Annu. Rev. Neurosci. 2001. 24:167–202

AN INTEGRATIVE THEORY OF PREFRONTAL CORTEX FUNCTION

Earl K. Miller
Center for Learning and Memory, RIKEN-MIT Neuroscience Research Center and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139; e-mail: ekm@ai.mit.edu

Jonathan D. Cohen
Center for the Study of Brain, Mind, and Behavior and Department of Psychology, Princeton University, Princeton, New Jersey 08544; e-mail: jdc@princeton.edu

Key Words frontal lobes, cognition, executive control, working memory, attention

■ **Abstract** The prefrontal cortex has long been suspected to play an important role in cognitive control, in the ability to orchestrate thought and action in accordance with internal goals. Its neural basis, however, has remained a mystery. Here, we propose that cognitive control stems from the active maintenance of patterns of activity in the prefrontal cortex that represent goals and the means to achieve them. They provide bias signals to other brain structures whose net effect is to guide the flow of activity along neural pathways that establish the proper mappings between inputs, internal states, and outputs needed to perform a given task. We review neurophysiological, neurobiological, neuroimaging, and computational studies that support this theory and discuss its implications as well as further issues to be addressed.

INTRODUCTION

One of the fundamental mysteries of neuroscience is how coordinated, purposeful behavior arises from the distributed activity of billions of neurons in the brain. Simple behaviors can rely on relatively straightforward interactions between the brain's input and output systems. Animals with fewer than a hundred thousand neurons (in the human brain there are 100 billion or more neurons) can approach food and avoid predators. For animals with larger brains, behavior is more flexible. But flexibility carries a cost: Although our elaborate sensory and motor systems provide detailed information about the external world and make available a large repertoire of actions, this introduces greater potential for interference and confusion. The richer information we have about the world and the greater number of options for behavior require appropriate attentional, decision-making, and coordinative functions, lest uncertainty prevail. To deal with this multitude of

0147-006X/01/0301-0167$14.00

possibilities and to curtail confusion, we have evolved mechanisms that coordinate lower-level sensory and motor processes along a common theme, an internal goal. This ability for cognitive control no doubt involves neural circuitry that extends over much of the brain, but it is commonly held that the prefrontal cortex (PFC) is particularly important.

The PFC is the neocortical region that is most elaborated in primates, animals known for their diverse and flexible behavioral repertoire. It is well positioned to coordinate a wide range of neural processes: The PFC is a collection of interconnected neocortical areas that sends and receives projections from virtually all cortical sensory systems, motor systems, and many subcortical structures (Figure 1). Neurophysiological studies in nonhuman primates have begun to define many of the detailed properties of PFC, and human neuropsychology and neuroimaging studies have begun to provide a broad view of the task conditions under which it is engaged. However, an understanding of the mechanisms by which the PFC executes control has remained elusive. The aim of this article is to describe a theory of PFC function that integrates these diverse findings, and more precisely defines its role in cognitive control.

The Role of the PFC in Top-Down Control of Behavior

The PFC is not critical for performing simple, automatic behaviors, such as our tendency to automatically orient to an unexpected sound or movement. These behaviors can be innate or they can develop gradually with experience as learning mechanisms potentiate existing pathways or form new ones. These "hardwired" pathways are advantageous because they allow highly familiar behaviors to be executed quickly and automatically (i.e. without demanding attention). However, these behaviors are inflexible, stereotyped reactions elicited by just the right stimulus. They do not generalize well to novel situations, and they take extensive time and experience to develop. These sorts of automatic behaviors can be thought of as relying primarily on "bottom-up" processing; that is, they are determined largely by the nature of the sensory stimuli and well-established neural pathways that connect these with corresponding responses.

By contrast, the PFC is important when "top-down" processing is needed; that is, when behavior must be guided by internal states or intentions. The PFC is critical in situations when the mappings between sensory inputs, thoughts, and actions either are weakly established relative to other existing ones or are rapidly changing. This is when we need to use the "rules of the game," internal representations of goals and the means to achieve them. Several investigators have argued that this is a cardinal function of the PFC (Cohen & Servan-Schreiber 1992, Passingham 1993, Grafman 1994, Wise et al 1996, Miller 1999). Two classic tasks illustrate this point: the Stroop task and the Wisconsin card sort task (WCST).

In the Stroop task (Stroop 1935, MacLeod 1991), subjects either read words or name the color in which they are written. To perform this task, subjects must selectively attend to one attribute. This is especially so when naming the color

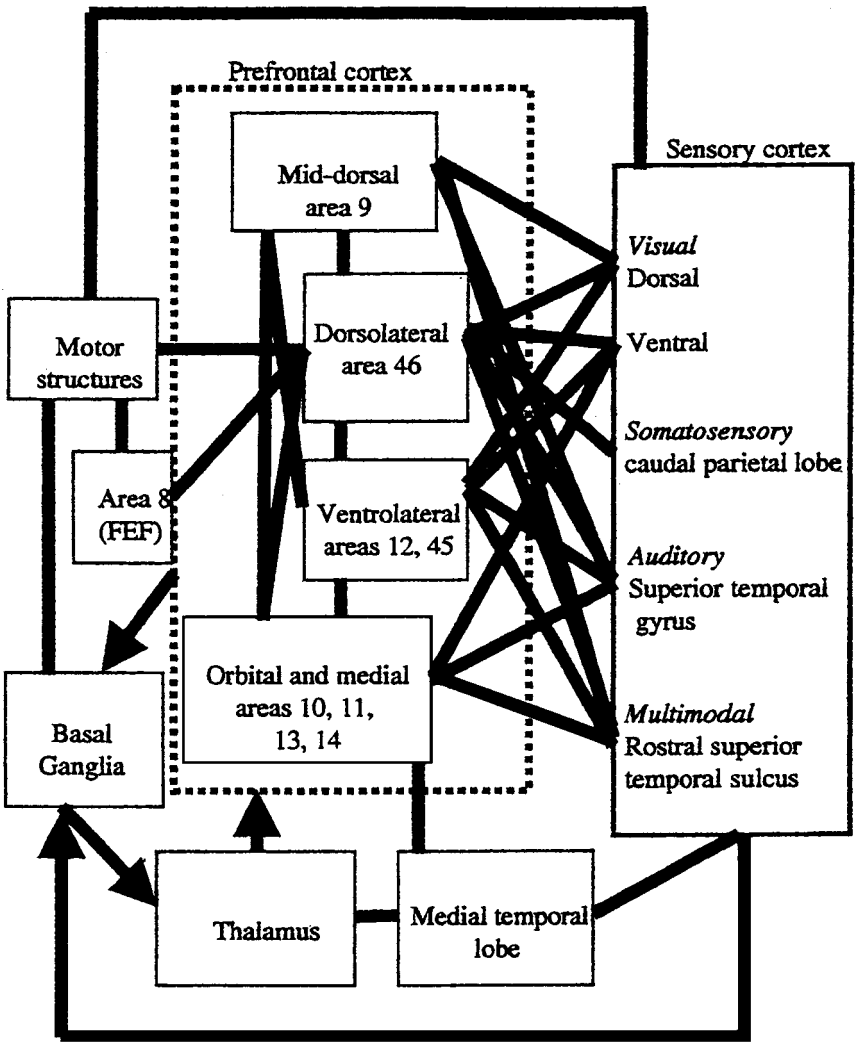


Figure 1 Schematic diagram of some of the extrinsic and intrinsic connections of the prefrontal cortex. The partial convergence of inputs from many brain systems and internal connections of the prefrontal cortex (PFC) may allow it to play a central role in the synthesis of diverse information needed for complex behavior. Most connections are reciprocal; the exceptions are indicated by arrows. The frontal eye field (FEF) has variously been considered either adjacent to, or part of, the PFC. Here, we compromise by depicting it as adjacent to, yet touching, the PFC.

of a conflict stimulus (e.g. the word GREEN displayed in red), because there is a strong prepotent tendency to read the word ("green"), which competes with the response to the color ("red"). This illustrates one of the most fundamental aspects of cognitive control and goal-directed behavior: the ability to select a weaker, task-relevant response (or source of information) in the face of competition from an otherwise stronger, but task-irrelevant one. Patients with frontal impairment have difficulty with this task (e.g. Perrett 1974, Cohen & Servan-Schreiber 1992, Vendrell et al 1995), especially when the instructions vary frequently (Dunbar & Sussman 1995, Cohen et al 1999), which suggests that they have difficulty adhering to the goal of the task or its rules in the face of a competing stronger (i.e. more salient or habitual) response.

Similar findings are evident in the WCST. Subjects are instructed to sort cards according to the shape, color, or number of symbols appearing on them and the sorting rule varies periodically. Thus, any given card can be associated with several possible actions, no single stimulus-response mapping will work, and the correct one changes and is dictated by whichever rule is currently in effect. Humans with PFC damage show stereotyped deficits in the WCST. They are able to acquire the initial mapping without much difficulty but are unable to adapt their behavior when the rule varies (Milner 1963). Monkeys with PFC lesions are impaired in an analog of this task (Dias et al 1996b, 1997) and in others when they must switch between different rules (Rossi et al 1999).

The Stroop task and WCST are variously described as tapping the cognitive functions of either selective attention, behavioral inhibition, working memory, or rule-based or goal-directed behavior. In this article, we argue that all these functions depend on the representation of goals and rules in the form of patterns of activity in the PFC, which configure processing in other parts of the brain in accordance with current task demands. These top-down signals favor weak (but task-relevant) stimulus-response mappings when they are in competition with more habitual, stronger ones (such as in the Stroop task), especially when flexibility is needed (such as in the WCST). We believe that this can account for the wide range of other tasks found to be sensitive to PFC damage, such as A-not-B (Piaget 1954, Diamond & Goldman-Rakic 1989), Tower of London (Shallice 1982, 1988; Owen et al 1990), and others (Duncan 1986, Duncan et al 1996), Stuss & Benson 1986).

We build on the fundamental principle that processing in the brain is competitive: Different pathways, carrying different sources of information, compete for expression in behavior, and the winners are those with the strongest sources of support. Desimone & Duncan (1995) have proposed a model that clearly articulates such a view with regard to visual attention. These authors assume that visual cortical neurons processing different aspects of a scene compete with each other via mutually inhibitory interactions. The neurons that "win" the competition and remain active reach higher levels of activity than those with which they share inhibitory interactions. Voluntary shifts of attention result from the influence of excitatory top-down signals representing the to-be-attended features of the scene. These bias the competition among neurons representing the scene, increasing the activity of

neurons representing the to-be-attended features and, by virtue of mutual inhibition, suppressing activity of neurons processing other features. Desimone & Duncan suggest that the PFC is an important source of such top-down biasing. However, they left unspecified the mechanisms by which this occurs. That is the focus of this article.

We begin by outlining a theory that extends the notion of biased competition and proposes that it provides a fundamental mechanism by which the PFC exerts control over a wide range of processes in the service of goal-directed behavior. We describe the minimal set of functional properties that such a system must exhibit if it can serve as a mechanism of cognitive control. We then review the existing literature that provides support for this set of properties, followed by a discussion of recent computational modeling efforts that illustrate how a system with these properties can support elementary forms of control. Finally, we consider unresolved issues that provide a challenge for future empirical and theoretical research.

