, rather than by a single temporal differentiator. This separation avoids the problem of the Suri and Schultz (1998) model by allowing transient rather than sustained signals to cancel the primary reward signal, thereby enabling precisely timed reward-canceling signals to be trained, and preventing spurious sustained inhibitory signals to the dopamine cells. This separation also allows the inhibitory system to follow and precisely cancel the real-time dynamics of the primary reward signal, as in Figure 1*B*, where the striosomal signals cancel the dopamine burst despite its

asymmetry. Where temporal uncertainty exists in reward prediction, the tails of inhibition (Fig. 2*E*) in the data are explained by the model's ability to learn temporally distributed net inhibitory signals that track the temporal dispersion of reward.

Like our model, the TD model of Schultz et al. (1997) uses transient rather than sustained timing signals. However, because this model does not separate the computation of excitation and inhibition, each transient pulse is temporally differentiated to produce an onset burst followed by an offset depression. Over the course of many trials, the onset burst strengthens its preceding timed signal weight, thereby recursively chaining backward until all timed signal weights between the CS and R have been activated by learning. This predicts that the dopamine burst gradually travels backward in time and that the reward response extinguishes well before the CS response occurs. The data show instead that dopamine bursts do not occur systematically in the middle of the ISI during training, and moreover, the dopamine burst occurs concurrently at both CS and R during individual training trials (Ljungberg et al., 1992).

The Contreras-Vidal and Schultz (1997) model of the dopamine cell system is based partly on the ART2 model (Carpenter and Grossberg, 1987). They first suggested that striosomes may generate a spectrum of adaptively timed reward predictions, based on the earlier spectral timing models of Grossberg and colleagues (Grossberg and Schmajuk, 1989; Grossberg and Merrill, 1992, 1996; Fiala et al., 1996). Their striosomal model nonetheless faces problems because it relies on lateral inhibition among striosomal cells, rather than intracellular timing mechanisms. GABAergic lateral inhibition among striosomal cells is weak (Jaeger et al., 1994; Wilson, 1995) and may not be strong enough to mediate the competitive choices required by their model. In addition, their model assumes adaptively timed inhibitory reward prediction learning at the striosomal–SNc synapses instead of at the corticostriosomal synapses. This fails to incorporate data on corticostriatal LTP/LTD (Wickens and Kotter, 1995). In their model, corticostriatal LTP/LTD would cause erroneous timing predictions because the cell with the strongest corticostriatal input becomes active first and generates its adaptively timed signal, whereas it suppresses its competing neighbor cells via strong lateral inhibition. After this, the winning cell remains refractory, and the cell with the next strongest corticostriosomal weight becomes active, and so on. If learning occurs in the corticostriosomal path, as much evidence suggests, then the rank ordering of corticostriosomal weights may change as the synaptic weights change relative to each other. This would cause erroneous reward timing predictions, because the model striosomal cells would become active in the wrong sequential order. Our model avoids these problems by describing an intracellular mGluR-mediated adaptive timing mechanism rather than an extracellular one.

Another significant difference between the present model and that of Contreras-Vidal and Schultz (1997) is the source of excitation to the dopamine cells. Their model assumes that matrisomal cells provide the excitatory input to SNc cells indirectly, via double inhibition through the substantia nigra pars reticulata (SNr). This polysynaptic, matrisomal cell-SNr-SNc pathway cannot be ruled out as a source of net excitation to the dopamine cells, but as we have shown above, it is not the main pathway of SNc excitation. It should also be pointed out that although the present model attempts to represent the principal circuitry responsible for dopamine cell responses, additional afferent circuitry exists that may also be capable of eliciting phasic dopamine

cell responses, e.g., the SNr–SNc projection, and the STN–PPTN and STN–SNc projections.

Houk et al. (1995) modeled dopamine cell firing using the direct and indirect basal ganglia pathways. They assumed that the polysynaptic, net excitatory indirect path through the basal ganglia is faster than the monosynaptic, direct path. The indirect path is proposed to generate the initial excitatory dopamine burst, whereas the direct path is proposed to mediate the slower inhibition of the dopamine cells.

