31

On the Control of Control: The Role of Dopamine in Regulating Prefrontal Function and Working Memory

Todd S. Braver and Jonathan D. Cohen

ABSTRACT An important aspect of cognitive control is the ability to appropriately select, update, and maintain contextual information related to behavioral goals, and to use this information to coordinate processing over extended periods. In our novel, neurobiologically based, connectionist computational model, the selection, updating, and maintenance of context occur through interactions between the prefrontal cortex (PFC) and dopamine (DA) neurotransmitter system. Phasic DA activity serves two simultaneous and synergistic functions: (1) a gating function, which regulates the access of information to active memory mechanisms subserved by PFC; and (2) a learning function, which allows the system to discover what information is relevant for selection as context. We present a simulation that establishes the computational viability of these postulated neurobiological mechanisms for subserving control functions.

The need for a control mechanism in cognition has been long noted within psychology. Virtually all theorists agree that some mechanism is needed to guide, coordinate, and update behavior in a flexible fashion—particularly in novel or complex tasks (Norman and Shallice 1986). In particular, control over processing requires that information related both to current context and to behavioral goals be actively represented, such that these representations can bias behavior in favor of goal-directed activities over extended periods. Indeed, most computationally explicit theories of human behavior have included such a mechanism as a fundamental component. For example, in production system models, goal states represented in declarative memory are used to coordinate the sequence of production firings involved in complex behaviors (e.g., Anderson 1983). One critical feature of goal representations in production systems is that they must be actively represented and maintained throughout the course of a sequence of behaviors. Such formulations of a control (or "executive") mechanism closely parallel theorizing about the nature of frontal lobe function (Bianchi 1922; Damasio 1985; Luria 1969), and clinical observations of patients with frontal lesions who often exhibit impairments in tasks requiring control over behavior—the so-called dysexecutive syndrome. Shallice (Norman and Shallice 1986; Shallice, 1982, 1988) explicitly noted this relationship, using the production system framework to describe his theory of a "supervisory attentional system" (SAS) as a mechanism by which the frontal lobes coordinate complex cognitive processes and select nonroutine actions. While these efforts have provided insights into the types of processes that may be engaged by cognitive control, they do not map transparently onto underlying neural mechanisms. They have also not fully addressed several critical issues, such as how a control system can develop through learning.

A number of recently proposed connectionist models of prefrontal function incorporate some of the central features of control processes in production system models, such as the active maintenance of goal representations (Dehaene and Changeux 1992; Guigon et al. 1991; Levine and Prueitt 1989). Connectionist models have the advantage of both being mechanistically explicit and using a computational architecture that maps more naturally onto neural mechanisms than traditional production system models. In this chapter, we report on work that uses this framework to address a critical question about cognitive control: How can a system learn to choose and appropriately update representations in active memory that can be used to control behavior? This is an extension of our ongoing effort to specify the neural underpinnings of cognitive control (Braver et al. 1995a; Cohen, Braver, and O'Reilly 1996; Cohen and Servan-Schreiber 1992), reviewed briefly below as background.

A central hypothesis in our work is that a cardinal function of prefrontal cortex (PFC) is to actively maintain context information. We use the general term *context* to include not only goal representations, which have their influence on planning and overt behavior, but also representations that may have their effect earlier in the processing stream, on interpretive or attentional processes. We assume that a primary function of PFC is to maintain task-relevant context representations in an active state. These active context representations serve to mediate control by modulating the flow of information within task-specific pathways such that processing in the task-relevant pathway is favored over a (possibly stronger) competing pathway. This function of PFC can also be thought of as a component of *working memory* (WM), commonly defined as the collection of mechanisms responsible for the on-line maintenance and manipulation of information necessary to perform a cognitive task (Baddeley and Hitch 1994). From this perspective, context can be viewed as the subset of representations within WM that govern how other representations are used.

As noted above, there is long-standing recognition that control involves representation and maintenance of context information (e.g., goals). However, a more complete account of cognitive control has additional requirements. Here we focus on four. Context information must be (1) appropriately selected for maintenance; (2) held for arbitrary lengths of time; (3) protected against interference; and (4) updated at appropriate junctures. Inasmuch as we assume that context information is represented in PFC, our interest is in the mechanisms that regulate the selection and updating of representations in PFC. One type of system meeting these requirements uses a gating mechanism to regulate the flow of activity into PFC: when the gate is opened, activity can flow into the PFC and activate the appropriate context representations; when the gate is closed, the activated representations are protected from interference, and therefore can be maintained and exert control for extended periods. Such a system, however, must know when it is appropriate to deploy the gate. This additional requirement threatens to introduce a regress in the control of processing: If the gating mechanism controls the controller, "who" is controlling the gating mechanism? Moreover, how can this component of control be learned, and how can this be mediated in a neurobiologically plausible way?

In this chapter, we propose a computational and neurobiological solution to this dilemma that involves the dopamine (DA) neurotransmitter system. Specifically, we suggest that DA projections to PFC serve to gate access of context representations into active memory through simple neuromodulatory effects on processing units in the PFC. These effects serve both gating and learning functions, which enable the system to discover what information must be maintained for performing a given task, and to regulate when that information is updated. This avoids the "homunculus" that plagues many theories of executive control. Below, we review evidence for this hypothesis, including evidence that PFC supports active memory, computational analyses of simple and gated active memory systems, and evidence that the modulatory effects of DA can support both its gating and learning functions. Following this review, we present a simulation that establishes the model's computational viability.

31.1 A NEURALLY BASED ACCOUNT OF THE CONTROL OF ACTIVE MEMORY

Prefrontal Cortex and Control

Neurobiological Evidence The role of control mechanisms in PFC has long been suggested by neuropsychological evidence. Increased distractibility and perseveration are hallmarks of neurological damage to PFC (Damasio 1985; Engle, Kane, and Tuholski 1999; Milner 1963; Owen et al. 1991; Stuss and Benson 1986) and of psychiatric disorders known to involve PFC such as schizophrenia (Malmo 1974; Nuechterlein and Dawson 1984). Neurophysiological studies have begun to provide a more detailed characterization of PFC function. Miller (chap. 22, this volume) provides an excellent review of this literature, which demonstrates that units in PFC (1) selectively code information relevant to task performance and not distractor information; (2) can code multimodal, task-relevant contingencies (including sensory information from different modalities and sensorimotor mappings); (3) can maintain such information over extended delays, in the absence of sustained sensory input; and (4) exhibit a pattern of temporal dynamics that suggests they are the source of attentional bias in posterior systems directly responsible for sensory and motor processing. These findings are consistent with the control function that we have ascribed to PFC. Recent neuroimaging studies using event-related fMRI have begun to corroborate these neurophysiological findings in human subjects, demonstrating sustained activity of PFC during delay intervals in working memory tasks (Cohen et al. 1997; Courtney et al. 1997) and in tasks that engage the "executive" functions of working memory (D'Esposito and Postle, chap. 15, this volume; Frith, chap. 24, this volume)

Computational Analysis As noted above, we have hypothesized that PFC exerts control by biasing processing in the pathways responsible for task performance. This biasing function is illustrated by Cohen and colleagues' previous models of the Stroop task (Cohen, Dunbar, and McClelland 1990; Cohen and Huston 1994; Cohen and Servan-Schreiber 1992), in which activation of a context representation corresponding to the relevant task dimension (e.g., color) sends activity to all the hidden units in the processing pathway corresponding to that dimension. This favors the flow of activity along that pathway, allowing it to compete effectively with information flowing along an otherwise stronger but irrelevant pathway (word naming). Thus activation of the context representation biases processing in favor of the task-relevant dimension, establishing the sensorimotor mapping necessary to perform the task.