Overview of the Theory

We assume that the PFC serves a specific function in cognitive control: the active maintenance of patterns of activity that represent goals and the means to achieve them. They provide bias signals throughout much of the rest of the brain, affecting not only visual processes but also other sensory modalities, as well as systems responsible for response execution, memory retrieval, emotional evaluation, etc. The aggregate effect of these bias signals is to guide the flow of neural activity along pathways that establish the proper mappings between inputs, internal states, and outputs needed to perform a given task. This is especially important whenever stimuli are ambiguous (i.e. they activate more than one input representation), or when multiple responses are possible and the task-appropriate response must compete with stronger alternatives. From this perspective, the constellation of PFC biases—which resolves competition, guides activity along appropriate pathways, and establishes the mappings needed to perform the task—can be viewed as the neural implementation of attentional templates, rules, or goals, depending on the target of their biasing influence.

To help understand how this might work, consider the schematic shown in Figure 2. Processing units are shown that correspond to cues (C1, C2, C3). They can be thought of as neural representations of sensory events, internal states (e.g. stored memories, emotions, etc), or combinations of these. Also shown are units corresponding to the motor circuits mediating two responses (R1 and R2), as well as intervening or "hidden" units that define processing pathways between cue and response units. We have set up the type of situation for which the PFC is thought to be important. Namely, one cue (C1) can lead to either of two responses (R1 or R2) depending on the situation (C2 or C3), and appropriate behavior depends on establishing the correct mapping from C1 to R1 or R2. For example, imagine you are standing at the corner of a street (cue C1). Your natural reaction is to look left before crossing (R1), and this is the correct thing to do in most

of the world (C2). However, if you are in England (C3), you should look right (R2). This is a classic example of a circumstance requiring cognitive control, which we assume depends on the PFC. How does the PFC mediate the correct behavior?

We assume that cues in the environment activate internal representations within the PFC that can select the appropriate action. This is important when the course of action is uncertain, and especially if one of the alternatives is stronger (i.e. more habitual or more salient) but produces the incorrect behavior. Thus, standing at the corner (C1), your "automatic" response would be to look left (R1). However, other cues in the environment "remind" you that you are in England (C3). That is, the cues activate the corresponding PFC representation, which includes information about the appropriate action. This produces excitatory bias signals that guide neural activity along the pathway leading you to look right (e.g. C1 $\rightarrow \cdots \rightarrow$ R2). Note that activation of this PFC representation is necessary for you to perform the correct behavior. That is, you had to keep "in mind" the knowledge that you were in England. You might even be able to cross a few streets correctly while keeping this knowledge in mind, that is, while activity of the appropriate representation is maintained in the PFC. However, if this activity subsides—that is, if you "forget" you are in England—you are likely to revert to the more habitual response and look left. Repeated selection can strengthen the pathway from C1 to R2 and allow it to become independent of the PFC. As this happens, the behavior becomes more automatic, so you can look right without having to keep in mind that you are in England. An important question is how the PFC develops the representations needed to produce the contextually appropriate response.

In an unfamiliar situation you may try various behaviors to achieve a desired goal, perhaps starting with some that have been useful in a similar circumstance (looking to the left for oncoming traffic) and, if these fail, trying others until you meet with success (e.g. by looking right). We assume that each of these is associated with some pattern of activity within the PFC (as in Figure 2). When a behavior meets with success, reinforcement signals augment the corresponding pattern of activity by strengthening connections between the PFC neurons activated by that behavior. This process also strengthens connections between these neurons and those whose activity represents the situation in which the behavior was useful, establishing an association between these circumstances and the PFC pattern that supports the correct behavior. With time (and repeated iterations of this process), the PFC representation can be further elaborated as subtler combinations of events and contingencies between them and the requisite actions are learned. As is discussed below, brainstem neuromodulatory systems may provide the relevant reinforcement signals, allowing the system to "bootstrap" in this way.

Obviously, many details need to be added before we fully understand the complexity of cognitive control. But we believe that this general notion can explain many of the posited functions of the PFC. The biasing influence of PFC feedback signals on sensory systems may mediate its role in directing attention (Stuss & Benson 1986; Knight 1984, 1997; Banich et al 2000), signals to the motor system

may be responsible for response selection and inhibitory control (Fuster 1980, Diamond 1988), and signals to intermediate systems may support short-term (or working) memory (Goldman-Rakic 1987) and guide retrieval from long-term memory (Schachter 1997, Janowsky et al 1989, Gershberg & Shimamura 1995). Without the PFC, the most frequently used (and thus best established) neural pathways would predominate or, where these don't exist, behavior would be haphazard. Such impulsive, inappropriate, or disorganized behavior is a hallmark of PFC dysfunction in humans (e.g. Bianchi 1922, Duncan 1986, Luria 1969, Lhermitte 1983, Shallice & Burgess 1996, Stuss & Benson 1986).

Minimal Requirements for a Mechanism of Top-Down Control

There are several critical features of our theory. First, the PFC must provide a source of activity that can exert the required pattern of biasing signals to other structures. We can thus think of PFC function as "active memory in the service of control." It follows, therefore, that the PFC must maintain its activity robustly against distractions until a goal is achieved, yet also be flexible enough to update its representations when needed. It must also house the appropriate representations, those that can select the neural pathways needed for the task. Insofar as primates are capable of tasks that involve diverse combinations of stimuli, internal states, and responses, representations in the PFC must have access to and be able to influence a similarly wide range of information in other brain regions. That is, PFC representations must have a high capacity for multimodality and integration. Finally, as we can acquire new goals and means, the PFC must also exhibit a high degree of plasticity. Of course, it must be possible to exhibit all these properties without the need to invoke some other mechanism of control to explain them, lest our theory be subject to perennial concerns of a hidden "homunculus."

The rapidly accumulating body of findings regarding the PFC suggests that it meets these requirements. Fuster (1971, 1973, 1995), Goldman-Rakic (1987, 1996), and others have extensively explored the ability of PFC neurons to maintain task-relevant information. Miller et al (1996) have shown that this is robust to interference from distraction. Fuster has long advocated the role of the PFC in integrating diverse information (Fuster 1985, 1995). The earliest descriptions of the effects of frontal lobe damage suggested its role in attention and the control of behavior (Ferrier 1876, Bianchi 1922), and investigators since have interpreted the pattern of deficits following PFC damage as a loss of the ability to acquire and use behavior-guiding rules (Shallice 1982, Duncan 1986, Passingham 1993, Grafman 1994, Wise et al 1996). Recent empirical studies have begun to identify neural correlates of plasticity in the PFC (Asaad et al 1998, Bichot et al 1996, Schultz & Dickinson 2000), and recent computational studies suggest how these may operate as mechanisms for self-organization (Braver & Cohen 2000, Egelman et al 1998). Our purpose in this article is to bring these various observations and arguments together, and to illustrate that a reasonably coherent, and mechanistically explicit,

theory of PFC function is beginning to emerge. The view presented here draws on previous work that has begun to outline such a theory (e.g. Cohen & Servan-Schreiber 1992; Cohen et al 1996; O'Reilly et al 1999; Miller 1999, 2000). In the sections that follow, we review neurobiological, neuropsychological, and neuroimaging findings that support this theory, and computational modeling studies that have begun to make explicit the processing mechanisms involved.

PROPERTIES OF THE PFC

Convergence of Diverse Information

One of the critical features for a system of cognitive control is the requirement that it have access to diverse information about both the internal state of the system and the external state of the world. The PFC is anatomically well situated to meet this requirement. The cytoarchitectonic areas that comprise the monkey PFC are often grouped into regional subdivisions, the orbital and medial, the lateral, and the mid-dorsal (see Figure 1). Collectively, these areas have interconnections with virtually all sensory systems, with cortical and subcortical motor system structures, and with limbic and midbrain structures involved in affect, memory, and reward. The subdivisions have partly unique, but overlapping, patterns of connections with the rest of the brain, which suggests some regional specialization. However, as in much of the neocortex, many PFC connections are local; there are extensive connections between different PFC areas that are likely to support an intermixing of disparate information. Such intermixing provides a basis for synthesizing results from, and coordinating the regulation of, a wide variety of brain processes, as would be required of a brain area responsible for the orchestration of complex behavior.

Sensory Inputs The lateral and mid-dorsal PFC is more closely associated with sensory neocortex than is the ventromedial PFC (see Figure 1). It receives visual, somatosensory, and auditory information from the occipital, temporal, and parietal cortices (Barbas & Pandya 1989, 1991; Goldman-Rakic & Schwartz 1982; Pandya & Barnes 1987; Pandya & Yeterian 1990; Petrides & Pandya 1984, 1999; Seltzer & Pandya 1989). Many PFC areas receive converging inputs from at least two sensory modalities (Chavis & Pandya 1976; Jones & Powell 1970). For example, the dorsolateral (DL) (areas 8, 9, and 46) and ventrolateral (12 and 45) PFC both receive projections from visual, auditory, and somatosensory cortex. Furthermore, the PFC is connected with other cortical regions that are themselves sites of multimodal convergence. Many PFC areas (9, 12, 46, and 45) receive inputs from the rostral superior temporal sulcus, which has neurons with bimodal or trimodal (visual, auditory, and somatosensory) responses (Bruce et al 1981, Pandya & Barnes 1987). The arcuate sulcus region (areas 8 and 45) and area 12 seem to be particularly multimodal. They contain zones that receive overlapping inputs from three sensory modalities (Pandya & Barnes 1987). In all these cases,

the PFC is directly connected with secondary or "association" but not primary sensory cortex.

Motor Outputs The dorsal PFC, particularly DL area 46, has preferential connections with motor system structures that may be central to how the PFC exerts control over behavior. The DL area 46 is interconnected (*a*) with motor areas in the medial frontal lobe such as the supplementary motor area, the pre–supplementary motor area, and the rostral cingulate, (*b*) with the premotor cortex on the lateral frontal lobe, and (*c*) with cerebellum and superior colliculus (Bates & Goldman-Rakic 1993, Goldman & Nauta 1976, Lu et al 1994, Schmahmann & Pandya 1997). The DL area 46 also sends projections to area 8, which contains the frontal eye fields, a region important for voluntary shifts of gaze. There are no direct connections between the PFC and primary motor cortex, but they are extensive with premotor areas that, in turn, send projections to primary motor cortex and the spinal cord. Also important are the dense interconnections between the PFC and basal ganglia (Alexander et al 1986), a structure that is likely to be crucial for automating behavior. The basal ganglia receives inputs from much of the cerebral cortex, but its major output (via the thalamus) is frontal cortex (see Figure 1).

Limbic Connections The orbital and medial PFC are closely associated with medial temporal limbic structures critical for long-term memory and the processing of internal states, such as affect and motivation. This includes direct and indirect (via the medial dorsal thalamus) connections with the hippocampus and associated neocortex, the amygdala, and the hypothalamus (Amaral & Price 1984, Barbas & De Olmos 1990, Barbas & Pandya 1989, Goldman-Rakic et al 1984, Porrino et al 1981, Van Hoesen et al 1972). Other PFC regions have access to these systems both through connections with the orbital and medial PFC and through other intervening structures.