With regard to the fast excitation of the dopamine cells, Houk et al. (1995) cite data showing that striatal stimulation results in a fast EPSP followed by a slower IPSP in the globus pallidus (Kita and Kitai, 1991). However, it is unlikely that the EPSPs are polysynaptic, because they could be elicited with as little as 2 msec latency (Kita and Kitai, 1991). Likewise, the fast EPSP that results from cortical excitation (Kita, 1992) might be better explained as from a cortical-STN-pallidal route. Moreover, STN activity may modulate rather than excite the SNc (Smith and Grace, 1992). These data contradict Houk and colleagues' (1995) assumption of net striatal–SNc excitation via the model indirect pathway. The data are probably caused by STN–SNr excitation and subsequent SNr–SNc inhibition (Hajos and Greenfield, 1994; Tepper et al., 1995).

With regard to the slow inhibition of the dopamine cells, Houk et al. (1995) proposed that the direct path provides a prolonged inhibition of the dopaminergic cells, which persists from the time of the reward-predicting CS through the time at which the reward occurs. This is inconsistent with the data in two distinct but related ways. First, when the reward-predicting CS occurs, it produces a dopamine burst, but the dopamine cell firing then immediately returns to baseline. There is no persistent depression in dopamine cell firing, although the Houk et al. (1995) model must predict such a persistent depression. Second, when an expected reward is omitted, there is a brief depression in the dopamine cell firing, after which it immediately returns to baseline. The Houk et al. (1995) model instead predicts a prolonged (although below baseline) response rather than a transient response to the omission of expected reward.

The Berns and Sejnowski (1998) model suggests that the primary source of net SNc excitation is the pallidum, via a hypothetical inhibitory neuron. No suggestion is given regarding the location of this neuron or from which pallidal segment (internal or external) the signal originates. As in our model, the Berns and Sejnowski (1998) model assumes that the striosomal cells are the main source of inhibition to the SNc, but their model does not treat dopamine cell temporal dynamics, which would be necessary for it to explain the data of Figures 1 and 2.

The new spectral timing model of nigral dopamine activity provides functional explanations of known SNc afferents. The model suggests how the ventral basal ganglia stream learns an excitatory prediction of reward via the PPTN, whereas the striosomal cells learn an adaptively timed inhibitory prediction of reward. This analysis clarifies how the nigral dopamine cells are linked to four other cell types that are directly or indirectly afferent to the SNc: ventral striatal cells, PPTN cells, striosomal cells of the basal ganglia, and cells in the lateral hypothalamus. The model predicts that an adaptive timing mechanism occurs at the striosomal cells. Key explanatory limitations of previous models, including TD and direct/indirect pathway models of nigral dopamine cell responses, are overcome by the present model.

Table 1. Model variables

S	Ventral striatal cell
I_R	Reward input signal from lateral hypothalamus
W_{iS}	CS-to-striatum synaptic weights
N^+	Above-baseline dopamine burst signal
N^-	Below-baseline dopamine dip signal
I_i	CS input signal
P	PPTN cell activity
U_P	PPTN cell afterhyperpolarization
x_{ij}	Striosomal metabotropic response
$G_{ij}Y_{ij}$	Striosomal calcium concentration
Z_{ij}	CS input-to-striosomal synaptic weights
D	Dopamine cell activity
$\bar{D}$	Baseline average dopamine signal
r_j	Striosomal activity buildup rate parameter
M	Membrane potential driving integrate-and-fire (IAF) spiking model
ε	Gaussian noise input to IAF model

APPENDIX

This section lists the mathematical equations and parameters of the model. The circuit in Figure 4 was modeled using neurons with a single-voltage compartment. The model variables are summarized in Table 1, and the fixed parameters are summarized in Table 2. The variables in Figure 4 obey the following equations. Model ventral striatal cell activity S responds at rate τ_S and is excited by primary reward inputs I_R and by CS inputs I_i that are gated by adaptive weights W_{iS}:

$$\frac{1}{\tau_S}\frac{d}{dt}S = -A_s S + (1-S)\left[\sum_i I_i W_{iS} + I_R W_{RS}\right]. \quad (1)$$

The CS-to-striatal weights W_{iS} change only when S is positive. They are potentiated by a "positively reinforcing" dopamine burst N^+ and depressed by a "negatively reinforcing" dopamine depression N^-, described below. The weights W_{iS} range between a minimum of zero and a maximum of $W_S^{max}I_i$, and they decay at a rate β_{WS} with negative reinforcement:

$$\frac{1}{\tau_{WS}}\frac{d}{dt}W_{iS} = S[N^+(I_i W_S^{max} - W_{iS}) - \beta_{WS}N^- W_{iS}]. \quad (2)$$