For context representations to bias processing, however, they must be actively maintained for the duration of the task. Although the previous models noted above did not include a mechanism for doing so; a number of mechanisms can support the short-term maintenance of information in connectionist models. The most commonly employed and best understood of these are fixed-point attractor networks (e.g., Hopfield 1982; Zipser 1991), which possess recurrent connections that "recirculate" activation among units, and are thus capable of supporting sustained activity. Such networks typically settle into stable states called "attractors," in which a particular pattern of activity is maintained, and which therefore can be used to store information actively. A number of computational models of simple maintenance tasks have demonstrated that both physiological and behavioral data regarding PFC function can be captured using attractor networks (Braver, Cohen, and Servan-Schreiber 1995a; Dehaene and Changeux 1989; Moody et al. 1998; Zipser et al. 1993).

On the other hand, simple attractor systems have limitations that pose problems in more realistic tasks. The state of an attractor system is determined by its inputs, so that presentation of any new input will drive the system into a new attractor state, overwriting previously stored information (Bengio, Frasconi, and Simard 1993; Mozer 1993), and making the system subject to interference from task-irrelevant inputs (i.e., distractors). Although attractor networks can be configured to display resistance to disruption from distractors (i.e., hysteresis), this impairs their ability to be easily updated. One way in which attractor networks can overcome these difficulties is through the addition of a gating mechanism. Gated networks respond to inputs, changing their attractor state only when the "gate" is opened. Compared to other types of recurrent networks, networks with a gating mechanism were found better able to learn and perform complex short-term memory tasks, especially when the tasks involved noisy environments, frequent updating, and relatively long periods of storage (Hochreiter and Schmidhuber 1997). These and other computational studies suggest that gated attractor systems can meet many of the requirements for active memory in a control system. Moreover, the physiological evidence reviewed above is consistent with the hypothesis that prefrontal cortex implements such a system. Zipser and colleagues (Moody et al. 1998; Zipser 1991; Zipser et al. 1993) have proposed gated attractor models of short-term memory, and have used these to simulate the patterns of delay period activity observed for PFC neurons, although these models have specified neither the source of the gating signal nor how its timing is learned.

Dopamine Modulation of Information Processing

Dopamine and Cognitive Control There has been a growing appreciation of the role of dopamine (DA) in higher cognitive function (see Robbins and Rogers, chap. 21, this volume). Several lines of evidence have shown a link between DA function and cognitive control. These include studies of cognitive deficits in patients suffering from brain disorders involving DA pathology, such as Parkinson's disease and schizophrenia (e.g., Cohen et al. 1999; Gold 1992; Robbins et al. 1994), pharmacological studies manipulating DA activity locally in the PFC of nonhuman primates (Brozoski et al. 1979; Sawaguchi and Goldman-Rakic 1991, 1994; Sawaguchi, Matsumura, and Kubota 1990), and systemic manipulation of DA in humans (Kimberg, D'Esposito, and Farah 1997; Luciana, Collins, and Depue 1995; Luciana et al. 1992; Servan-Schreiber et al. 1998). Based on these findings, several authors have proposed that DA activity serves to modulate the cognitive control functions mediated by PFC (Cohen and Servan-Schreiber 1992; Goldman-Rakic and Selemon 1997). Here, we extend this idea, by proposing more specifically that the DA system provides a mechanism for learning to predict reward and to update the contents of active memory correspondingly, so as to maximize the chance of receiving reward. We propose that this function is carried out by simple, but appropriately timed neuromodulatory effects on target neurons. We hypothesize that one effect of DA is to modulate the responsivity of PFC units to their input, allowing DA to gate inputs to PFC. Another effect of DA is to modulate the strength of the connection between these inputs and the DA neurons themselves, allowing the DA system to discover what information should trigger this gate, and thereby to update the contents of active memory in PFC appropriately. There is a substantial corpus of neurobiological data to support this view of DA function.

Modulatory Effects of Dopamine Like other catecholamines, dopamine is known to produce modulatory effects on target neurons (Chiodo and Berger 1986; Hernandez-Lopez et al. 1997; Penit-Soria, Audinat, and Crepel 1987). Our previous models, by implementing this neuromodulatory action as a change in the slope (or gain) of the activation function of processing units, have simulated a variety of the effects of DA, at both the physiological and behavioral levels (Braver, Cohen, and Servan-Schreiber 1995a; Cohen and Servan-Schreiber 1993; Servan-Schreiber et al. 1998; Servan-Schreiber, Printz, and Cohen 1990). A change in gain modulates the responsivity of units to their afferent input, and thus can function as a gate on the flow of activity into PFC. Detailed anatomic studies of PFC suggest that DA projections are well positioned to influence both excitatory inputs and local inhibitory interactions (Lewis et al. 1992; Sesack, Snyder, and Lewis 1995; Williams and Goldman-Rakic 1993), a pattern that is consistent with a role of DA in gating PFC (discussed below). Furthermore, although neuromodulatory influences are typically assumed to be slow acting and nonspecific in information content (Moore and Bloom 1978), recent findings have suggested that DA cells can exhibit fast and stimulus-specific responses, as required to serve a gating function (Grace 1991; Schultz, Apicella, and Ljungberg 1993).

Timing of Dopamine Responses Schultz and colleagues (Schultz 1992) have observed rapid, stimulus-locked and stimulus-specific activity in DA neurons (~100 msec in duration, occurring 80–150 msec after stimulus onset). For example, following training in a spatial delayed-response task requiring active maintenance (Schultz, Apicella, and Ljungberg 1993), DA cells came to respond to the cue to be maintained. The cue was the first stimulus in the sequence that itself was unpredictable, but that predicted subsequent reward (even when there were intervening distractors). This is precisely the timing that might be expected of a control mechanism responsible for updating context representations. When an unexpected cue indicates that a new desired state can be achieved, then this cue should elicit an updating of the context representation (e.g., goal) in active memory, replacing the current representation with one that will guide behavior toward the desired state.

Learning effects of Dopamine Findings from reward-conditioning paradigms suggest how the gating signal could be learned. DA has long been recognized to play a role in reward learning (Wise and Rompre 1989). In the Schulz and colleagues studies referred to above, DA responds initially only to the rewarding event, but with training this response "migrates" to predictive cues. Montague, Dayan, and Sejnowski (1996) have proposed a formal analysis of the role of DA in reward conditioning, in terms of a temporal difference (TD) learning algorithm (Sutton 1988; Sutton and Barto 1990). The TD algorithm provides a mechanism by which learning can chain backward in time, allowing the DA system to identify successively earlier predictors of reward, until the earliest possible predictor is found that cannot itself be predicted. In the Montague, Dayan, and Sejnowski model, DA responses are simulated as being proportional to the prediction error in the TD algorithm (i.e., the degree of mismatch between expected and received rewards), and DA release modulates the strength of synapses from units representing cues that predict reward to the DA units themselves.[1] In simulations as in empirical studies, the DA response decreases to events as they become more predictable (e.g., an expected reward), whereas it increases to events that predict reward but are themselves unpredicted. Intriguingly, the parameter used by Montague, Dayan, and Sejnowski to simulate the effects of DA on learning is analogous to the parameter we have used to simulate DA effects on unit responsivity. This raises the possibility that a single parameter can be used to account for both effects, which may occur simultaneously, in turn providing a means by which the gating signal can be learned.

A New Theory Although we have previously theorized that PFC is critical for the active maintenance of context information, and that DA activity serves to modulate the responsivity of PFC neurons to external input (Cohen and Servan-Schreiber 1992), the findings just discussed suggest a number of hypotheses revising and extending our original theory. These hypotheses provide an account of both the ability to update context representations and the means of learning how to do so:

- Context representations are actively maintained in a gated attractor system within PFC.
- Phasic changes in DA activity serve two functions:
  1. to gate information into active memory in PFC;
  2. to strengthen associations between stimuli that predict reward and the DA response.
- Both effects rely on a similar neuromodulatory mechanism.
- The gating effect occurs through the transient potentiation of both excitatory afferent and local inhibitory effects in PFC.
- The learning effect occurs through modulation of synaptic weights, driven by errors between predicted and received rewards (i.e., the TD learning algorithm).