Intrinsic Connections Most PFC regions are interconnected with most other PFC regions. There are not only interconnections between all three major subdivisions (ventromedial, lateral, and mid-dorsal) but also between their constituent areas (Barbas & Pandya 1991, Pandya & Barnes 1987). The lateral PFC is particularly well connected. Ventrolateral areas 12 and 45 are interconnected with DL areas 46 and 8, with dorsal area 9, as well as with ventromedial areas 11 and 13. Intrinsic connections within the PFC allow information from regional afferents and processes to be distributed to other parts of the PFC. Thus, the PFC provides a venue by which information from wide-ranging brain systems can interact through relatively local circuitry.

Convergence and Plasticity

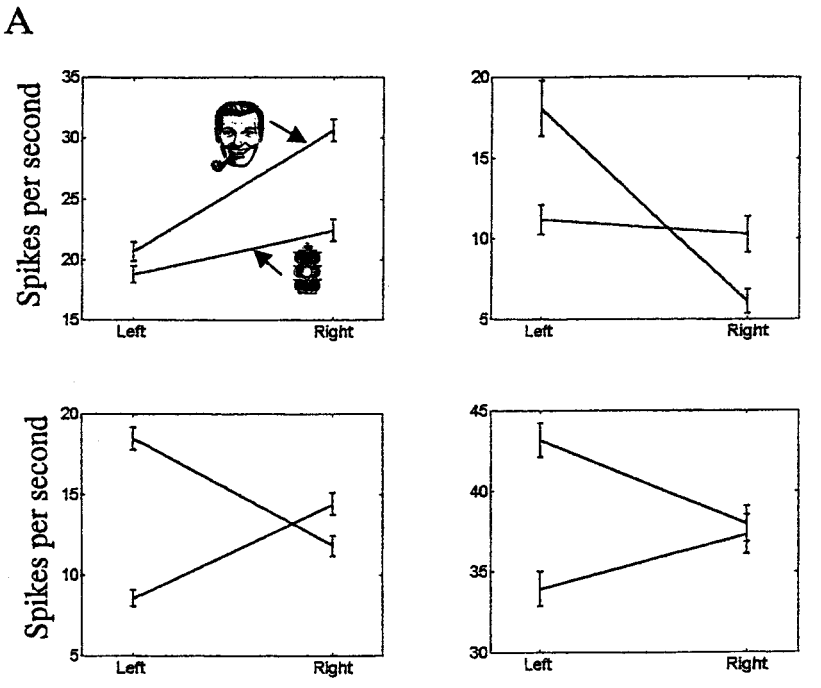
Given that goal-directed behavior depends on our ability to piece together relationships between a wide range of external and internal information, it stands to

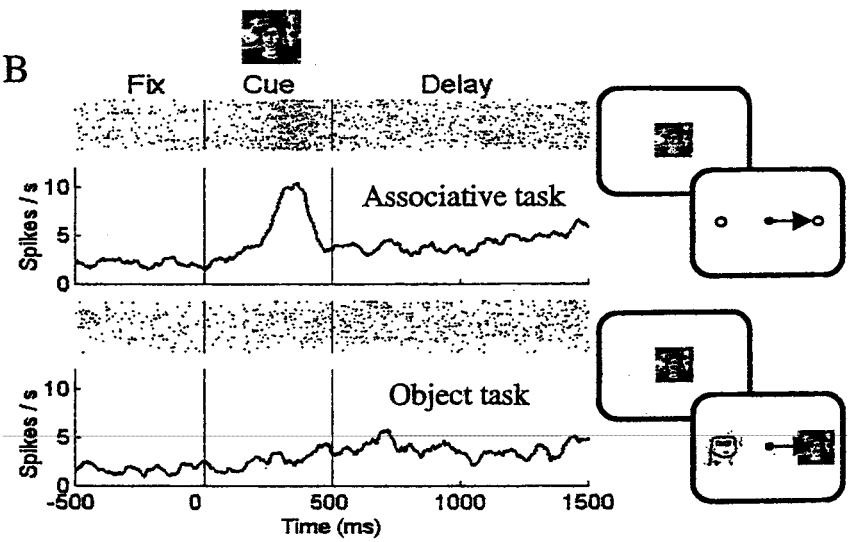
reason that top-down control must come from PFC representations that reflect a wide range of learned associations. There is mounting neurophysiological evidence that this is the case. Asaad et al (1998) trained monkeys to associate, on different blocks of trials, each of two cue objects with a saccade to the right or a saccade to the left. They found relatively few lateral PF neurons whose activity simply reflected a cue or response. Instead, the modal group of neurons (44% of the population) showed activity that reflected the current association between a visual cue and a directional saccade it instructed. For example, a given cell might only be strongly activated when object "A" instructed "saccade left" and not when object "B" instructed the same saccade or when object "A" instructed another saccade (Figure 3*A*). Lateral PFC neurons can also convey the degree of association between a cue and a response (Quintana & Fuster 1992).

Other studies indicate that PFC neurons acquire selectivity for features to which they are initially insensitive but are behaviorally relevant. For example, Bichot et al (1996) observed that neurons in the frontal eye fields (in the bow of the arcuate sulcus)—ordinarily not selective to the form and color of stimuli—became so as the animal learned eye movements that were contingent on these features. Similarly, Watanabe (1990, 1992) has trained monkeys to recognize that certain visual and auditory stimuli signaled whether or not, on different trials, a reward (a drop of juice) would be delivered. He found that neurons in lateral PFC (around the arcuate sulcus and posterior end of the principal sulcus) came to reflect specific cue-reward associations. For example, a given neuron could show strong activation to one of the two auditory (and none of the visual) cues, but only when it signaled reward. Other neurons were bimodal, activated by both visual and auditory cues but also strongly modulated by their reward status.

More complicated behaviors depend not on simple contingencies between cues and responses or rewards but on general principles or rules that may involve more-complex mapping. PFC activity also seems to represent this information. Barone & Joseph (1989) observed cells near the arcuate sulcus that were responsive to specific light stimuli, but only when they occurred at a particular point in a particular

Figure 3 (*A*) Shown is the activity of four single prefrontal (PF) neurons when each of two objects, on different trials, instructed either a saccade to the right or a saccade to the left. The lines connect the average values obtained when a given object cued one or the other saccade. The error bars show the standard error of the mean. Note that in each case, the neuron's activity depends on both the cue object and the saccade direction and that the tuning is nonlinear or conjunctive. That is, the level of activity to a given combination of object and saccade cannot be predicted from the neuron's response to the other combinations. [Adapted from Asaad et al (1998).] (*B*) A PF neuron whose neural response to a cue object was highly dependent on task context. The bottom half shows an example of a single PF neuron's response to the same cue object during an object task (delayed matching to sample) and during an associative task (conditional visual motor). Note that the neuron is responsive to the cue during one task but not during the other, even though sensory stimulation is identical across the tasks. [Adapted from Asaad et al (2000).]


A
Spikes per second
35
30
25
20
15
Left
Right
20
15
10
5
Left
Right
Spikes per second
20
15
10
5
Left
Right
45
40
35
30
Left
Right


B
Fix
Cue
Delay
Associative task
Object task
Spikes / s
10
5
0
-500
0
500
1000
1500
Time (ms)

sequence that the monkey had to imitate. White & Wise (1999) trained a monkey to orient to a visual target according to two different rules. One of four cue patterns briefly appeared at one of four locations. The cue indicated where a target would eventually appear. It did so by one of two rules, a spatial rule (the cue appeared at the same location that the target would appear) or an associative rule (the identity of the cue instructed the location, e.g. cue A indicated the right, cue B the left, etc). They found that up to half of lateral PFC neurons showed activity that varied with the rule. Another example was provided by Asaad et al (2000), who trained monkeys to alternate between tasks that employed the same cues and responses but three different rules: matching (delayed matching to sample), associative (conditional visuomotor), and spatial (spatial delayed response). Over half of lateral PFC neurons were rule dependent. Neural responses to a given cue or forthcoming saccade often depended on which rule was current in the task (Figure 3*B*). Plus, the baseline activity of many neurons (54%) varied with the rule. Hoshi et al (1998) have also observed PFC neurons that were modulated by whether the monkey was using a shape-matching or location-matching rule. Recently, Wallis et al (2000) have shown that lateral and orbitofrontal PFC neurons reflect whether the monkey is currently using a "matching" or "nonmatching" rule to select a test object.

Studies of monkeys and humans with PFC damage also suggest that the PFC is critical for learning rules. For example, Petrides found that following PFC damage, patients could no longer learn arbitrary associations between visual patterns and hand gestures (Petrides 1985, 1990). In monkeys, damage to ventrolateral area 12 or to the arcuate sulcus region also impairs the ability to learn arbitrary cue-response associations (Halsband & Passingham 1985, Murray et al 2000, Petrides 1982, 1985). Learning of visual stimulus-response conditional associations is also impaired by damage to PFC inputs from the temporal cortex (Eacott & Gaffan 1992, Gaffan & Harrison 1988, Parker & Gaffan 1998). Passingham (1993) argues that most, if not all, tasks that are disrupted following PFC damage depend on acquiring conditional associations (if-then rules).

In sum, these results indicate that PFC neural activity represents the rules, or mappings required to perform a particular task, and not just single stimuli or forthcoming actions. We assume that this activity within the PFC establishes these mappings by biasing competition in other parts of the brain responsible for actually performing the task. These signals favor task-relevant sensory inputs (attention), memories (recall), and motor outputs (response selection) and thus guide activity along the pathways that connect them (conditional association).

Feedback to Other Brain Areas

Our model of PFC function requires feedback signals from the PFC to reach widespread targets throughout the brain. The PFC has the neural machinery to provide these feedback signals; it sends projections to much of the neocortex (Pandya & Barnes 1987, Pandya & Yeterian 1990). Physiological studies have yielded results consistent with this notion.

Fuster et al (1985) and Chafee & Goldman-Rakic (2000) have found that deactivating the lateral PFC cortex attenuates the activity of visual cortical (inferior temporal and posterior parietal) neurons to a behaviorally relevant cue. Tomita et al (1999) directly explored the role of top-down PFC signals in the recall of visual memories stored in the inferior temporal (IT) cortex. Appearance of a cue object instructed monkeys to recall and then choose another object that was associated with the cue during training. In the intact brain, information is shared between IT cortices in the two cerebral hemispheres. By severing the connecting fibers, each IT cortex could only "see" (receive bottom-up inputs from) visual stimuli in the contralateral visual field. The fibers connecting the PFC in each hemisphere were left intact. When Tomita et al examined activity of single neurons in an IT cortex that could not "see" the cue, it nonetheless reflected the recalled object, albeit with a long latency. It appeared that visual information took a circuitous route, traveling from the opposite IT cortex (which could "see" the cue) to the still-connected PFC in each hemisphere and then down to the "blind" IT cortex. This was confirmed by severing the PFC in the two hemispheres and eliminating the feedback, which abolished the IT activity and disrupted task performance.