The PPTN activity P is excited by striatal inputs S and primary reward inputs I_R:

$$\frac{1}{\tau_P}\frac{d}{dt}P = -[1 + U_P W_{UP}]P + (1-P)[S W_{SP} + I_R W_{RP}]. \quad (3)$$

Accommodation, or habituation, of PPTN activity is modeled as a lasting afterhyperpolarization, which reduces the excitability of the PPTN in an activity-dependent way:

$$\frac{1}{\tau_{UP}}\frac{d}{dt}U_P = -U_P + (1-U_P)P. \quad (4)$$

The dopamine cell activity D is excited by the rectified PPTN activity $[P - \Gamma_P]^+$, where Γ_P is a signal threshold, and a tonic arousal signal I_D. The notation $[x]^+ = max(x,0)$ denotes rectification. The dopamine cell activity D is inhibited in an adaptively timed fashion by the summed spectrum of signals:

Table 2. Model parameters

Symbol	Description	Value
α_r	Striosomal spectrum spacing	50.0
β_r	Striosomal spectrum offset	1.0
Γ_G	Calcium spike threshold	0.37
α_G	Calcium activation rate	5.0
β_G	Calcium passive decay rate	20.0
B_G	Calcium concentration maximum	5.0
α_y	Calcium recovery rate	1.0
β_y	Activity-dependent calcium inactivation rate	80.0
Γ_Y	Calcium inactivation threshold	0.18
Γ_S	Striosomal output threshold	0.2
γ_s	Striosomal learning gain	10000
α_z	Striosomal learning rate	0.1
w_{RS}	Hypothalamus-to-ventral striatum synaptic weight	1.2
τ_S	Ventral striatal cell response time constant	30.0
τ_{WS}	CS-to-ventral striatal learning rate	20.0
W_S^{max}	Maximum CS-to-ventral striatal synaptic weight	2.5
β_{WS}	CS-to-ventral striatal weight decay rate	0.2
A_S	Ventral striatal activity passive decay rate	0.7
Γ_N	Phasic dopamine signal threshold	0.0
τ_P	PPTN cell response time constant	200.0
t_{UP}	PPTN afterhyperpolarization time constant	4.0
τ_D	Dopamine cell response time constant	15.0
W_{PD}	PPTN-to-dopamine cell synaptic weight	50.0
W_{SP}	Ventral striatal-to-PPTN cell synaptic weight	2.0
W_{RP}	Hypothalamus-to-PPTN cell synaptic weight	0.8
W_{UP}	PPTN afterhyperpolarization gain	140.0
Γ_P	PPTN output signal threshold	0.135
$\tau_{\bar{D}}$	Baseline average dopamine time constant	4.0
I_D	Tonic input to dopamine cell	0.15
h_D	Dopamine cell maximum hyperpolarization	0.1
V_I	Integrate-and-fire (IAF) model output	0.5
R	IAF model membrane resistance	1333
C	IAF model membrane capacitance	0.025
σ_{noise}	IAF Gaussian noise input	0.4
R_{DA}	IAF dopamine cell membrane resistance	80
R_{PPTN}	IAF PPTN cell membrane resistance	6667
C_{PPTN}	IAF PPTN cell membrane capacitance	0.005
σ_{PPTN}	IAF PPTN cell Gaussian noise input	0.1

$$\sum_{i,j} [G_{ij}Y_{ij} - \Gamma_S]^+ Z_{ij} \tag{5}$$

from the striosomal cells:

$$\frac{1}{\tau_D} \frac{d}{dt} D = -D + (1 - D)[[P - \Gamma_P]^+ W_{PD} + I_D]$$

$$- (D + h_D) \sum_{i,j} [G_{ij}Y_{ij} - \Gamma_S]^+ Z_{ij}. \tag{6}$$

A tonic dopamine signal is computed as a time average of the momentary dopamine cell potential:

$$\frac{1}{\tau_{\bar{D}}} \frac{d}{dt} \bar{D} = D - \bar{D}. \tag{7}$$

Transient deviations from this tonic signal constitute reinforcement learning signals (Wickens et al., 1996). The positive reinforcement learning signal N derives from excitatory phasic fluctuations of the dopamine signal above the baseline:

$$N^+ = [D - \bar{D} - \Gamma_N]^+. \tag{8}$$

The complementary negative reinforcement learning signal is derived from inhibitory phasic fluctuations of the dopamine signal below baseline:

$$N^- = [\bar{D} - D - \Gamma_N]^+. \tag{9}$$

Spectral timing in the striosomal cells is mediated by a number of interacting factors, which are represented by the simplified intracellular system of Equations 10–14. A model of spectral timing in the cerebellum has elsewhere proposed detailed biochemical correlates of this type of learning in terms of mGluR1, Ca^{2+}, Ca-dependent K^+ channels, and intracellular second messengers. See Fiala et al. (1996) for this biochemically detailed treatment. Here we simplify and adapt this model to provide a phenomenological account of intracellular processes that does not attempt to predict the exact concentrations of particular chemical species.

Subscript i indexes which CS activates the cells, whereas subscript j indexes the response rate of the jth population of cell sites in the striosomal cell. It is important to note that the model does not require a different cell for each CS at each response rate, or delay, which would lead to a combinatoric explosion. Instead, multiple CSs synapse onto a single set of striosomal cells that span a spectrum of delays. In addition, not all CSs may be represented. Ventral prefrontal cortex (which provides much of the striosomal input signals) seems to preferentially represent CSs that have some motivational salience (Tremblay and Schultz, 1999).

The spectrum-sharing property of the model is made possible by the intracellular rather than extracellular delay timing mechanism, which allows a dissociation between the cortical (CS)-to-striosomal connection strength and the striosomal cell fixed Ca spike delays. The possibility of interference among coactive CSs would still necessitate more than a single striosomal spectrum, possibly at different dendritic sites (cf. Fiala et al., 1996). Cell recordings in SNc, PPTN, ventral striatal, and limbic cortical cells during multiple overlapping stimulus-delayed reward tasks might elucidate the nature of cortical CS representations and the extent to which CS signals may converge or interfere with each other in the excitatory and inhibitory pathways. The model predicts that multiple excitatory CS signals converging on the same dopamine cell will elicit multiple dopamine bursts in the trained animal, provided that the CSs are not predictably paired during training. Likewise, the model predicts that multiple CSs converging on the same striosomal cell may impair the ability of that particular cell to predict later rewards in a series during overlapping tasks. These predictions have yet to be tested. The spectral timing dynamics of the model are defined as follows. Striosomal cell activity x_{ij} responds to the ith CS at rate r_j:

$$\frac{d}{dt} x_{ij} = r_j[-x_{ij} + (1 - x_{ij})I_i]. \tag{10}$$

To provide a range of adaptively timed Ca^{2+} spikes, the striosomal buildup rate parameter spans a range of values for a given set of cells:

$$r_j = \frac{\alpha_r}{\beta_r + j}, \quad j = 1, 2, \ldots, n. \quad (11)$$

The activities x_{ij} induce intracellular calcium dynamics to cause transient calcium spikes at delays that are determined by r_j. These Ca^{2+} spikes determine the times at which the corresponding cells can learn from a dopamine burst. In particular, quantity $[G_{ij}Y_{ij}]^+$ represents an intracellular Ca^{2+} spike (Grossberg and Merrill, 1992), where

$$\frac{d}{dt} G_{ij} = \alpha_G(B_G - G_{ij})f_G(x_{ij} - \Gamma_G) - \beta_G G_{ij} \quad (12)$$

and

$$\frac{d}{dt} Y_{ij} = \alpha_Y(1 - Y_{ij}) - \beta_y[G_{ij}Y_{ij} - \Gamma_Y]^+. \quad (13)$$