- The coincidence of the gating and learning signals produces cortical associations between the information being gated and a triggering of the gating signal in the future.

In the studies presented below, we test the plausibility of these claims in a computer simulation of a model that implements our theory. Specifically, the simulation examines the hypothesis, suggested in the previous two subsections, that appropriate timing of gating signals can be acquired during task performance through reward-based learning mechanisms.

31.2 SIMULATION: REWARD-BASED LEARNING OF GATING SIGNALS

This study was conducted to establish the computational validity of the hypothesis that DA implements both gating and learning effects, and that such a system can learn to appropriately gate relevant context information into active memory. Although previous work has demonstrated that DA activity can be simulated accurately in a system governed by reinforcement learning (Montague, Dayan, and Sejnowski 1996), it has not been shown that the dynamics of DA activity can simultaneously be exploited as (and used to learn the timing of) a gating signal. Furthermore, this hypothesis poses the following dilemma. If gating the appropriate context representations into active memory is learned through a reward-based mechanism, but reward itself depends on gating the appropriate context representations, then how can the process get started? This is a classic "bootstrapping" problem, solutions for which are often best demonstrated by simulation. To do so, we constructed a model of a simple cognitive control task, where context information must be actively maintained across delay periods during which intervening distractor events may occur, and properly updated on a trial-to-trial basis.

Task

We used a variant of a delayed-response paradigm (the AX version of the continuous performance test, or AX-CPT; Nuechterlein and Dawson 1984; Rosvold et al. 1956) that we have used extensively to study the processing of context and its relationship to PFC and DA function in behavioral (Cohen et al. 1999; Cohen and Servan-Schreiber 1993; Servan-Schreiber, Cohen, and Steingard 1996), psychopharmacological (Braver 1997), and neuroimaging (Barch et al. 1997; Carter et al. 1998) studies. The AX-CPT paradigm has also been the subject of previous modeling work (Braver 1997; Braver, Cohen, and Servan-Schreiber 1995b; Cohen, Braver, and O'Reilly 1996). In this paradigm, a cue is presented at the beginning of each trial (e.g., the letter *A* or *B*), followed by a delay of variable length, and then a probe (e.g., the letter *X*) to which one of two responses must be made. The correct response to the probe is contingent on the identity

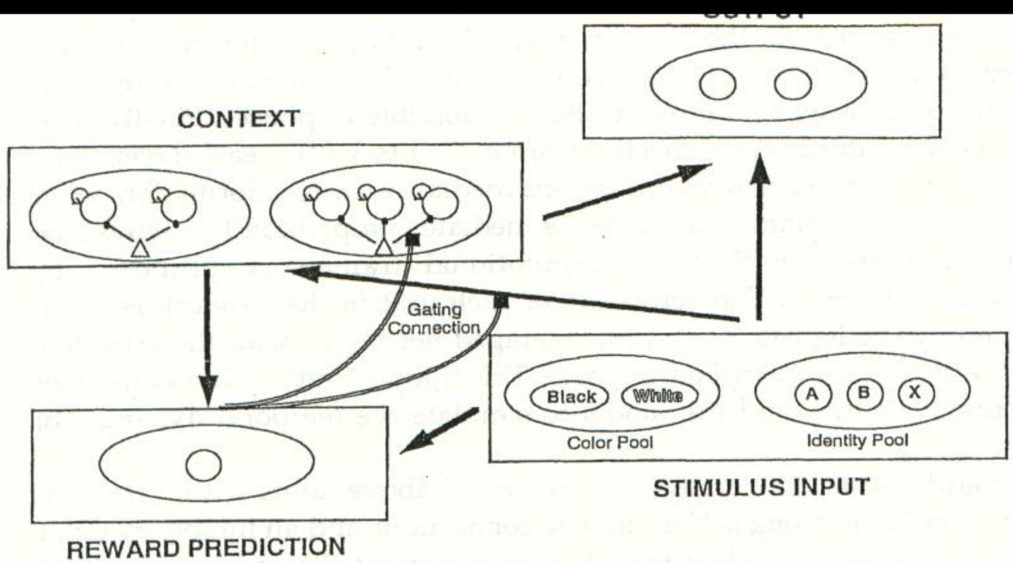

Figure 31.1 Learning/gating model used in simulation. Excitatory connections exist between layers (indicated by arrows), whereas lateral inhibitory connections exist within each layer (not shown). Input units make one-to-one connections with context layer units. Context units have self-excitatory connections, providing a mechanism for active maintenance. Low levels of baseline activity in the context layer are enforced by local inhibitory bias units (indicated by small triangles). The input and context layers are fully connected to the reward prediction/gating (RPG) unit. This unit, in turn, makes a gating connection with both afferent excitatory and local inhibitory input to the context layer. The RPG unit also modulates learning in all modifiable connections of the network.

of the cue. One reponse (e.g., press the left button) is required if the probe follows a specified cue (e.g., A-X, which we will refer to as "AX" trials), and the other response (e.g., right button) is required for all other cue-probe sequences (e.g., BX). Thus responding correctly to the probe requires maintenance of context information provided by the cue. Additionally, distractor stimuli are presented randomly, interspersed during the cue-probe delay and intertrial interval (ITI). Distractors are distinguished from the cue and probe stimuli by a particular feature (e.g., the color of the letters), but can have the same identity as the cue (e.g., A or B). Thus the AX-CPT paradigm engages cognitive control, insofar as correct performance requires the abilities to actively maintain context over a variable delay, ignore distractors, and update context selectively in response to cue stimuli but not distractors.

Architecture and Processing

Our model of this task is shown in figure 31.1. The network is composed of a stimulus layer (5 units), a context layer (5 units), a response layer (2 units), and a reward prediction/gating (RPG) unit. The stimulus and context layers are each separated into two pools, the first used to represent stimulus identity (*A*, *B*, *X*), and the second, stimulus color (black, white).

Units in the stimulus layer have one-to-one excitatory connections to corresponding units in the context layer. All units within the stimulus and context layers have excitatory connections to both units in the response (output) layer, which represent the two possible responses. Finally, there are lateral inhibitory connections among units within each layer. Thus between-layer excitatory connections mediate flow of information, while within-layer inhibitory connections mediate competition for representation, consistent with the computational framework proposed by McClelland (1993). The activation of each unit in the network is determined by the logistic of its time-averaged net input (with the exception of the RPG unit described below).[2] This allows units to integrate their inputs over time, and the model to simulate the temporal dynamics of processing.

In addition to the connectivity described above, units within the context layer have strong self-excitatory connections and an inhibitory input from a tonically active bias unit. This arrangement allows context units to assume a relatively low baseline of activity, yet self-sustain a higher level of activity following a sufficiently strong input, even after the input is removed.[3] We use this behavior to simulate active maintenance of context information in PFC. The weights of the one-to-one connections from the stimulus units to the context units, and among the context units, are fixed at values such that stimulus unit activity can activate context units when the entire context pool is at rest (i.e., no context units are active), but stimulus unit activity cannot alter an existing pattern of context unit activity.[4] Thus stimulus units are not able on their own to update the state of activated context units; this requires the "intervention" of the RPG unit (discussed below). The "hardwiring" of these connections reflects our assumption that the active maintenance properties of PFC, and its connections with task-specific processing pathways, arise by mechanisms different from the reward-based learning mechanisms described below, beyond the scope of current consideration.[5] The connection weights to and from the RPG unit and from the stimulus and context units to the output units are modifiable, and adjusted according to the learning rule described below.