Other evidence suggestive of PFC-IT interactions also comes from investigations, by Miller & Desimone (1994) and Miller et al (1996), of the respective roles of the PFC and IT cortex in working memory. During each trial, monkeys were shown first a sample stimulus. Then, one to four test stimuli appeared in sequence. If a test stimulus matched the sample, the monkey indicated so by releasing a lever. Sometimes, one of the intervening nonmatch stimuli could be repeated. For example, the sample stimulus "A" might be followed by "B ... B ... C ... A." The monkey was only rewarded for responding to the final match ("A") and thus had to maintain a specific representation of the sample rather than respond to any repetition of any stimulus. As noted in the next section, neurons were found in the PFC that exhibited sustained sample-specific activity that survived the presentation of intervening distractors. This was not so for IT cortex. However, neurons in both areas showed a selective enhancement of responses to a match of the sample. The fact that IT neurons had not maintained a representation of this stimulus suggests that their enhanced response to the match might have resulted from interactions with the representation maintained in the PFC. This is consistent with the recent finding indicating that, in a target detection task, target-specific activity appears simultaneously within the PFC and the visual cortex (Anderson et al 1999). Together, these findings suggest that identification of an intended stimulus relies on interactions between the PFC and the posterior cortex.

Active Maintenance

If the PFC represents the rules of a task in its pattern of neural activity, it must maintain this activity as long as the rule is required. Usually this extends beyond the eliciting event and must span other intervening, irrelevant, and potentially

interfering events. The capacity to support sustained activity in the face of interference is one of the distinguishing characteristics of the PFC.

Sustained neural activity within the PFC was first reported by Fuster (1971) and Kubota & Niki (1971) and has subsequently been reported in a large number of studies. These have demonstrated that neurons within the PFC remain active during the delay between a transiently presented cue and the later execution of a contingent response. Such delay period activity is often specific to a particular type of information, such as the location and/or identity of a stimulus (di Pellegrino & Wise 1991; Funahashi et al 1989; Fuster 1973; Fuster & Alexander 1971; Kubota & Niki 1971; Rainer et al 1998a,b, 1999; Rao et al 1997; Romo et al 1999), forthcoming actions (Asaad et al 1998, Ferrera et al 1999, Quintana & Fuster 1992), expected rewards (Leon & Shadlen 1999, Tremblay et al 1998, Watanabe 1996), and more-complex properties such as the sequential position of a stimulus within an ordered series (Barone & Joseph 1989) or a particular association between a stimulus and its corresponding response (Asaad et al 1998). Functional neuroimaging studies have begun to yield similar results with humans (Cohen et al 1997, Courtney et al 1997, Prabhakaran et al 2000).

Other areas of the brain exhibit a simple form of sustained activity. For example, in many cortical visual areas, a brief visual stimulus will evoke activity that persists from several hundred milliseconds to several seconds (Fuster & Jervey 1981, Gnadt & Andersen 1988, Miller et al 1993, Miyashita & Chang 1988). What appears to distinguish the PFC is the ability to sustain such activity in the face of intervening distractions. When monkeys must sustain the memory of a sample object over a delay filled with visual distractors, each of which must be attended and processed, sustained activity in the PFC can maintain the sample memory across the distractors (Miller et al 1996). By contrast, sustained activity in extrastriate visual areas (such as the IT and posterior parietal cortex) is easily disrupted by distractors (Constantinidis & Steinmetz 1996; Miller et al 1993, 1996). Thus, posterior cortical neurons seem to reflect the most recent input regardless of its relevance, whereas the PFC selectively maintains task-relevant information.

Learning "Across Time" Within the PFC

Typically, the internal representation of goals and associated rules must be activated in anticipation of the behavior they govern. Furthermore, as we have seen, rules often involve learning associations between stimuli and behaviors that are separated in time. How can associations be learned between a rule or event that occurs at one point in time and contingent behaviors or rewards that occur later? The capacity of the PFC for active maintenance, coupled with its innervation by brainstem dopaminergic systems, suggests one way in which this might occur.

The capacity to actively maintain representations over time is fundamental to associative learning, as it allows information about fleeting events and actions to comingle that would otherwise be separated in time (Fuster 1985). For example, consider the Asaad et al (1998) study discussed above, in which the monkey needed

to associate a cue object with the direction of a saccade that had to be made after the cue was no longer present. Presumably, this was made possible by sustained activity within the PFC that insured that a representation of the cue persisted until a saccade was made. Furthermore, as learning progressed, activity related to the forthcoming saccade direction was triggered progressively earlier. Thus, even though the cue and action were separated in time, information about each was simultaneously present in the PFC, permitting an association to be formed between them. In addition to associating temporally separate events that were needed to form task rules, the PFC must be able to associate those rules, and the conditions that elicited them, with subsequent reward. This is necessary if patterns of PFC activity responsible for achieving a goal are to be reinforced, so that they are likely to recur under the appropriate conditions in the future. This function may be served by dopaminergic projections from the midbrain ventral tegmental area.

Midbrain dopamine (DA) neurons exhibit relatively low levels of spontaneous firing but give bursts of activity to behaviorally salient events, especially the delivery of unpredicted, desirable stimuli, such as food or juice rewards (Mirenowicz & Schultz 1994, 1996). As learning progresses, however, DA neurons become activated progressively earlier in time, by events that predict reward, and cease their activation to the now-expected reward (Schultz et al 1993). If the predicted reward fails to appear, activity is inhibited at the expected time of its delivery (Hollerman & Schultz 1998), and if the reward (or an event that has come to predict it) appears earlier than expected, it will again elicit DA neural responses. Thus, midbrain DA neurons seem to be coding "prediction error," the degree to which a reward, or a cue associated with reward, is surprising (Montague et al 1996, Schultz 1998). There is growing evidence that this mechanism operates within PFC. Neurons throughout the PFC (e.g. in lateral as well as ventromedial areas) convey information about expected rewards and show enhanced activity as the size and desirability of an expected reward increases (Leon & Shadlen 1999; Tremblay & Schultz 1999; Watanabe 1990, 1992, 1996). Similar observations have been made in human neuroimaging studies (London et al 2000, O'Doherty et al 2000).

The aim of the cognitive system is not only to predict reward but to pursue the actions that will ensure its procurement. The prediction error signal could help mediate this learning by selectively strengthening not only connections among neurons that provide information about the prediction of reward (Schultz et al 1997), but also their connections with representations in the PFC that guide the behavior needed to achieve it. The role of the PFC in mediating this relationship is suggested by the observation that frontally damaged patients exhibit disturbances in learning and decision-making tasks that involve the evaluation of reward (Bechara et al 1994, 1997; Rolls 2000). Evidence that a predictive reinforcement-learning mechanism may operate within the PFC was also observed by Asaad et al (1998). In this experiment, monkeys learned to associate visual cues with one of two saccadic responses. Initially, the monkeys chose their responses at random, but learned the correct cue-response pairing over a few trials. As they learned the association, neural activity representing the forthcoming saccadic response

appeared progressively earlier on successive trials. In other words, the initiation of response-related delay activity gradually shifted, with learning, from a point in time just before the execution of the response and reward delivery to progressively earlier points in time, until it was nearly coincident with the presentation of the cue. This evolution with learning closely resembles that of the reward-prediction signal thought to be mediated by the DA system (Montague et al 1996). In the section that follows, we discuss computational modeling work that provides further support for the plausibility of a DA-based learning mechanism that can establish an association between anticipation of reward and activation of representations in the PFC needed to achieve it.

Summary

Our review of studies in monkeys and humans shows that the PFC exhibits the properties required to support a role in cognitive control: sustained activity that is robust to interference; multimodal convergence and integration of behaviorally relevant information; feedback pathways that can exert biasing influences on other structures throughout the brain; and ongoing plasticity that is adaptive to the demands of new tasks. Of course, these properties are not unique to the PFC. They can be found elsewhere in the brain, to varying degrees and in various combinations. However, we argue that the PFC represents a specialization along this particular combination of dimensions that is optimal for a role in the brain-wide control and coordination of processing. In the section that follows, we consider the theoretical implications of this model, with reference to computational analyses that illustrate these in explicit and concrete form.

A GUIDED ACTIVATION THEORY OF PFC FUNCTION

Explorations of how properties of the PFC might be implemented to mediate control have been conducted using neural network models. Such models attempt to simulate the behavioral performance of human subjects (or animals) in cognitive tasks using neurobiologically plausible mechanisms (e.g. the spread of activity among simple processing units along weighted connections) in order to identify the principles that are most relevant to behavior. Using this approach, Dehaene & Changeux (1989, 1992), Levine & Prueitt (1989), Cohen & Servan-Schreiber (1992), and Braver et al (1995) have all described models of PFC function and have used these to simulate the performance of normal and frontally damaged patients in tasks that are sensitive to PFC damage, such as the Stroop task, WCST, and others. These models capture many of the PFC properties reviewed in the first part of this article and suggest how they might interact to engender cognitive control.

A Simple Model of PFC Function

Most neural network models that address the function of the PFC simulate it as the activation of a set of "rule" units whose activation leads to the production of a

response other than the one most strongly associated with a given input. In most models, the PFC units themselves are not responsible for carrying out input-output mappings needed for performance. Rather, they influence the activity of other units whose responsibility is making the needed mappings. This principle is illustrated in its simplest form by a model of the Stroop task developed by Cohen et al (1990).

This model (Figure 4) is made up of five sets of units required for carrying out each of the two possible tasks in a Stroop experiment, color naming and word reading: two sets of input units representing each of the two types of stimulus features (e.g. the colors red and green, and the orthographic features associated with the words RED and GREEN); a set of output units representing each potential response (e.g. the articulatory codes for "red" and "green"); and two sets of intermediate units that provide a pathway between each set of input units and the output units. Connections along the word-reading pathway are stronger as a consequence of more extensive and consistent use. The result is that when a Stroop conflict stimulus is presented (such as the word GREEN printed in red ink), information flowing along the word pathway dominates competition at the response level, and the model responds to the word (see Figure 4*A*). This captures the fact that in the absence of instructions, subjects routinely read the word (i.e. say "green"). However, when they are instructed to do so, subjects can instead name the color (i.e. say "red").

The ability to engage the weaker pathway requires the addition of a set of units (labeled "control" in Figure 4), which in this case, represent the two dimensions of the stimulus (color and word). Each of these control units is connected to intermediate units in the corresponding processing pathway. Activating one of these units biases processing in favor of that pathway by providing additional input to (i.e. "priming") the intermediate units along that pathway. In the case of the color pathway, this allows them to more effectively compete with and prevail over activity flowing along the stronger word pathway (see Figure 4*B*). This biasing effect corresponds to the role of top-down attentional control in the biased competition model proposed by Desimone & Duncan (1995). We assume that the control units in the Stroop model represent the function of neurons within the PFC; they establish the mapping between stimuli and responses required to perform the task. This model provides a concrete implementation of the scheme diagrammed in Figure 2. It and closely related models have been used to simulate quantitative features of the performance of both normal subjects and patients with frontal damage in a wide range of tasks that rely on cognitive control (Braver et al 1995; Cohen & Servan-Schreiber 1992; Cohen et al 1992, 1994a,b, 1996; Dehaene & Changeux 1992; Mozer 1991; Phaf et al 1990).