In Equation 12, $f_G(x)$ is a step function: 0 for $x < 0$, 1 for $x > 0$. Parameters Γ_G and Γ_Y in Equations 12 and 13 are signal thresholds. When G_{ij} is activated by suprathreshold striosomal cell firing at a rate that varies with r_j, it rapidly increases the intracellular Ca^{2+}. As the calcium concentration rises to its maximal level, the available Ca^{2+} (Y_{ij}) rapidly decreases, causing a rapid falloff in the Ca^{2+} concentration. The Ca^{2+} concentration remains low as long as the mGluR1 receptors receive tonic input. Subsequent Ca spikes occur only when the tonic input is removed long enough for reset, in which the mGluR1 receptor and available Ca return to baseline. In the brief interval when the calcium concentration exceeds the activity threshold Γ_S in Equation 6, striosomal cell transmitter release is significantly enhanced, and the CS–striosomal weight Z_{ij} is potentiated via LTP if a dopamine burst is received:

$$\frac{d}{dt} Z_{ij} = \alpha_z[G_{ij}Y_{ij} - \Gamma_S]^+(-Z_{ij} + \gamma_S(N^+ + N^-)). \quad (14)$$

Simulated spike trains were generated with an integrate-and-fire (IAF) model using the cell membrane potentials M as input (defined for cells in Eq. 1, 3, 6, and 10 above, by variables S, P, D, and x_{ij}, respectively, and shown in Figs. 3*B* (*S*), 3*A* (*P*), 1 and 2 (*D*), and 3*C* (x_{ij})):

$$\frac{d}{dt} V = \frac{M + \epsilon}{C} - \frac{1}{RC} V. \quad (15)$$

The noise term ϵ was Gaussian with variance σ^2_{noise}. When the voltage exceeded a threshold V_I value, a spike was generated, and the voltage was reset to 0. Model outputs were computed from the model spiking response for 20 trials, and the model spikes were grouped into 20 msec-wide bins to compute histograms. The default IAF parameters (Table 2) were $V_I = 0.5$, $R = 1333$, $C = 0.025$, $\sigma_{noise} = 0.4$, except that for the dopamine cell, $R = 80$; for the PPTN cell, $R = 6667$, $C = 0.005$, and $\sigma_{noise} = 0.1$. The different R and C values were necessary to model the different firing properties of the cells.

The model performed a series of simulated learning trials. Each trial lasted 10 sec. The CS was active for 2 sec, and the R was active for 750 msec during the CS, beginning 1.2 sec after CS onset. Numerical integration was performed with an adaptive step size fourth-order Runge-Kutta method except for the IAF model, which used a first-order method and a discrete stepsize of 0.001 sec. The adaptive stepsize output was converted to a fixed stepsize by linear interpolation, so that it could be used to drive the IAF model. The CS was active from $t = 2$ sec into the trial, and it shut off when the primary reward signal shut off, or after $t = 3.95$, whichever was earlier. The primary reward signal typically began at $t = 3.2$ and lasted for 750 msec, with a magnitude of 1.0. The CS input (I_{CS}) had an amplitude of 0.6.

REFERENCES

Berns G, Sejnowski T (1998) A computational model of how the basal ganglia produce sequences. J Cognit Neurosci 10:108–121.

Berridge K, Robinson T (1998) What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Res Rev 28:309–369.

Brog J, Salyapongse A, Deutch A, Zahm D (1993) The patterns of afferent innervation of the core and shell in the "accumbens" part of the rat ventral striatum: immunohistochemical detection of retrogradely transported fluoro-gold. J Comp Neurol 338:255–278.

Buonomano DV, Mauk MD (1994) Neural network model of the cerebellum: temporal discrimination and the timing of motor responses. Neural Comput 6:38–55.

Calabresi P, Maj R, Pisani A, Mercuri N, Bernardi G (1992a) Long-term synaptic depression in the striatum: physiological and pharmacological characterization. J Neurosci 12:4224–4233.

Calabresi P, Pisani A, Mercuri N, Bernardi G (1992b) Long-term potentiation in the striatum is unmasked by removing the voltage-dependent magnesium block of NMDA receptor channels. Eur J Neurosci 4:929–935.

Calabresi P, Pisani A, Mercuri N, Bernardi G (1994) Post-receptor mechanisms underlying striatal long-term depression. J Neurosci 14:4871–4881.

Carpenter G, Grossberg S (1987) ART 2: Self-organization of stable category recognition codes for analog input patterns. Appl Optics 26:4919–4930.

Conde H (1992) Organization and physiology of the substantia nigra. Exp Brain Res 88:233–248.