The reward prediction/gating unit receives connections from all units in the stimulus and context layers. Its activity is computed as the weighted sum of the input received from the stimulus and context units on the current time step (current predicted reward) and the value of the actual reward for that trial (+1 for correct response and −1 for incorrect response) minus the stimulus and context input received on the previous time step (previously predicted reward), which is the temporal difference (TD) error.[6] The behavior of this unit serves as our simulation of phasic changes in dopamine activity, as in Montague, Dayan, and Sejnowski 1996. Accordingly, the activity of this unit (i.e., the value of the TD error) serves as a learning signal, used to adjust all modifiable weights in the network according to the TD learning algorithm.[7]

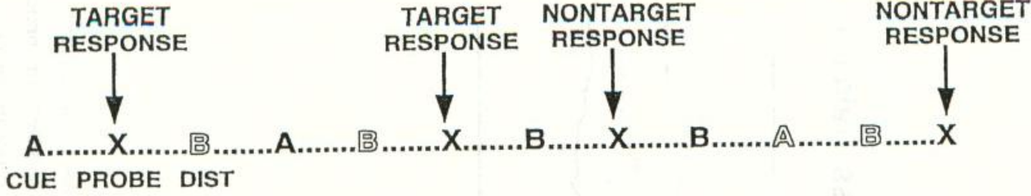

Figure 31.2 Typical sequence of training trials. Distractor stimuli are shown in outline type; task-relevant stimuli in solid type. Delay and intertrial interval (ITI) time steps are indicated by dashes. Distractors could occur during the delay period or ITI. Cue stimuli (*A* or *B*) need to be maintained over the delay in order to make the correct response to the probe (*X*)—one response when it follows an *A* cue, and a different response when it follows a *B* cue.

The RPG unit also exerts a gating effect on the context layer, allowing the current stimulus to change the state of (i.e., active representation in) the context layer. This occurs through potentiation of the strength of both afferent input (excitatory connections from stimulus to context) and local inhibition (inhibitory connections from the tonically active bias unit).[8] These potentiating effects have the following consequences. If a context unit, active when a gating signal occurs, does not receive excitation from any stimulus unit, but another context unit does, then the gating signal will favor activation of the competing context unit (due to potentiation of its excitatory input) and suppression of the current context unit (due to potentiation of inhibition from the competing context unit). Thus the gating influence of the RPG unit provides a mechanism for updating the state of activity in the context layer.

Training

We trained the network with a continuous sequence of task trials. (figure 31.2) Each trial consisted of the following events (simulated by activating the appropriate stimulus units): cue (*A* or *B*), delay interval, probe (*X*), and intertrial interval. Stimuli were presented for 3 time steps each; the minimum interval period was 7 time steps. Distractor events could be presented within both the delay and ITI, and each distractor increased the length of the interval by an additional 10 time steps (3 time steps for stimulus presentation +7 additional time steps for delay interval). The probability of a distractor appearing during any interval period was 0.50 for the first distractor in that period. The probability of an additional distractor appearing in that period decreased by half as the number of distractors increased (i.e., the probability of a second distractor appearing was 0.25, the probability of a third distractor appearing was 0.125, etc.). Each stimulus identity (*A* or *B*) was presented with equal frequency for both cues and distractors.

All modifiable weights were initialized to small random values prior to training (−0.25, 0.25). During training, weights were adjusted on every

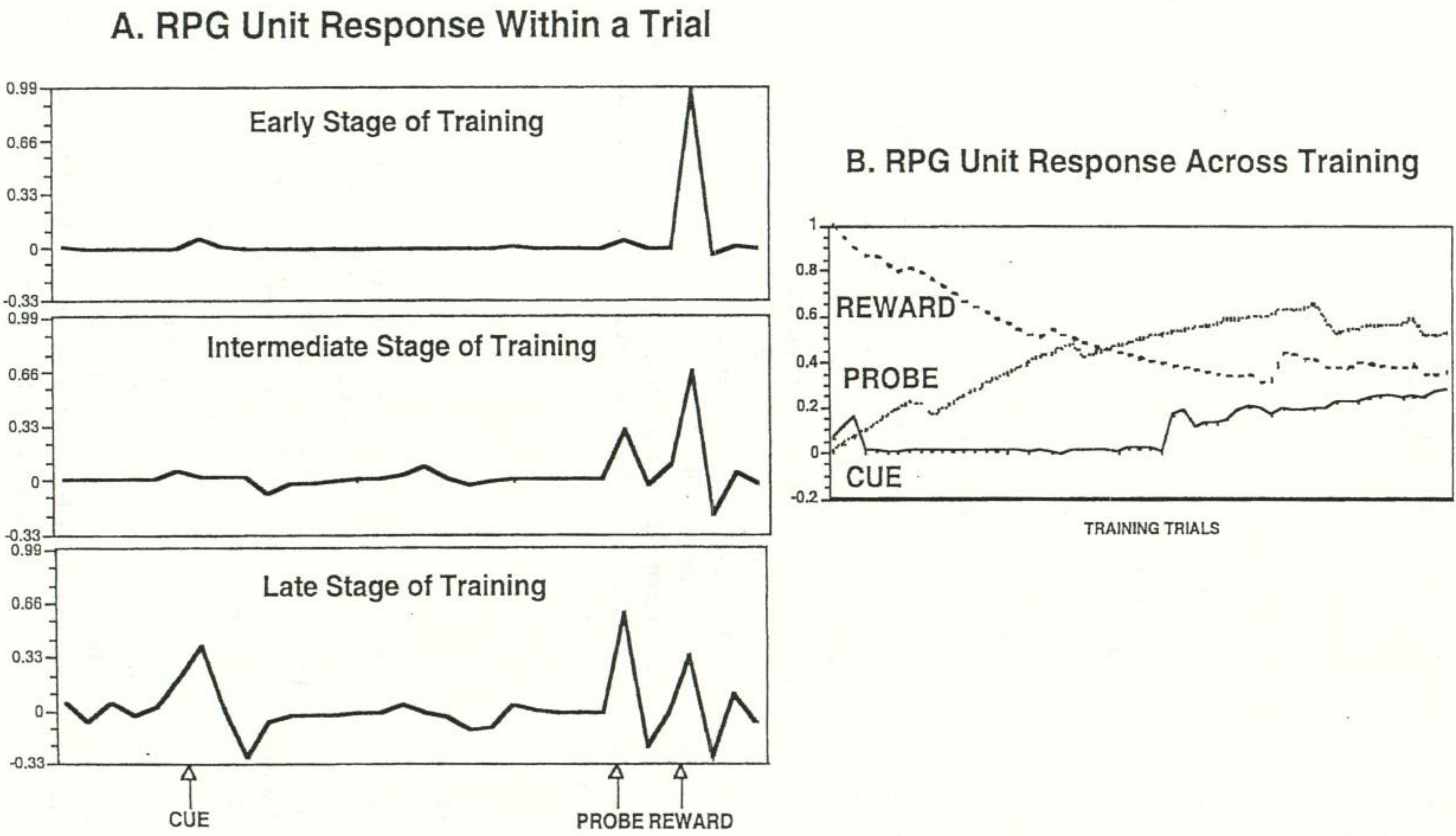

Figure 31.3 Reward prediction/gating (RPG) unit learning dynamics. A. RPG unit response (temporal difference error) within a trial in early, middle, and late stages of training. Note that in the early stage, the RPG unit only responds during reward delivery. In the middle stage, a response to the probe develops, whereas the reward response decreases. In the late stage, a response develops to the cue stimulus as well, whereas probe and reward responses continue to decrease. B. RPG unit response across training, where learning dynamics are shown continuously, as a function of the number of training trials.