Guided Activation as a Mechanism of Cognitive Control

The Stroop model brings several features of our theory into focus. First, it emphasizes that the role of the PFC is modulatory rather than transmissive. That is, the pathway from input to output does not "run through" the PFC. Instead, the PFC guides activity flow along task-relevant pathways in more posterior and/or

subcortical areas. In this respect, the function of the PFC can be likened to that of a switch operator in a system of railroad tracks. We can think of the brain as a set of tracks (pathways) connecting various origins (e.g. stimuli) to destinations (responses). The goal is to get the trains (activity carrying information) at each origin to their proper destination as efficiently as possible, avoiding any collisions. When the track is clear (i.e. a train can get from its origin to destination without risk of running into any others), then no intervention is needed (i.e. the behavior can be carried out automatically and will not rely on the PFC). However, if two trains must cross the same bit of track, then some coordination is needed to guide them safely to their destinations. Patterns of PFC activity can be thought of as a map that specifies which pattern of "tracks" is needed to solve the task. In the brain, this is achieved by the biasing influence that patterns of PFC activity have on the flow of activity in other parts of the brain, guiding it along pathways responsible for task performance, just as activation of the color-control unit in the Stroop model biased processing in favor of the color-naming pathway. Note that this function need not be restricted to mappings from stimuli to responses but applies equally well to mappings involving internal states (e.g. thoughts, memories, emotions, etc), either as "origins" or "destinations," or both. Thus, depending on their target of influence, we can think of representations in the PFC as attentional templates, retrieval cues, rules, or goals, depending on whether the biasing influences target sensory processes, internal processes, particular courses of action, or their intended outcomes.

This distinction between modulation vs transmission is consistent with the classic pattern of neuropsychological deficits associated with frontal lobe damage. The components of a complex behavior are usually left intact, but the subject is not able to coordinate them in a task-appropriate way (for example, a patient who, when preparing coffee, first stirred and then added cream) (Shallice 1982, Levine et al 1998, Duncan et al 1996). The notion that the function of the PFC is primarily modulatory also makes some interesting and testable predictions. For example, in neuroimaging studies, it should be possible to find circumstances that activate more posterior cortical areas without activation of the PFC whereas it should be much less common to activate the PFC without associated posterior structures. In other words, although there should be circumstances under which transmission can occur without the need for modulation (e.g. word reading in the Stroop task), it does not make sense to have modulation in the absence of transmission.

Active Maintenance in the Service of Control

The Stroop model also illustrates another critical feature of our theory: the importance of sustained activity as a mechanism of control. For a representation to have a biasing influence, it must be activated over the course of performing a task. This feature of the model brings theories of PFC function into direct contact with cognitive psychological constructs, such as the relationship between controlled (PFC-mediated) and automatic processing. The model suggests that this

is a continuum, defined by the relative strength of the pathway supporting a task-relevant process compared with those carrying competing information (Cohen et al 1990). Thus, weaker pathways (such as for color naming) rely more on top-down support (i.e. the activity of control units), especially when they face competition from a stronger pathway (e.g. word reading).

This suggests that an increase in the demand for control requires greater or more-enduring PFC activation, which concurs with accumulating evidence from the neuroimaging literature that tasks thought to rely more heavily on controlled processing consistently engage the PFC (Baker et al 1996; Cohen et al 1994a,b, 1997; Frith et al 1991; MacDonald et al 2000; Smith & Jonides 1999; Banich et al 2000). Furthermore, it provides a mechanistic account of the long-standing observation that as a task becomes more practiced, its reliance on control (and the PFC) is reduced. This happens because practice strengthens the connections along the task-relevant pathway in other brain structures. Simulations using the Stroop model (Cohen et al 1990) capture detailed quantitative effects of practice both on measures of performance (e.g. power law improvements in speed of response) and concurrent changes in the reliance on control (e.g. Stroop interference). From a neural perspective, as a pathway is repeatedly selected by PFC bias signals, activity-dependent plasticity mechanisms can strengthen them. Over time, these circuits can function independently of the PFC, and performance of the task becomes more automatic. This concurs with studies showing that PFC damage impairs new learning while sparing well-practiced tasks (Rushworth et al 1997) and neuroimaging and neurophysiological studies that demonstrate greater PFC activation during initial learning and weaker activity as a task becomes more practiced (Knight 1984, 1997; Yamaguchi & Knight 1991; Asaad et al 1998; Shadmehr & Holcomb 1997, Petersen et al 1998).

Our view also provides an interpretation of the relationship between PFC function and working memory. Traditional theories of working memory have distinguished between storage and executive components (Baddeley 1986), with the former responsible for maintaining information online (i.e. in an activated state), and the latter responsible for its manipulation (i.e. the execution of control). Neuropsychological interpretations have placed the storage component in more posterior sensory and motor systems (e.g. Gathercole 1994), whereas the executive control component has been assigned to the PFC. By contrast, as reviewed above, early monkey neurophysiological studies have emphasized the role of the PFC in maintenance. Our theory offers a possible resolution of this dilemma. It suggests that executive control involves the active maintenance of a particular type of information: the goals and rules of a task. This view concurs with cognitive psychological theories based on production system architectures (e.g. ACT*) (Anderson 1983), which posit that executive control relies on the activation of representations that correspond to the goals of a behavior and the rules for achieving it.

This perspective also provides a unifying view of the role of the PFC in other cognitive functions with which it has been associated, most commonly attention and inhibition. Both can be seen as varying reflections, in behavior, of the operation

of a single underlying mechanism of cognitive control: the biasing effects of PFC activity on processing in pathways responsible for task performance. As suggested by the biased competition model of Desimone & Duncan (1995), selective attention and behavioral inhibition are two sides of the same coin: Attention is the effect of biasing competition in favor of task-relevant information, and inhibition is the consequence that this has for the irrelevant information. Note that according to this view, inhibition occurs because of local competition among conflicting representations (e.g. between the two responses in the Stroop model) rather than centrally by the PFC. The "binding" function of selective attention (e.g. Treisman & Gelade 1980) can also be explained by such a mechanism if, in this case, we think of PFC representations as selecting the desired combination of stimulus features to be mapped onto the response over other competing combinations.

Finally, it is important to distinguish the form of activity-dependent control that we have ascribed to PFC from other forms of control that may occur in the brain. In particular, we believe that PFC-mediated control is complemented by another form of control dependent on the hippocampal system. The hippocampus is important for binding together information into a memory of a specific episode (Eichenbaum et al 1999, McClelland et al 1995, Squire 1992, Zola-Morgan & Squire 1993). By contrast, we suggest that the PFC, like other neocortical areas, is more important for extracting the regularities across episodes—in the case of the PFC, those corresponding to goals and task rules, rather than episodic memories of actually performing the task. We further posit that the PFC uses "activity-based" control; that is, its ongoing activity specifies the pattern of neural pathways that are currently needed. If PFC activity changes, so does the selected pattern of pathways. By contrast, the hippocampus may provide a form of "weighted-based" control; it helps consolidate permanent associative links between the pieces of information that define a long-term memory (Cohen & O'Reilly 1996, O'Reilly et al 1999, O'Reilly 2000). To use the railroad metaphor, the hippocampus is responsible for laying down new tracks and the PFC is responsible for flexibly switching between them. As noted below, interactions between the PFC and the hippocampus may provide a basis for understanding prospective forms of control, such as planning.

Updating of PFC Representations

In the real world, cognitive control is highly dynamic. People move from one task to the next, and new goals replace old ones. A major benefit of the activity-based mechanism of control that we have proposed is that it is highly flexible. So long as suitable representations exist within the PFC, activating them can quickly invoke a goal or rule, which can be flexibly switched to others as circumstances demand. That is, it is easier and faster (and perhaps less costly) to switch between existing tracks than it is to lay new ones down. This is clearly illustrated by models of PFC function in the WCST (Dehaene & Changeux 1992) and recent variants of this task (RC O'Reilly, DC Noelle, TS Braver, JD Cohen, submitted for publication), and it is supported by the fact that damage to the PFC impairs such flexibility

(Milner 1963, Dias et al 1996b, Rossi et al 1999). It should be noted, however, that the mechanisms responsible for updating representations within the PFC must be able to satisfy two conflicting demands: On the one hand, they must be responsive to relevant changes in the environment (adaptive); on the other, they must be resistant to updating by irrelevant changes (robust). As described above, neurophysiological studies suggest that PFC representations are selectively responsive to task-relevant stimuli (Rainer et al 1998b), yet they are robust to interference from distractors (Miller et al 1996). Conversely, two hallmarks of damage to the PFC are perseveration (inadequate updating) and increased distractibility (inappropriate updating) (e.g. Mishkin 1964, Chao & Knight 1997). These observations suggest the operation of mechanisms that ensure the appropriate updating of PFC activity in response to behavioral demands.

Cohen et al (1996) and Braver & Cohen (2000) have proposed that DA may play an important role in this function. They hypothesize that DA release may "gate" access to the PFC by modulating the influence of its afferent connections. A similar role for DA in the PFC has been suggested by Durstewitz et al (1999, 2000). Timing is a critical feature of such a gating mechanism: The signal-to-gate input must be rapid and coincide with the conditions under which an update is needed. This is consistent with recent studies indicating that DA release (once thought to be slow and nonspecific) has a phasic component with timing characteristics consistent with its proposed role in gating (Schultz 1998, Schultz & Dickinson 2000). As discussed above, midbrain DA neurons give bursts of activity to stimuli that are not predicted but that predict a later reward. This is precisely the timing required for a gating signal responsible for updating goal representations. For example, imagine that you are walking to work and out of the corner of your eye you notice a $20 bill lying on the ground. This unpredicted stimulus predicts reward, but only if you update your current goal and bend down to pick up the bill.

It is intriguing that the properties of midbrain DA neurons that may drive PFC associative learning mechanisms (discussed above) are formally equivalent to those used in models that simulate PFC updating mechanisms. These dual and concurrent influences of DA on gating and learning suggest that if the system learns while it gates, then perhaps it can learn on its own when to gate. That is, if an exploratory DA-mediated gating signal leads to a successful behavior, its coincident reinforcing effects will strengthen the association of this signal with cues representing the current context and the pattern of activity within the PFC that produced the behavior. This will increase the probability that in the future, the same context will elicit a gating signal and reactivation of the PFC activity that led to the rewarded behavior, which in turn will produce further reinforcement of these associations, etc. Recent computational modeling studies establish the plausibility of this bootstrapping mechanism (Braver & Cohen 2000). This ability to self-organize averts the problem of a theoretical regress regarding control (i.e. the invocation of a "homunculus") by allowing the system to learn on its own what signals should produce an updating of the contents of the PFC and when this should occur. However, an important issue that requires further exploration

is whether this system can support "subgoaling," or hierarchical updating—that is, the updating of some representations (e.g. the next chess move to make) while preserving others (e.g. the strategy being pursued). A closely related issue concerns the proper sequencing of actions. These abilities are fundamental to virtually all higher cognitive faculties, such as reasoning and problem solving (Newell & Simon 1972), and are known to involve the PFC (e.g. Duncan 1986; Duncan et al 1996; Baker et al 1996; Koechlin 1999; Nichelli et al 1994; Shallice 1982, 1988). Recent modeling work suggests that hierarchical updating and the sequencing of actions may rely on interactions between the PFC and the basal ganglia (Gobbel 1995, Houck 1995); however, a full elaboration of the mechanisms involved remains a challenge for future work.

REMAINING ISSUES

A detailed consideration of all the issues relevant to cognitive control is beyond the scope of this article. However, some additional issues are important for a more complete theory of PFC function and cognitive control. Here, we briefly review several important remaining issues.