Conde H, Dormont J, Farin D (1998) The role of the pedunculopontine tegmental nucleus in relation to conditioned motor performance in the cat. II. Effects of reversible inactivation by intracerebral microinjections. Exp Brain Res 121:411–418.

Contreras-Vidal J, Schultz W (1997) A predictive reinforcement model of dopamine neurons for learning approach behavior. First International Conference on Vision, Recognition, and Action: Neural Models of Mind and Machine. Department of Cognitive and Neural Systems, Boston University, Boston, MA, May 1997.

Crepel F, Hemart N, Jaillard D, Daniel H (1996) Cellular mechanisms of long-term depression in the cerebellum. Behav Brain Sci 19:347–353.

Dias R, Robbins T, Roberts A (1996) Dissociation in prefrontal cortex of affective and attentional shifts. Nature 380:69–72.

Dormont J, Conde H, Farin D (1998) The role of the pedunculopontine tegmental nucleus in relation to conditioned motor performance in the cat I. Context-dependent and reinforcement-related single unit activity. Exp Brain Res 121:401–410.

Eblen F, Graybiel A (1995) Highly restricted origin of prefrontal cortical inputs to striosomes in the macaque monkey. J Neurosci 15:5999–6013.

Fiala J, Grossberg S, Bullock D (1996) Metabotropic glutamate receptor activation in cerebellar purkinje cells as substrate for adaptive timing of the classically conditioned eye-blink response. J Neurosci 16:3760–3774.

Finch EA, Augustine GJ (1998) Local calcium signalling by inositol-1,4,5-trisphosphate in Purkinje cell dendrites. Nature 396:753–756.

Funahashi S, Bruce CJ, Goldman-Rakic PS (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. J Neurophysiol 61:331–349.

Futami T, Takakusaki K, Kitai S (1995) Glutamatergic and cholinergic inputs from the pedunculopontine tegmental nucleus to dopamine neurons in the substantia nigra pars compacta. Neurosci Res 21:331–342.

Gallagher M, Chiba A (1996) The amygdala and emotion. Curr Opin Neurobiol 6:221–227.

Garris PA, Kilpatrick M, Bunin MA, Michael D, Walker QD, Wightman RM (1999) Dissociation of dopamine release in the nucleus accumbens from intracranial self-stimulation. Nature 398:67–69.

Gerfen C (1992) The neostriatal mosaic: multiple levels of compartmental organization in the basal ganglia. Annu Rev Neurosci 15:285–320.

Grossberg S, Merrill J (1992) A neural network model of adaptively timed reinforcement learning and hippocampal dynamics. Cognit Brain Res 1:3–38.

Grossberg S, Merrill J (1996) The hippocampus and cerebellum in adaptively timed learning, recognition, and movement. J Cogn Neurosci 8:257–277.

Grossberg S, Schmajuk N (1989) Neural dynamics of adaptive timing and temporal discrimination during associative learning. Neural Networks 2:79–102.

Hajos M, Greenfield S (1994) Synaptic connections between pars compacta and pars reticulata neurones: electrophysiological evidence for functional modules within the substantia nigra. Brain Res 660:216–224.

Hollerman J, Schultz W (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. Nat Neurosci 1:304–309.

Houk J, Adams J, Barto A (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 (Houk J, Davis J, Beiser D, eds), pp 249–270. Cambridge, MA: MIT.

Jaeger D, Kita H, Wilson C (1994) Surround inhibition among projections neurons is weak or nonexistent in the rat neostriatum. J Neurosci 72:2555–2558.

Joiner WJ, Tang MD, Wang LY, Dworetzky SI, Boissard CG, Gan L, Gribkoff VK, Kaczmarek LK (1998) Formation of intermediate-conductance calcium-activated potassium channels by interaction of Slack and Slo subunits. Nat Neurosci 1:462–469.

Kita H (1992) Responses of globus pallidus neurons to cortical stimulation: intracellular study in the rat. Brain Res 589:84–90.

Kita H, Kitai S (1991) Intracellular study of rat globus pallidus neurons: membrane properties and responses to neostriatal, subthalamic and nigral stimulation. Brain Res 564:296–305.

Kojima J, Yamaji Y, Matsumura M, Nambu A, Inase M, Tokuno H, Takada M, Imai H (1997) Excitotoxic lesions of the pedunculopontine tegmental nucleus produce contralateral hemiparkinsonism in the monkey. Neurosci Lett 226:111–114.