time step in proportion to the activity of the RPG unit. Following presentation of the probe, the RPG unit received an input of +1 if the response was correct and −1 reinforcement if the response was incorrect, in addition to its usual input from the stimulus and context units. A response was considered correct if the activity of the left output unit was greater than 0.5 and greater than the right output unit for AX trials, and if the right output unit was greater than 0.5 and greater than the left output for BX trials. Thus, to perform the task correctly, the network had to learn to activate the context representation for the cue (*A* or *B*), maintain this over the delay, prevent distractor stimuli from disrupting this representation, and then use it to determine the correct response to the probe. During training, Gaussian noise was added to the net input of both context and output units, and was reduced in amplitude as error decreased (i.e., through a simple annealing schedule), consistent with the practice in other reinforcement learning simulations of having noise levels inversely related to the level of reward predicted (Gullapalli 1990).[9]

Results

Ten runs of the simulation were performed, each with randomly assigned initial weights for the modifiable connections in the network. The network converged to perfect performance on all ten runs. Learning followed a consistent pattern, comprising three stages (see figure 31.3). In the first stage, the connections from the stimulus and context units to the RPG unit remained weak, reflecting the lack of prediction or expectations of reward. Consequently, TD error (and the activity of the RPG unit) increased when reward was received because its delivery was unpredicted. In the intermediate stage, the stimulus unit for the probe (*X*) developed a positive connection with the RPG unit. Because reward (when it occurred) was delivered only following presentation of the probe, the network learned that the probe was a good predictor of reward. In reinforcement learning terms, the probe became a "secondary reinforcer," reducing the TD error (i.e., unexpectedness) at the time of reward delivery, and the response of the RPG unit to reward. Because the network had not yet learned to maintain the cue, however, the response to the probe was not always accurate, and reward was not delivered on every trial. Thus the probe was not a perfect predictor of reward, and a moderate level of TD error (and RPG unit activity) persisted for reward delivery. The third stage was reached when the TD algorithm allowed the network to learn the assocation between the cue stimuli and reward. Strong positive connections developed from the cue identity units (*A* and *B*) and the cue color unit (black) to the RPG unit, and a strong negative connection from the distractor color unit (white) to the RPG. As a consequence, activity in the RPG unit increased following presentation of a cue, but not following presentation of distractors. This increase in RPG unit activity produced a gating signal, which allowed the cue information to properly update the context representation, and be actively maintained over the delay. Moreover, because the cue information was being maintained as context, the context units began developing positive weights to the prediction unit, so that reward could be predicted based on the cue information. Once the cue information became a good predictor of reward, (because maintaining the cue increased the probability that reward was delivered), it became a "tertiary reinforcer," which further reduced the TD error both to the probe and reward delivery.

Note that noise in the context and output layers played a critical role in learning. In the output layer, noise encouraged response exploration, allowing the network to discover the correct response to the probe. Similarly, in the context layer, noise provided a way for the appropriate context unit to be active at the time of probe presentation (through random updating on some proportion of trials), before the network had learned to maintain the cue. This was critical for "bootstrapping" to take place. To summarize, the association between reward prediction and gating, coupled with noise, provided a mechanism for the network to discover how to regulate active memory so that cue information could selectively update the context representation.

Discussion

The results of this simulation provide preliminary support for the hypothesis that control over active maintenance of context representations can be achieved using a gating signal triggered by reward prediction errors. The pattern of RPG unit activity over the course of learning is very similar to that observed for DA neuronal activity over the course of learning in a delayed-response task (Schultz, Apicella, and Ljungberg 1993). In this respect, the results of our simulation replicate those of Montague, Dayan, and Sejnowski (1996), providing physiological support for the theory. However, our results go beyond those of Montague and colleagues, by demonstrating that the learning system can work synergistically with a gating signal to regulate control over active maintenance. By using the cue to predict reward, the network was also able to gate context information provided by the cue into active memory, where it could be used to bias subsequent responding. As a result, the probability of making the correct response was increased, and more rewards were achieved. Furthermore, because only cue stimuli elicited gating of the context layer, distractor stimuli were unable to disrupt the information maintained in the context layer. Thus the results also demonstrate that this type of control mechanism can protect context representations from the effects of interference. Moreover, the simulation makes clear how each of the two effects of the RPG unit are interdependent for learning the task properly. If RPG unit activity did not serve a gating function, the context representation would not be updated following cue presentation (or would be disrupted by every distractor). If the RPG unit activity did not modulate weight strengths based on reward prediction, the presentation of the cue input (*A* or *B* unit plus black unit) would never develop positive weights to the RPG unit, such that it could be activated by future cue presentations. Thus the simulations illustrate how both computational mechanisms associated with the RPG unit (gating and reward prediction learning) cooperate in the development of cognitive control over behavior in the task. The simulation also raises a number of more general conceptual issues regarding active maintenance, cognitive control, and reinforcement learning, which are discussed below.

Representation over Time A fundamental and unresolved issue in the application of reinforcement learning to classical and operant conditioning phenomena concerns the representation of perceptual information over time (Schultz, Dayan, and Montague 1997). For an organism to learn a relationship between a naturally reinforcing event (i.e., an unconditioned stimulus or US) and a predictive sensory cue (i.e., a conditioned stimulus or CS), the cue must still be represented when the reinforcement occurs. With very short delays, some perceptual trace of the cue may remain at the time of reinforcement. Although this is not likely at longer delays, when perceptual representations have presumably decayed.[10] To account for learning over such delays, some investigators (e.g., Sutton and Barto 1990) have proposed the mechanism of a decaying synaptic eligibility trace, which allows weights to be updated even when the cue is no longer actively represented. This does not solve an additional problem, however. Predictions of reward must continue at every time step from cue presentation until reward delivery for TD error to decrease and TD-based algorithms to function properly. Consequently, some form of active representation of the cue over an arbitrary period of time is required. Accordingly, most models of reinforcement learning represent each sensory cue as a vector, each element of which corresponds to the activity of that cue at a different point in time. In other words, the temporal dynamics of a cue are transformed into an explicit spatial representation (often referred to as a "complete serial compound" or CSC representation). Although it allows the system to learn an independent prediction of reward for every point in time (implemented as the connection strength from each element of the vector to the reward prediction unit), the CSC representation has a number of drawbacks, perhaps the most important of which is its neurobiological implausibility (Schultz, Dayan, and Montague 1997).

Our model implements a different solution to these problems. The context layer actively maintains representations that provide a continuous source of reward prediction necessary for TD learning to occur. Thus we propose that active maintenance within PFC may provide a mechanism for continuous reward prediction necessary for TD learning. As Hochreiter and Schmidhuber (1997) have observed, learning in difficult short-term memory tasks requires "constant error flow," which can be provided by computational units with activation that remains constant over time. One concern with such a solution, however, is that mechanisms for active maintenance must already be present for reward-based learning to occur. There are three principal ways that this could arise: (1) recurrent connectivity that develops as part of some intrinsic maturational process in PFC; (2) non-TD-based learning mechanisms that operate either prior to or interactively with reward-based learning (i.e., as another "bootstrapping" process); or (3) some other, innate mechanism (such as intrinsic bistability of neuronal activation states) that is preferentially expressed in PFC neurons. The available data do not adjudicate among these possibilities, although all three represent neurobiologically plausible mechanisms that are consistent with our model.