Representational Power of the PFC

The tremendous range of tasks of which people are capable raises important questions about the ability of the PFC to support the necessary scope of representations. The large size of the PFC (over 30% of the cortical mass), coupled with its anatomic connectivity discussed earlier, suggests that it can support a wide number and range of mappings. However, there are a nearly limitless number of tasks that a person can be asked to perform (e.g. "wink whenever I say bumblydoodle"), and it seems unlikely that all possible mappings are represented within the PFC. It is possible that some as-yet-undiscovered representational scheme supports a wide enough range of mappings to account for the flexibility of human behavior. However, more likely, it seems that plasticity may also play an important role in PFC function, establishing new representations as they are needed. This need is accentuated by virtue of the fact that the PFC must be able to modulate processes in other parts of the brain that are themselves plastic. Conversely, the circuitry within the PFC that supports older, well-established behaviors is likely to be "reclaimed" as PFC-independent pathways become responsible for them. Above, we reviewed evidence that the PFC exhibits a high degree of plasticity. As yet, however, the mechanisms that govern this plasticity are largely unknown, at either the neurobiological or the computational levels. At the neural level, this could involve the modification of existing synapses, the formation of new ones (perhaps with the assistance of rapid-learning mechanisms in the hippocampus), or even the recruitment of entirely new neurons (Gould et al 1999). At the computational level, we have suggested how predictive reinforcement-learning mechanisms can (*a*) strengthen patterns of PFC

activity that appropriately guide behavior and (*b*) associate them with the circumstances in which they are useful. However, the ability of such mechanisms to account for the wide range and extraordinary flexibility of human behavior remain to be established.

PFC Functional Organization

One important question concerns how PFC representations are functionally organized. Understanding the principles of PFC organization is likely to provide insights into how its representations develop and function. Various schemes have been proposed. For example, one possibility is that the PFC is organized by function, with different regions carrying out qualitatively different operations. One long-standing view is that orbital and medial areas are associated with behavioral inhibition, whereas ventrolateral and dorsal regions are associated with memory or attentional functions (Fuster 1989, Goldman-Rakic 1987). Another recent suggestion is that ventral regions support maintenance of information (memory), whereas dorsal regions are responsible for the manipulation of such information (Owen et al 1996, Petrides 1996). Such distinctions have heuristic appeal. However, our theory suggests an intriguing alternative.

If different regions of the PFC emphasize different types of information, then perhaps variations in the biasing signals that they provide can account for apparent dissociations of function. For example, both activity and deficits of orbital PFC are most frequently associated with tasks involving social, emotional, and appetitive stimuli (Hecaen & Albert 1978, O'Doherty et al 2000, Price 1999, Stuss & Benson 1986, Swedo et al 1989), whereas more-dorsal regions are activated in tasks involving more-"cognitive" dimensions of stimuli (form, location, sequential order, etc). Social and appetitive stimuli are "hot," meaning they are more likely to elicit reflexive (and often inappropriate) reactions. Thus, the impression that the orbital PFC subserves an inhibitory function may be explained by the fact that it is more involved in biasing task-relevant processes against strong competing alternatives. In contrast, more-cognitive stimuli (e.g. shapes, locations, etc) that engage more-dorsal regions are "cold," meaning they are less likely to engage responses with such asymmetries of strength. Thus, their competition is likely to be less fierce. Neuropsychological studies of monkeys support this notion (Dias et al 1996a, Roberts & Wallis 2000), and recent computational modeling has shown that dissociations in performance interpreted as evidence of a distinction between inhibitory and memory processes within the PFC can, alternatively, be explained in terms of a single processing mechanism operating over different types of representations (RC O'Reilly, DC Boelle, TS Braver, JD Cohen, submitted for publication).

Other organizational schemes have also been proposed for the PFC, including those based on stimulus dimensions, sensory vs motor, and sequential order (e.g. Barone & Joseph 1989, Wilson et al 1993, Wagner 1999, Casey et al 2000). Although our theory does not provide deep insights into which, if any, is most

likely to be correct, it does make strong claims that are related to this issue. First, although it allows for the possibility of broad categories or gradients of organization, it suggests that it is unlikely that different classes of information will be represented in a modular, or discretely localized, form. Complex behavior requires that we recognize and respond to relationships across diverse dimensions, and the role we have ascribed to the PFC involves representing these relationships. Both neurophysiological (Asaad et al 1998, Bichot et al 1996, Quintana & Fuster 1992, Rainer et al 1998a, Rao et al 1997, Watanabe 1990, White & Wise 1999) and neuorimaging (Cohen 2000, Nystrom et al 2000, Prabhakaran et al 2000) findings support this view, which suggests that most regions of the PFC can respond to a variety of different types of information.

Second, our theory suggests that learning will play an important role in the formation of representations in the PFC and, thus, may have an important influence on representational organization. This has been illustrated in a computational model of the PFC. Braver et al (1996) trained a network on a "spatial" memory task and an "object" memory task. The PFC module included some units with projections from both spatial and object pathways, and some units with projections from only one or the other. When the network was trained on each task separately, the PFC module relied primarily on the segregated location units and object units to perform the task. When training on the tasks was intermixed, the activity and number of the multimodal units were increased. This result has found support in empirical studies. Evidence of segregation in the PFC by stimulus domain has been reported by studies that separate training of different stimulus attributes or in monkeys passively viewing stimuli (i.e. in tasks that do not engage PFC function) (O Scalaidhe et al 1997, Wilson et al 1993). Evidence for PFC integration has come from tasks in which stimulus domains are intermixed or their integration is relevant (Asaad et al 1998, Bichot et al 1996, Fuster et al 1982, Prabhakaran et al 2000, Rainer et al 1998a, Rao et al 1997, White & Wise 1999).

Monitoring and the Allocation of Control

Previous discussion has focused on the need to appropriately update representations in the PFC as new goals arise and new rules are applied. However, as noted earlier, people also show a facility for adapting the degree of control they allocate to a task. For example, you pay closer attention to the road on a dark and rainy night than on a bright, sunny day. Such adjustments are adaptive, in view of the well-recognized capacity limits on cognitive control (discussed below). In our model, such adjustments would correspond to strength of the PFC pattern of activity (e.g. the strength of the color unit in the Stroop model). A stronger pattern of activity within the PFC produces stronger biasing effects for a particular pathway, but possibly at the expense of other ones (e.g. through competition among PFC representations). Recent studies have suggested that the allocation of control may depend on signals from the anterior cingulate cortex (ACC) that detect conflict in processing (e.g. Carter et al 1998, 2000; Botvinick et al 1999). Drawing on our

train track analogy, conflict occurs when two trains are destined to cross tracks at the same time. In neural terms, this corresponds to the coactivation of competing (i.e. mutually inhibitory) sets of units (e.g. for the responses "red" and "green" in the Stroop task). Such conflict produces uncertainly in processing, and an increased probability of errors. Thus, conflict signals the need for the allocation of additional control. Modeling work (M Botvinick, TS Brauer, CS Carter, DM Barch, JD Cohen, submitted for publication) has shown that coupling the conflict signal (detected by ACC) to adjustments in the allocation of control (amplification of PFC pattern of activity) can accurately simulate trial-based adjustments that subjects make in their behavior in experimental tasks (e.g. Botvinick et al 1999, Gratton et al 1992, Laming 1968, Logan et al 1983, Tzelgov et al 1992).

The tight coupling of conflict detection and allocation of control may explain the pervasive finding of coactivation of the PFC and the ACC in most neuroimaging studies (Owen & Duncan 2000). However, the distinct roles of the PFC and ACC are illustrated by a dissociation of their activity in a recent functional magnetic resonance imaging study by MacDonald et al (2000), using an instructed version of the Stroop task (Figure 5). In each trial, subjects were given a cue indicating whether they were to name the color or read the word in the subsequent display. The cue was followed by a delay of several seconds, and then either a congruent or a conflict stimulus was displayed. Figure 5 shows that during the delay, increasing activity was observed within a region of the DL PFC, greater for color naming (the more conrol demanding task) than word reading. There was no differential activation observed within ACC during this period. In contrast, strong activation was observed in ACC during the period of stimulus presentation and responding. This activity was greater for conflict than congruent stimuli. There was no differential response for these trial types within PFC during this period. These findings provide strong support for several of the hypotheses we have discussed: The demands for control are associated with an increase in PFC activity; tasks demanding greater control elicit stronger activity within the PFC; and the ACC responds selectively to conflict in processing. However, further work is needed to establish the causal relationship between detection of conflict within the ACC and the augmentation of control by the PFC.

Mechanisms of Active Maintenance

Our theory of the PFC, like many others, emphasizes its capacity for active maintenance. However, there has been relatively little empirical research on the mechanisms responsible for sustained activity. There are a number of theoretical possibilities, which can be roughly divided into two classes: cellular and circuit based. Cellular models propose neuron bistability as the basis of sustained activity, which is dependent on the biophysical properties of individual cells. The transitions between states are triggered by inputs to the PFC but maintained via the activation of specific voltage-dependent conductances (Wang 1999). Circuit-based models, on the other hand, propose that the recirculation of activity through closed (or

"recurrent") loops of interconnected neurons, or attractor networks (Hopfield 1982), support self-sustained activity (Zipser et al 1993). These loops could be intrinsic to the PFC (Pucak et al 1996, Melchitzky et al 1998), or they might involve other structures, such as the cortex-striatal-globus pallidus-thalamus-cortex loops (Alexander et al 1996). In either case, it should be noted that a mechanism is needed for regulating the updating of activity within the PFC, as discussed above.

Capacity Limits of Control

A better understanding of the mechanisms underlying active maintenance may provide insight into one of the most perplexing properties of cognitive control: its severely limited capacity. This has long been recognized in cognitive psychology (Broadbent 1958, Posner & Snyder 1975, Shiffrin & Schneider 1977) and is painfully apparent to anyone who has tried to talk on the phone and read e-mail at the same time. It is important to distinguish between this form of capacity, which has to do with how many representations can be actively maintained at the same time, and the issue of representational power discussed above, which has to do with the range of representations that are available to draw from in the PFC. It may also be important to distinguish the capacity limits of cognitive control from those of short-term storage of item information (e.g. verbal or visual short term memory) (Miller 1956, Baddeley 1986). Limits of control presumably reflect properties of PFC function. The limited capacity of short-term memory may involve mechanisms (e.g. articulatory rehearsal) and structures (e.g. sustained activity in posterior cortical areas) that are not central to cognitive control, and that may or may not rely on PFC function. The capacity limits of cognitive control have been used to explain many features of human cognition (e.g. Cowen 1988, Engle et al 1999, Just & Carpenter 1992, Posner & Snyder 1975, Shiffrin & Schneider 1977). However, to date, no theory has provided an explanation of the capacity limitation itself. This could reflect an inherent physiological constraint, such as the energetic requirements of actively maintaining representations in the PFC. More likely, it reflects fundamental computational properties of the system, such as an inherent limit on the number of representations that can be actively maintained and kept independent of one another within an attractor network (e.g. Usher & Cohen 1999). In any event, capacity constraints are a sine qua non of cognitive control and, therefore, provide an important benchmark for theories that seek to explain its underlying mechanisms.