Ljungberg T, Apicella P, Schultz W (1992) Responses of monkey dopamine neurons during learning of behavioral reactions. J Neurophysiol 67:145–163.

McDonald R, White N (1993) A triple dissociation of memory systems: hippocampus, amygdala, and dorsal striatum. Behav Neurosci 107:3–22.

Mirenowicz J, Schultz W (1994) Importance of unpredictability for reward responses in primate dopamine neurons. J Neurophysiol 72:1024–1027.

Mogenson G, Wu M (1986) Subpallidal projections to the mesencephalic locomotor region investigated with a combination of behavioral and electrophysiological recording techniques. Brain Res Bull 16:383–390.

Montague P, Dayan P, Sejnowski T (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J Neurosci 16:1936–1947.

Nakamura K, Ono T (1986) Lateral hypothalamus neuron involvement in integration of natural and artificial rewards and cue signals. J Neurophysiol 55:163–181.

Pisani A, Calabresi P, Centonze D, Bernardi G (1996) Enhancement of NMDA responses by group I metabotropic glutamate receptor activation in striatal neurones. Br J Pharmacol 120:1007–1014.

Scarnati E, Hajdu F, Pacitti C, Tombol T (1988) An EM and Golgi study on the connection between the nucleus tegmenti pedunculopontinus and the pars compacta of the substantia nigra in the rat. J Hirnforsch 29:95–105.

Schultz W (1998) Predictive reward signal of dopamine neurons. J Neurophysiol 80:1–27.

Schultz W, Apicelli P, Scarnati E, Ljungberg T (1992) Neuronal activity in monkey ventral striatum related to the expectation of reward. J Neurosci 12:4595–4610.

Schultz W, Apicella P, Ljungberg T (1993) Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. J Neurosci 13:900–913.

Schultz W, Romo R, Ljungberg T, Mirenowicz J, Hollerman J, Dickinson A (1995) Reward-related signals carried by dopamine neurons. In: Models of information processing in the basal ganglia (Houk J, Davis J, Beiser D eds), pp 11–27. Cambridge, MA: MIT.

Schultz W, Dayan P, Montague P (1997) A neural substrate of prediction and reward. Science 275:1593–1598.

Semba K, Fibiger H (1992) Afferent connections of the laterodorsal and the pedunculopontine tegmental nuclei in the rat: a retro- and anterograde transport and immunohistochemical study. J Comp Neurol 323:387–410.

Smith I, Grace A (1992) Role of the subthalamic nucleus in the regulation of nigral dopamine neuron activity. Synapse 12:287–303.

Suri R, Schultz W (1998) Learning of sequential movements by neural network model with dopamine-like reinforcement signal. Exp Brain Res 121:350–354.

Takakusaki K, Shiroyama T, Kitai S (1997) Two types of cholinergic neurons in the rat tegmental pedunculopontine nucleus: electrophysiological and morphological characterization. Neuroscience 79:1089–1109.

Takechi H, Eilers J, Konnerth A (1998) A new class of synaptic response involving calcium release in dendritic spines. Nature 396:757–760.

Tepper J, Martin L, Anderson D (1995) GABAA receptor-mediated inhibition of rat substantia nigra dopaminergic neurons by pars reticulata projection neurons. J Neurosci 15:3092–3103.

Tremblay L, Schultz W (1999) Relative reward preference in primate orbitofrontal cortex. Nature 398:704–708.

Wickens J, Kotter R (1995) Cellular models of reinforcement. In: Models of information processing in the basal ganglia (Houk J, Davis J, Beiser D, eds), pp 187–214. Cambridge, MA: MIT.

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

Wilkie DM (1987) Stimulus intensity affects pigeons' timing behavior: implications for an internal clock model. Anim Learn Behav 15:35–39.

Wilson C 1995 The contribution of cortical neurons to the firing pattern of striatal spiny neurons. In: Models of information processing in the basal ganglia (Houk J, Davis J, Beiser D, eds), pp 29–50. Cambridge, MA: MIT.

Yang C, Mogenson G (1987) Hippocampal signal transmission to the pedunculopontine nucleus and its regulation by dopamine D2 receptors in the nucleus accumbens: an electrophysiological and behavioural study. Neuroscience 23:1041–1055.