Alternative Control Mechanisms Another fundamental issue raised by the current study is whether gating is computationally required as a control mechanism for updating context representations. In principle, the answer is no. All that is required is a signal that differentiates task-relevant from task-irrelevant information and is derived in some form in the sensory input. This does not require a gating mechanism. For example, updating could occur through the proper conjunction of input features, previously maintained context representations, or both, coupled with the appropriate connection weights from input to context units (e.g., the conjunction of the *A* stimulus and the color black is sufficient to activate the *A* unit in the context layer and to overcome competition from other units in that layer). Thus, for any network that uses a gating signal to regulate access to active memory, an equivalent network can be constructed to perform the same functions without gating. There is a question, however, whether such a nongated network could be effectively learned through error-driven learning algorithms (either classical supervised or reward-based). The appropriate conjunction of weights required might be so precise as to be very difficult to learn using gradient descent procedures. We suspect that gated attractor networks coupled with TD learning provide a more powerful and robust computational mechanism for learning to perform tasks that require regulation of access to active memory. Although consistent with Hochreiter and Schmidhumber's analyses (1997) of simple recurrent networks and supervised learning algorithms, this conjecture remains to be tested for networks using TD learning to control the gating mechanism.

31.3 GENERAL DISCUSSION

In this chapter, we have presented a new model of the mechanisms underlying an important dimension of cognitive control: the ability to and the ability to learn how to do this. Furthermore, we have described simulation results that establish the computational plausibility of this model. On the other hand, the current model has important limitations, and significant challenges remain for a comprehensive theory of cognitive control. For example, we have not demonstrated that the mechanisms we propose can learn to gate into memory task-relevant information that itself is not directly predictive of reward. We have not provided an account of performance in more complex tasks, such as those which involve subgoaling. We have also not addressed the nature of context representations in the PFC—how these come about and how, without requiring infinite capacity, they can support the remarkable range and flexibility of behaviors of which humans are capable. These all remain challenges for further theoretical work.

Nevertheless, we believe that this model, even in its current limited form, has the potential to enrich our understanding of cognitive control. The model makes strong predictions about the engagement of PFC and DA in performance of simple control tasks, such as the AX-CPT, as well as the effects that disturbances of PFC and DA should have on task performance. We have begun to garner support for some of these predictions in related work using a wide variety of cognitive neuroscience methods. First, in behavioral studies, we have shown that patients with schizophrenia, who are thought to suffer from DA abnormalities in PFC, show a specific pattern of performance deficits in the AX-CPT consistent with a deficit in actively maintaining context (Braver, Barch, and Cohen 1999b; Cohen et al. 1999; Servan-Schreiber, Cohen, and Steingard 1996). Moreover, we have found a strikingly similar pattern of deficits in healthy subjects performing the AX-CPT under interference conditions (Braver, Barch, and Cohen 1999b). Second, in simulation studies we have found that the gating model can capture both of these patterns of deficits in terms of disturbances to the DA system (i.e., the reward prediction/gating unit). In particular, the model suggests that the deficits observed in schizophrenia might be due to increased noise in the RPG unit (Braver, Barch, and Cohen 1999a; Braver and Cohen 1999), while the deficits observed under interference can be captured by assuming that the distractor stimuli produce partial RPG unit activation (Braver, Cohen, and McClelland 1997). Third, preliminary results from a pharmacological study suggest that the interference-induced deficits in AX-CPT performance in healthy subjects may be ameliorated by low doses of amphetamine, a potent enhancer of DA transmission (Braver 1997). Finally, in functional neuroimaging studies, we have directly demonstrated the role of PFC in the active maintenance of context. During performance of the AX-CPT under conditions where the delay between cue and probe was manipulated, we observed greater PFC activity in long versus short delay trials, and further found that this activity was sustained throughout the delay period (Barch et al. 1997). In contrast, we observed that in the interference version of the AX-CPT, this activity is not sustained, but rather decays during the delay period, when distractors are presented (Braver, Barch, and Cohen 1999b).

Our model may also lead to new insights regarding cognitive control at the psychological level. For example, gated attractors may provide a useful theoretical framework within which to consider the effects of task switching that are addressed in detail in other contributions to this volume (e.g., Allport and Wylie, chap. 2, Jolicoeur, Dell'Acqua, and Crebolder, chap. 13, Goschke, chap. 14, De Jong, chap. 15, and Meiran, chap. 16, this volume). More generally, our model may help drive a reexamination of the relationship between motivational processes and cognitive control. The account of dopamine provided here suggests that it plays a unified role in motivation and cognition by configuring the system to optimize its predictions of reward and by regulating cognitive processes to increase the frequency with which rewards are obtained. This, in turn, offers an interesting perspective on prefrontal cortex function: the active maintenance of information in the service of maximizing rewards. From this perspective, one might imagine that PFC evolved at least in part to take control over the deployment of DA-mediated reinforcement by chaining together complex internal representations of reward prediction, and thus to support the construction of elaborate goal structures necessary for complex, temporally extended behaviors. This perspective suggests the intriguing possibility that the literature on the cognitive functions of PFC and DA can be linked with the growing, but heretofore separate, literatures on the affective and motivational functions of these brain systems (Bechara et al. 1996; Davidson and Sutton 1995; Willner and Scheel-Kruger 1991).

At the most general level, the model we have presented provides an illustration of how a system built of simple processing elements and general principles of learning can organize itself to regulate its own behavior in an adaptive fashion, without invoking the problem of a "homunculus." It also provides an example of how implementing a theory as an explicit computational model can lead to new and unexpected insights. Our hypotheses concerning the modulatory effects of dopamine (i.e., its role in gating) bear little surface resemblance to theories regarding the role of DA in reinforcement learning. It was only through a comparison of the formalisms of specific models that we were led to the observation that similar parameters were being used to implement these seemingly different DA effects, and to the idea that these effects may have synergistic effects. Our work also illustrates how efforts to understand the neural underpinnings of cognition can lead to insights at the psychological level. Our insights into the potential relationship between reward-based learning and gated attractors as a mechanism for the control of processing were driven in large measure by observations about the effects of a particular neurotransmitter and by efforts to account for its function. Thus, even in light of the limitations of our current model, we hope that our work may indicate how theories that draw simultaneously from, and bridge between, the neurobiological, psychological and computational domains can help advance our understanding of the mechanisms underlying cognitive control.

NOTES

1. The claim that dopamine modulates synaptic plasticity has received support in the neurophysiological literature (Calabresi et al. 1997; Law-Tho, Deuce, and Crepel 1995; Wickens, Begg, and Arbuthnott 1996).

2. The activation of unit a_i at time t is given as

$$a_i(t) = \frac{1}{1+e^{-\gamma \mathrm{net}_i(t)}},$$

where γ is the gain on the activation function, while $\mathrm{net}_i(t)$ is given as

$$\mathrm{net}_i(t) = \tau \sum_j a_j(t)\ wi_{ij} + (1-\tau)\ \mathrm{net}_i(t-1),$$

where τ is the time constant for averaging the net input (set at 0.5 for all simulations), and w_{ij} is the weight of the connection from each unit j that projects to unit i.

3. We should note that single, continuous-valued processing units in our model are used to simulate cell assemblies in the cortex (e.g., Amit 1989), and that recurrent self-connections simulate mutual excitatory synapses among cells belonging to a particular assembly.

4. Input-to-context module weights were set to $+3.0$; self-excitatory connections within the context module, to $+5.5$: lateral inhibitory connections within the context level, to -4.0; and local inhibitory input from the bias unit, to -2.5.

5. In work currently in progress, we have found preliminary evidence that both the active maintenance properties and connectivity pattern of context layer representations can be independently discovered through the application of learning algorithms, such as LEABRA (O'Reilly 1996), that combine correlational with error-driven learning. It remains a question for future research to determine whether this type of learning algorithm can be integrated with TD learning to provide more sophisticated models (i.e., ones that can address larger data sets and more complex cognitive tasks) and to reduce the number of parameters that need be fixed prior to learning.