Prospective Control and Planning

Perhaps the most impressive feature of human cognition is its ability to plan for the future. We often forgo pursuing a given goal until a more appropriate time. Active maintenance cannot account for this. When we plan in the morning to go to the grocery story on the way home from work, it seems unlikely that we actively maintain this information in the PFC throughout the day. More likely, this information is stored elsewhere and then activated at the appropriate time. This

may involve interactions between the PFC and other brain systems capable of rapid learning, such as the hippocampus (cf Cohen & O'Reilly 1996, O'Reilly & McClelland 1994). Thus, it is possible that the hippocampus rapidly encodes an association between the desired goal representation(s) within the PFC (e.g. go to the grocery store) and features of the circumstance under which the goal should be evoked (e.g. commuting home). Then, as these circumstances arise, the appropriate representation within the PFC is associatively activated, guiding performance in accord with the goal and its associated rules (e.g. turning right at the light toward the store rather than the habitual left toward home). Neurophysiological studies suggest that the PFC is important for the ability to prospectively activate long-term memories (Rainer et al 1999, Tomita et al 1999). However, the detailed nature of such interactions, and their relationship to the dopaminergic gating and learning mechanisms described above, remain to be fully specified.

CONCLUSIONS

One of the great mysteries of the brain is cognitive control. How can interactions between millions of neurons result in behavior that is coordinated and appears willful and voluntary? There is consensus that it depends on the PFC, but there has been little understanding of the neural mechanisms that endow it with the properties needed for executive control. Here, we have suggested that this stems from several critical features of the PFC: the ability of experience to modify its distinctive anatomy; its wide-ranging inputs and intrinsic connections that provide a substrate suitable for synthesizing and representing diverse forms of information needed to guide performance in complex tasks; its capacity for actively maintaining such representations; and its regulation by brainstem neuromodulatory systems that provide a means for appropriately updating these representations and learning when to do so. We have noted that depending on their target of influence, representations in the PFC can function variously as attentional templates, rules, or goals by providing top-down bias signals to other parts of the brain that guide the flow of activity along the pathways needed to perform a task. We have pointed to a rapidly accumulating and diverse body of evidence that supports this view, including findings from neurophysiological, neuroanatomical, human behavioral and neuroimaging, and computational modeling studies.

The theory we have described provides a framework within which to formulate hypotheses about the specific mechanisms underlying the role of the PFC in cognitive control. We have reviewed a number of these, some of which have begun to take explicit form in computational models. We have also provided a sampling of the many questions that remain about these mechanisms and the functioning of the PFC. Regardless of whether the particular hypotheses we have outlined accurately describe PFC function, they offer an example of how neurally plausible mechanisms can exhibit the properties of self-organization and self-regulation required to account for cognitive control without recourse to a "homunculus." At the very

least, we hope that they provide some useful examples of how the use of a computational and empirical framework, in an effort to be mechanistically explicit, can provide valuable leads in this conceptually demanding pursuit. We believe that future efforts to address the vexing, but important, questions surrounding PFC function and cognitive control will benefit by ever tighter coupling of neurobiological experiments and detailed computational analysis and modeling.

ACKNOWLEDGMENTS

The order of authorship was arbitrarily determined; both authors made equal contributions to the research and preparation of this manuscript. Our work has been supported by grants from the NINDS, the NIMH, the RIKEN-MIT Neuroscience Research Center, the Pew Charitable Trusts, the McKnight Foundation, the Whitehall Foundation, the John Merck Fund, the Alfred P. Sloan Foundation, and NARSAD. We thank Wael Asaad, Mark Histed, Jonathan Wallis, Richard Wehby, and Marlene Wicherski for valuable comments on this manuscript. EKM would also like to thank Wael Asaad for valuable conversations. JDC would like to thank BJ Casey and Leigh Nystrom for thoughtful discussions and comments regarding this manuscript, as well as Todd Braver, David Noelle, and Randy O'Reilly, with whom many of the theoretical ideas presented in this article were developed.

LITERATURE CITED

Alexander GE, Delong MR, Strick PL. 1986. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* 9:357–81

Amaral DG, Price JL. 1984. Amygdalo-cortical projections in the monkey (Macaca fascicularis). *J. Comp. Neurol.* 230:465–96

Anderson JR. 1983. *The Architecture of Cognition*. Cambridge, MA: Harvard Univ. Press

Anderson KC, Asaad WF, Wallis JD, Miller EK. 1999. Simultaneous recordings from monkey prefrontal (PF) and posterior parietal (PP) cortices during visual search. *Soc. Neurosci. Abstr.* 25:885

Asaad WF, Rainer G, Miller EK. 1998. Neural activity in the primate prefrontal cortex during associative learning. *Neuron* 21:1399–407

Asaad WF, Rainer G, Miller EK. 2000. Task-specific neural activity in the primate prefrontal cortex. *J. Neurophysiol.* 84:451–59

Baddeley A. 1986. *Working Memory*. Oxford: Clarendon

Baker SC, Rogers RD, Owen AM, Frith CD, Dolan RJ, et al. 1996. Neural systems engaged by planning: a PET study of the Tower of London Task. *Neuropsychologia* 34:515–26

Banich MT, Milham MP, Atchley R, Cohen NJ, Webb A, et al. 2000. Prefrontal regions play a predominant role in imposing an attentional "set": evidence from fMRI. *Cogn. Brain Res.* In press

Barbas H, De Olmos J. 1990. Projections from the amygdala to basoventral and mediodorsal prefrontal regions in the rhesus monkey. *J. Comp. Neurol.* 300:549–71

Barbas H, Pandya D. 1991. Patterns of connections of the prefrontal cortex in the rhesus monkey associated with cortical architecture. In *Frontal Lobe Function and Dysfunction*, ed. HS Levin, HM Eisenberg, AL Benton,

pp. 35–58. New York: Oxford Univ. Press
Barbas H, Pandya DN. 1989. Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. *J. Comp. Neurol.* 286:353–75
Barone P, Joseph JP. 1989. Prefrontal cortex and spatial sequencing in macaque monkey. *Exp. Brain Res.* 78:447–64
Bates JF, Goldman-Rakic PS. 1993. Prefrontal connections of medial motor areas in the rhesus monkey. *J. Comp. Neurol.* 336:211–28
Bechara A, Damasio AR, Damasio H, Anderson SW. 1994. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* 50:7–15
Bechara A, Damasio H, Tranel D, Damasio AR. 1997. Deciding advantageously before knowing the advantageous strategy. *Science* 275:1293–95
Bianchi L. 1922. *The Mechanism of the Brain and the Function of the Frontal Lobes*. Edinburgh: Livingstone
Bichot NP, Schall JD, Thompson KG. 1996. Visual feature selectivity in frontal eye fields induced by experience in mature macaques. *Nature* 381:697–99
Botvinick M, Nystrom LE, Fissell K, Carter CS, Cohen JD. 1999. Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature* 402:179–81
Braver TS, Cohen JD, eds. 2000. *On the Control of Control: The Role of Dopamine in Regulating Prefrontal Function and Working Memory*. Cambridge, MA: MIT Press. In press
Braver TS, Cohen JD, Servan-Schreiber D. 1995. A computational model of prefrontal cortex function. In *Advances in Neural Information Processing Systems*, ed. DS Touretzky, G Tesauro, TK Leen, pp. 141–48. Cambridge, MA: MIT Press
Broadbent DE. 1958. *Perception and Communication*. London: Pergamon
Bruce C, Desimone R, Gross CG. 1981. Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J. Neurophysiol.* 46:369–84
Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD. 1998. Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* 280:747–49
Carter CS, Macdonald AM, Botvinick M, Ross LL, Stenger VA, et al. 2000. Parsing executive processes: strategic vs. evaluative functions of the anterior cingulate cortex. *Proc. Natl. Acad. Sci. USA* 97:1944–48
Casey BJ, Forman SD, Franzen P, Berkowitz A, Braver TS, et al. 2000. Sensitivity of prefrontal cortex to changes in target probability: a functional MRI study. *Hum. Brain Mapp.* In press
Chafee MV, Goldman-Rakic PS. 2000. Inactivation of parietal and prefrontal cortex reveals interdependence of neural activity during memory-guided saccades. *J. Neurophysiol.* 83:1550–66
Chao LL, Knight RT. 1997. Prefrontal deficits in attention and inhibitory control with aging. *Cereb. Cortex* 7:63–9
Chavis DA, Pandya DN. 1976. Further observations on cortico-frontal connections in the rhesus monkey. *Brain Res.* 117:369–86
Cohen JD. 2000. Special issue: functional topography of prefrontal cortex. *Neuroimage* 11:378–79
Cohen JD, Barch DM, Carter CS, Servan-Schreiber D. 1999. Schizophrenic deficits in the processing of context: converging evidence from three theoretically motivated cognitive tasks. *J. Abnorm. Psychol.* 108: 120–33
Cohen JD, Braver TS, O'Reilly RC. 1996. A computational approach to prefrontal cortex, cognitive control, and schizophrenia: Recent developments and current challenges. *Philos. Trans. Roy. Soc. London B.* 351:1515–1527
Cohen JD, Dunbar K, McClelland JL. 1990. On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychol. Rev.* 97:332–61
Cohen JD, Forman SD, Braver TS, Casey BJ, Servan-Schreiber D, Noll DC. 1994a. Activation of prefrontal cortex in a nonspatial

working memory task with functional MRI. *Hum. Brain Mapp.* 1:293–304

Cohen JD, O'Reilly RC. 1996. A preliminary theory of the interactions between the prefrontal cortex and hippocampus that contribute to planning and prospective memory. In *Prospective Memory: Theory and Applications*, ed. M Brandimonte, G Einstein, M McDaniel. Hillsdale, New Jersey: Erlbaum

Cohen JD, Perlstein WM, Braver TS, Nystrom LE, Noll DC, et al. 1997. Temporal dynamics of brain activation during a working memory task. *Nature* 386:604–8

Cohen JD, Romero RD, Farah MJ, Servan-Schreiber D. 1994b. Mechanisms of spatial attention: the relation of macrostructure to microstructure in parietal neglect. *J. Cogn. Neurosci.* 6:377–87

Cohen JD, Servan-Schreiber D. 1992. Context, cortex and dopamine: a connectionist approach to behavior and biology in schizophrenia. *Psychol. Rev.* 99:45–77

Cohen JD, Servan-Schreiber D, McClelland JL. 1992. A parallel distributed processing approach to automaticity. *Am. J. Psychol.* 105:239–69

Constantinidis C, Steinmetz MA. 1996. Neuronal activity in posterior parietal area 7a during the delay periods of a spatial memory task. *J. Neurophysiol.* 76:1352–55

Courtney SM, Ungerleider LG, Keil K, Haxby JV. 1997. Transient and sustained activity in a distributed neural system for human working memory. *Nature* 386:608–12

Cowen N. 1998. Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychol. Bull.* 104:163–91

Dehaene S, Changeux JP. 1989. A simple model of prefrontal cortex function in delayed-response tasks. *J. Cogn. Neurosci.* 1:244–61

Dehaene S, Changeux JP. 1992. The Wisconsin card sorting test: theoretical analysis and modeling in a neuronal network. *Cerebr. Cortex* 1:62–79