6. More precisely, the TD error is computed according to the equation, derived from Sutton 1988:

$$\delta(t) = r(t) + \lambda P(t) - P(t-1),$$

where $r(t)$ is the reward input at time t, $P(t)$ is the total prediction input at time t, and λ is a discount factor, fixed at 0.95 for all simulations. This formulation suggests that an unexpected actual reward (for which predictions are zero) would lead to an increase in TD error (i.e., phasic activation of the RPG unit). Additionally, in the absence of actual reward (i.e., $r(t) = 0$), TD error increases when the current state is thought to be more predictive of reward than the previous state (i.e., $P(t) > P(t-1)$), such as when a salient cue appears in the environment.

7. Modifiable network weights are adjusted according to the learning rule:

$$\Delta w = \eta \delta(t) x_i(t-1),$$

where $x(t-1)$ is the activity of the sending unit at time $(t-1)$, η is the learning rate, and $\delta(t)$ is the TD error at time t (see note 6).

8. The modulatory effect of gating on afferent excitatory and local inhibitory input to context units was given as

$$w_{ij}' = \gamma(t) w_{ij},$$

where

$$\gamma(t) = 1 + \frac{k-1}{1+e^{-(S\delta(t)-0.5B)}} \quad k > 1,$$

and where $\delta(t)$ is the TD error of the gating unit at time t, with k determining the maximum gain (γ) of the gating unit. The function $\gamma(t)$, a sigmoid in which gain monotonically increases with the level of TD error, is bounded such that the minimum gain is 1 and the maximum is k. S and B are additional parameters that determine the slope of the sigmoid and its baseline value (i.e., when $\delta(t)$ is zero). In the simulation, $k = 5$, $S = 20$, and $B = 5$. The results of the model were not found to depend critically on these parameter values, although it was important to choose a parameter that caused the slope of the function to be relatively steep, such that small increases in $\delta(t)$ had a nonnegligible effect. This allowed the RPG unit to exert a gating function early in the learning process, when activity is not very high.

9. The noise was drawn from a Gaussian distribution having zero mean. Its standard deviation was initialized to a value of 0.2. During training, this value was decreased by half whenever the TD error at the time of reinforcement delivery (averaged across a moving window of ten trials) also decreased by half. The noise parameter and annealing schedule primarily affected the speed of learning, and the results of the simulation were not found to depend upon the exact values used.

10. Although there is evidence that some presumably perceptual regions, such as posterior parietal and inferior temporal cortex, do show sustained active representations of stimuli over delay periods, these representations appear to be abolished by the presentation of new stimuli (Constantinidis and Steinmetz 1996; Miller, Erickson, and Desimone 1996). Thus they cannot serve as generally useful temporal representations of the sort desired for reward prediction learning.

REFERENCES

Amit, D. J. (1989). *Modeling brain function: the world of attractor neural networks*. Cambridge: Cambridge University Press.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Baddeley, A. D., and Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology, 8*, 485–493.

Barch, D. M., Braver, T. S., Nystrom, L., Forman, S. D., Noll, D. C., and Cohen, J. D. (1997). Dissociating working memory from task difficulty in human prefrontal cortex. *Neuropsychologia, 35*, 1373–1380.

Bechara, A., Tranel, D., Damasio, H., and Damasio, A. R. (1996). Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cerebral Cortex, 6*, 215–225.

Bengio, Y., Frasconi, P., and Simard, P. (1993). The problem of learning long-term dependencies in recurrent networks. In *Proceedings of the IEEE International Conference on Neural Networks*, pp. 1183–1188.

Bianchi, L. (1922). *The mechanism of the brain and the function of the frontal lobes*. Edinburgh: Livingstone.

Braver, T. S. (1997). Mechanisms of cognitive control: A neurocomputational model. Ph.D. diss., Carnegie Mellon University, Pittsburgh.

Braver, T. S., Barch, D. M., and Cohen, J. D. (1999a). Cognition and control in schizophrenia: A computational model of dopamine and prefrontal function. *Biological Psychiatry, 46,* 312–328.

Braver, T. S., Barch, D. M., and Cohen, J. D. (1999b). Mechanisms of cognitive control: Active memory, inhibition and the prefrontal cortex. Technical Report PDP. CNS. 99.1. Pittsburgh: Carnegie Mellon University.

Braver, T. S., and Cohen, J. D. (1999). Dopamine, cognitive control, and schizophrenia: the gating model. *Progress in Brain Research, 121,* 327–349.

Braver, T. S., Cohen, J. D., and McClelland, J. L. (1997). An integrated computational model of dopamine function in reinforcement learning and working memory. *Society for Neuroscience Abstracts,* 23, 775.

Braver, T. S., Cohen, J. D., and Servan-Schreiber, D. (1995a). A computational model of prefrontal cortex function. In D. S. Touretzky, G. Tesauro, and T. K. Leen (Eds.), *Advances in neural information-processing systems,* vol. 7, pp. 141–148. Cambridge, MA: MIT Press.

Braver, T. S., Cohen, J. D., and Servan-Schreiber, D. (1995b). Neural network simulations of schizophrenic performance in a variant of the CPT-AX: A predicted double dissociation. *Schizophrenia Research, 15,* 110.

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,* 929–931.

Calabresi, P., Pisani, A., Centonze, D., and Bernardi, G. (1997). Synaptic plasticity and physiological interactions between dopamine and glutamate in the striatum. *Neuroscience and Biobehavioral Reviews, 21,* 519–523.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., and Cohen, J. D. (1998). The role of the anterior cingulate cortex in error detection and the on line monitoring of performance. *Science, 280,* 747–749.

Chiodo, L., and Berger, T. (1986). Interactions between dopamine and amino-acid induced excitation and inhibition in the striatum. *Brain Research, 375,* 198–203.

Cohen, J. D., Barch, D. M., Carter, C. S., and Servan-Schreiber, D. (1999). Schizophrenic deficits in the processing of context: Converging evidence from three theoretically motivated cognitive tasks. *Journal of Abnormal Psychology, 108,* 120–133.

Cohen, J. D., Braver, T. S., and O'Reilly, R. (1996). A computational approach to prefrontal cortex, cognitive control, and schizophrenia: Recent developments and current challenges. *Philosophical Transactions of the Royal Society of London, B351,* 1515–1527.

Cohen, J. D., Dunbar, K., and McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97,* 332–361.

Cohen, J. D., and Huston, T. A. (1994). Progress in the use of interactive models for understanding attention and performance. In C. Umiltà and M. Moscovitch (Eds.), *Attention and Performance XV,* pp. 1–19. Cambridge, MA: MIT Press.

Cohen, J. D., Perstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., and Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature, 386,* 604–608.

Cohen, J. D., and Servan-Schreiber, D. (1992). Context, cortex and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review, 99*, 45–77.

Cohen, J. D., and Servan-Schreiber, D. (1993). A theory of dopamine function and cognitive deficits in schizophrenia. *Schizophrenia Bulletin, 19*, 85–104.

Constantinidis, C., and Steinmetz, M. A. (1996). Neuronal activity in posterior parietal area 7a during the delay periods of a spatial memory task. *Journal of Neurophysiology, 76*, 1352–1355.

Courtney, S. M., Ungerleider, L. G., Keil, K., and Haxby, J. V. (1997). Transient and sustained activity in a distributed neural system for human working memory. *Nature, 386*, 608–612.

Damasio, A. R. (1985). The frontal lobes. In K. M. Heilman and E. Valenstein (Eds.), *Clinical neuropsychology*, pp. 339–375. New York: Oxford University Press.

Davidson, R. J., and Sutton, S. K. (1995). Affective neuroscience: The emergence of a discipline. *Current Opinion in Neurobiology, 5*, 215–225.

Dehaene, S., and Changeux, J. P. (1989). A simple model of prefrontal cortex function in delayed-response tasks. *Journal of Cognitive Neuroscience, 1*, 244–261.

Dehaene, S., and Changeux, J. P. (1992). The Wisconsin Card-Sorting Test: Theoretical analysis and modeling in a neuronal network. *Cerebral Cortex, 1*, 62–79.

Engle, R. W., Kane, M., and Tuholski, S. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In A. Miyake and P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control*, pp. 102–134. New York: Cambridge University Press.

Gold, J. (1992). Prefrontal function and schizophrenic symptoms. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 5*, 253–261.

Goldman-Rakic, P. S., and Selemon, L. D. (1997). Functional and anatomical aspects of prefrontal pathology in schizophrenia. *Schizophrenia Bulletin, 23*, 437–458.

Grace, A. A. (1991). Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: A hypothesis of the etiology of schizophrenia. *Neuroscience, 41*, 1–24.

Guigon, E., Dorizzi, B., Burnod, Y., and Schultz, w. (1991). Neural correlates of learning in the prefrontal cortex of the monkey: A predictive model. *Cerebral Cortex, 5*, 135–147.

Gullapalli, V. (1990). A stochastic reinforcement learning algorithm for learning real-valued functions. *Neural Networks, 3*, 671–692.

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. *Journal of Neuroscience, 17*, 3334–3342.

Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*, 443–453.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences, U.S.A., 79*, 2554–2558.

Kimberg, D. Y., D'Esposito, M., and Farah, M. J. (1997). Effects of bromocriptine on human subjects depend on working memory capacity. *Neuroreport, 8*, 381–385.

Law-Tho, D., Deuce, J. M., and Crepel, F. (1995). Dopamine favours the emergence of long-term depression versus long-term potentiation in slices of rat prefrontal cortex. *Neuroscience Letters, 188*, 125–128.

Levine, D. S., and Prueitt, P. S. (1989). Modeling some effects of frontal lobe damage: Novelty and perseveration. *Neural Networks, 2,* 103–116.

Lewis, D. A., Hayes, T. L., Lund, J. S., and Oeth, K. M. (1992). Dopamine and the neural circuitry of primate prefrontal cortex: Implications for schizophrenia research. *Neuropsychopharmacology, 6,* 127–134.

Luciana, M., Collins, P. F., and Depue, R. A. (1995). DA and 5-HT influences on spatial working memory functions of prefrontal cortex. Paper presented at the Second Annual Meeting of the Cognitive Neuroscience Society, San Francisco, April.

Luciana, M., Depue, R. A., Arbisi, P., and Leon, A. (1992). Facilitation of working memory in humans by a D2 dopamine receptor agonist. *Journal of Cognitive Neuroscience, 4,* 58–68.

Luria, A. R. (1969). Frontal lobe syndromes. In P. J. Vinken and G. W. Bruyn (Eds.), *Handbook of clinical neurology,* vol. 2, pp. 725–757. New York: Elsevier.

Malmo, H. P. (1974). On frontal lobe functions: Psychiatric patient controls. *Cortex, 10,* 231–237.

McClelland, J. L. (1993). Toward a theory of information processing in graded, random, and interactive networks. In D. E. Meyer and S. Kornblum (Eds.), *Attention and Performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience,* pp. 655–688. Cambridge, MA: MIT Press.

Miller, E. K., Erickson, C. A., and Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience, 16,* 5154–5167.

Milner, B. (1963). Effects of different brain lesions on card sorting. *Archives of Neurology, 9,* 90–100.

Montague, P. R., Dayan, P., and Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience, 16,* 1936–1947.

Montague, P. R., and Sejnowski, T. J. (1994). The predictive brain: Temporal coincidence and temporal order in synaptic learning mechanisms. *Learning and Memory, 1,* 1–33.

Moody, S. L., Wise, S. P., di Pellegrino, G., and Zipser, D. (1998). A model that accounts for activity in primate frontal cortex during a delayed matching-to-sample task. *Journal of Neuroscience, 18,* 399–410.

Moore, R. Y., and Bloom, F. E. (1978). Central catecholamine neuron systems: Anatomy and physiology of the dopamine systems. *Annual Review of Neuroscience, 1,* 129–169.

Mozer, M. C. (1993). Neural net architectures for temporal sequence processing. In A. Weigend and N. Gershenfeld (Eds.), *Predicting the future and understanding the past,* pp. 243–264. Redwood City, CA: Addison-Wesley.

Norman, D. A., and Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In T. Shallice and D. A. Norman (Eds.), *Consciousness and self-regulation,* pp. 1–18. New York: Plenum Press.

Nuechterlein, K. H., and Dawson, M. E. (1984). Information processing and attentional functioning in the developmental course of schizophrenia disorders. *Schizophrenia Bulletin, 10,* 160–203.

O'Reilly, R. C. (1996). The LEABRA model of neural interactions and learning in the neocortex. Ph.D. diss., Carnegie Mellon University, Pittsburgh.

Owen, A. M., Roberts, A. C., Polkey, C. E., Sahakian, B. J., and Robbins, T. W. (1991). Extra-dimensional versus intra-dimensional set shifting performance following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia, 29*, 993–1006.

Penit-Soria, J., Audinat, E., and Crepel, F. (1987). Excitation of rat prefrontal cortical neurons by dopamine: An in vitro electrophysiological study. *Brain Research, 425*, 263–274.

Robbins, T., Roberts, A., Owen, A., Sahakian, B., Everitt, B., Wilkinson, L., Muir, J., DeSalvia, M., and Tovee, M. (1994). Monoaminergic-dependent cognitive functions of the prefrontal cortex in monkey and man. In A.-M. Thierry, J. Glowinski, P. Goldman-Rakic, and Y. Christen (Eds.), *Motor and cognitive functions of the prefrontal cortex*, pp. 93–111. Berlin: Springer.

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome, E. D., and Beck, L. H. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology, 20*, 343–350.

Sawaguchi, T., and Goldman-Rakic, P. S. (1991). D1 dopamine receptors in prefrontal cortex: Involvement in working memory. *Science, 251*, 947–950.

Sawaguchi, T., and Goldman-Rakic, P. (1994). The role of D1 dopamine receptors in working memory: Local injections of dopamine antagonists into the prefrontal cortex of rhesus monkeys performing an oculomotor delayed-response task. *Journal of Neurophysiology, 63*, 1401–1412.

Sawaguchi, T., Matsumura, M., and Kubota, K. (1990). Effects of dopamine antagonists on neuronal activity related to a delayed response task in monkey prefrontal cortex. *Journal of Neurophysiology, 63*, 1401–1410.

Schultz, W. (1992). Activity of dopamine neurons in the behaving primate. *Seminars in Neurosciences, 4*, 129–138.

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. *Journal of Neuroscience, 13*, 900–913.

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*, 1593–1599.

Servan-Schreiber, D., Bruno, R., Carter, C. S., and Cohen, J. D. (1998). Dopamine and the mechanisms of cognition: 1. A neural network model predicting dopamine effects on selective attention. *Biological Psychiatry, 43*, 713–722.

Servan-Schreiber, D., Cohen, J. D., and Steingard, S. (1996). Schizophrenic deficits in the processing of context: A test of a theoretical model. *Archives of General Psychiatry, 53*, 1105–1113.

Servan-Schreiber, D., Printz, H., and Cohen, J. D. (1990). A network model of catecholamine effects: Gain, signal-to-noise ratio, and behavior. *Science, 249*, 892–895.

Sesack, S. R., Snyder, C. L., and Lewis, D. A. (1995). Axon terminals immunolabeled for dopamine or tyrosine hydroxylase synapse on GABA-immunoreactive dendrites in rat and monkey cortex. *Journal of Comparative Neurology, 363*, 1735–1780.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London, 298*, 199–209.

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge: Cambridge University Press.

Stuss, D. T., and Benson, D. F. (1986). *The frontal lobes*. New York: Raven Press.

Sutton, R. (1988). Learning to predict by the method of temporal difference. *Machine Learning, 3*, 9–44.