Desimone R, Duncan J. 1995. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18:193–222

Diamond A. 1988. Abilities and neural mechanisms underlying AB performance. *Child Dev.* 59:523–27

Diamond A, Goldman-Rakic PS. 1989. Comparison of human infants and rhesus monkeys on Piaget's A-not-B task: evidence for dependence on dorsolateral prefrontal cortex. *Exp. Brain Res.* 74:24–40

Dias R, Robbins TW, Roberts AC. 1996a. Dissociation in prefrontal cortex of affective and attentional shifts. *Nature* 380:69–72

Dias R, Robbins TW, Roberts AC. 1996b. Primate analogue of the Wisconsin Card Sorting Test: effects of excitotoxic lesions of the prefrontal cortex in the marmoset. *Behav. Neurosci.* 110:872–86

Dias R, Robbins TW, Roberts AC. 1997. Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sort Test: restriction to novel situations and independence from "on-line" processing. *J. Neurosci* 17:9285–97

di Pellegrino G, Wise SP. 1991. A neurophysiological comparison of three distinct regions of the primate frontal lobe. *Brain* 114:951–78

Dunbar K, Sussman D. 1995. Toward a cognitive account of frontal lobe function: simulating frontal lobe deficits in normal subjects. *Ann. NY Acad. Sci.* 769:289–304

Duncan J. 1986. Disorganization of behaviour after frontal lobe damage. *Cogn. Neuropsychol.* 3:271–90

Duncan J, Emslie H, Williams P, Johnson R, Freer C. 1996. Intelligence and the frontal lobe: the organization of goal-directed behavior. *Cogn. Psychol.* 30:257–303

Duncan J, Owen AM. 2000. Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci.* In press

Durstewitz D, Kelc M, Gunturkun O. 1999. A neurocomputational theory of the dopaminergic modulation of working memory functions. *J. Neurosci.* 19:2807–22

Durstewitz D, Seamans JK, Sejnowski TJ. 2000. Dopamine-mediated stabilization of

delay-period activity in a network model of the prefrontal cortex. *J. Neurophysiol.* 83:1733–50

Eacott MJ, Gaffan D. 1992. Inferotemporal-frontal disconnection—the uncinate fascicle and visual associative learning in monkeys. *Eur. J. Neurosci.* 4:1320–32

Eichenbaum H, Dudchenko P, Wood E, Shapiro M, Tanila H. 1999. The hippocampus, memory, and place cells: Is it spatial memory or a memory space? *Neuron* 23:209–26

Engel RW, Kane M, Tuholski S. 1999a. Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In *Mechanisms of Active Maintenance and Executive Control*, ed. A Miyake, P Shah. New York: Cambridge Univ. Press

Engel RW, Tuholski SW, Laughlin JE, Conway AR. 1999b. Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *J. Exp. Psychol. Gen.* 128:309–31

Ferrera VP, Cohen J, Lee BB. 1999. Activity of prefrontal neurons during location and color delayed matching tasks. *NeuroReport* 10:1315–22

Ferrier D. 1876. *The Functions of the Brain*. London: Smith, Elder

Frith CD, Friston K, Liddle PF, Frackowiak RSJ. 1991. Willed action and the prefrontal cortex in man: a study with PET. *Proc. R. Soc. London Ser. B* 244:241–46

Funahashi S, Bruce CJ, Goldman-Rakic PS. 1989. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* 61:331–49

Fuster JM. 1973. Unit activity in prefrontal cortex during delayed-response performance: neuronal correlates of transient memory. *J. Neurophysiol.* 36:61–78

Fuster JM. 1980. *The Prefrontal Cortex*. New York: Raven

Fuster JM. 1985. The prefrontal cortex, mediator of cross-temporal contingencies. *Hum. Neurobiol.* 4:169–79

Fuster JM. 1989. *The Prefrontal Cortex*, Vol. 2. New York: Raven

Fuster JM. 1995. *Memory in the Cerebral Cortex*. Cambridge, MA: MIT Press

Fuster JM, Alexander GE. 1971. Neuron activity related to short-term memory. *Science* 173:652–54

Fuster JM, Bauer RH, Jervey JP. 1982. Cellular discharge in the dorsolateral prefrontal cortex of the monkey in cognitive tasks. *Exp. Neurol.* 77:679–94

Fuster JM, Bauer RH, Jervey JP. 1985. Functional interactions between inferotemporal and prefrontal cortex in a cognitive task. *Brain Res.* 330:299–307

Fuster JM, Jervey JP. 1981. Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science* 212:952–55

Gaffan D, Harrison S. 1988. Inferotemporal-frontal disconnection and fornix transection in visuomotor conditional learning by monkeys. *Behav. Brain Res.* 31:149–63

Gathercole SE. 1994. Neuropsychology and working memory: a review. *Neuropsychology* 8:494–505

Gershberg FB, Shimamura AP. 1995. Impaired use of organizational strategies in free recall following frontal lobe damage. *Neuropsychologia* 13:1305–33

Gnadt JW, Andersen RA. 1988. Memory related motor planning activity in posterior parietal cortex of macaque. *Exp. Brain Res.* 70:216–20

Gobbel JR. 1995. A biophysically-based model of the neostriatum as a reconfigurable network. *Proc. Swed. Conf. Connectionism, 2nd*. Hillsdale, NJ: Erlbaum

Goldman PS, Nauta WJ. 1976. Autoradiographic demonstration of a projection from prefrontal association cortex to the superior colliculus in the rhesus monkey. *Brain Res.* 116:145–49

Goldman-Rakic PS. 1987. Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In *Handbook of Physiology: The Nervous System*, ed. F Plum,

pp. 373–417. Bethesda, MD: Am. Physiol. Soc.

Goldman-Rakic PS. 1996. The prefrontal landscape: implications of functional architecture for understanding human mentation and the central executive. *Philos. Trans. R. Soc. London Ser. B* 351:1445–53

Goldman-Rakic PS, Schwartz ML. 1982. Interdigitation of contralateral and ipsilateral columnar projections to frontal association cortex in primates. *Science* 216:755–57

Goldman-Rakic PS, Selemon LD, Schwartz ML. 1984. Dual pathways connecting the dorsolateral prefrontal cortex with the hippocampal formation and parahippocampal cortex in the rhesus monkey. *Neuroscience* 12:719–43

Gould E, Reeves AJ, Graziano MS, Gross CG. 1999. Neurogenesis in the neocortex of adult primates. *Science* 286:548–52

Grafman J. 1994. Alternative frameworks for the conceptualization of prefrontal functions. In *Handbook of Neuropsychology*, ed. F Boller, J Grafman, pp. 187. Amsterdam: Elsevier

Gratton G, Coles MGH, Donchin E. 1992. Optimizing the use of information: strategic control of activation of responses. *J. Exp. Psychol.* 121:480–506

Halsband U, Passingham RE. 1985. Premotor cortex and the conditions for movement in monkeys. *Behav. Brain Res.* 18:269–76

Hecaen H, Albert ML. 1978. *Human Neuropsychology.* New York: Wiley

Hollerman JR, Schultz W. 1998. Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci. USA* 1:304–9

Hopfield JJ. 1982. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA* 79:2554–58

Hoshi E, Shima K, Tanji J. 1998. Task-dependent selectivity of movement-related neuronal activity in the primate prefrontal cortex. *J. Neurophysiol.* 80:3392–97

Houck JC. 1995. *Models of Information in the Basal Ganglia*. Cambridge, MA: MIT Press

Janowsky JS, Shimamura AP, Kritchevsky M, Squire LR. 1989. Cognitive impairment following frontal lobe damage and its relevance to human amnesia. *Behav. Neurosci.* 103:548–60

Jones EG, Powell TPS. 1970. An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. *Brain* 93:793–820

Just MA, Carpenter PA. 1992. A capacity theory of comprehension: individual differences in working memory. *Psychol. Rev.* 99:122–49

Knight RT. 1984. Decreased response to novel stimuli after prefrontal lesions in man. *Clin. Neurophysiol.* 59:9–20

Knight RT. 1997. Distributed cortical network for visual attention. *J. Cogn. Neurosci.* 9:75–91

Koechlin E, Basso G, Pietrini P, Panzer S, Grafman J. 1999. The role of the anterior prefrontal cortex in human cognition. *Nature* 399:148–51

Kubota K, Niki H. 1971. Prefrontal cortical unit activity and delayed alternation performance in monkeys. *J. Neurophysiol.* 34:337–47

Laming DRJ. 1968. *Information Theory of Choice-Reaction Times*. London: Academic

Leon MI, Shadlen MN. 1999. Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron* 24:415–25

Levine B, Stuss DT, Milberg WP, Alexander MP, Schwartz M, Macdonald R. 1998. The effects of focal and diffuse brain damage on strategy application: evidence from focal lesions, traumatic brain injury and normal aging. *J. Int. Neuropsychol. Soc.* 4:247–64

Levine DS, Prueitt PS. 1989. Modeling some effects of frontal lobe damage-novelty and perseveration. *Neural Networks* 2:103–16

Lhermitte F. 1983. "Utilization behaviour" and its relation to lesions of the frontal lobes. *Brain* 106:237–55

Logan GD, Zbrodoff NJ, Fostey ARW. 1983.

Costs and benefits of strategy construction in a speeded discrimination task. *Mem. Cogn.* 11:485–93

London ED, Ernst M, Grant S, Bonson K, Weinstein A. 2000. Orbitofrontal cortex and human drug abuse: functional imaging. *Cereb. Cortex* 10:334–42

Lu MT, Preston JB, Strick PL. 1994. Interconnections between the prefrontal cortex and the premotor areas in the frontal lobe. *J. Comp. Neurol.* 341:375–92

Luria AR. 1969. Frontal lobe syndromes. In *Handbook of Clinical Neurology*, ed. PJ Vinken, GW Bruyn, pp. 725–57. New York: Elsevier

MacDonald AW, Cohen JD, Stenger VA, Carter CS. 2000. Dissociating the role of dorsolateral prefrontal cortex and anterior cingulate cortex in cognitive control. *Science* 288:1835–38

MacLeod CM. 1991. Half a century of research on the Stroop effect: an integrative review. *Psychol. Bull.* 109:163–203

McClelland JL, McNaughton BL, O'Reilly RC. 1995. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* 102:419–57

Melchitzky DS, Sesack SR, Pucak ML, Lewis DA. 1998. Synaptic targets of pyramidal neurons providing intrinsic horizontal connections in monkey prefrontal cortex. *J. Comp. Neurol.* 390:211–24

Miller EK. 1999. The prefrontal cortex: complex neural properties for complex behavior. *Neuron* 22:15–17

Miller EK. 2000. The neural basis of top-down control of visual attention in the prefrontal cortex. In *Attention and Performance*, ed. S Monsell, J Driver, 18:In press. Cambridge, MA: MIT Press

Miller EK, Desimone R. 1994. Parallel neuronal mechanisms for short-term memory. *Science* 263:520–22

Miller EK, Erickson CA, Desimone R. 1996. Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J. Neurosci.* 16:5154–67

Miller EK, Li L, Desimone R. 1993. Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *J. Neurosci.* 13:1460–78

Miller GA. 1956. The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63:81–97

Milner B. 1963. Effects of different brain lesions on card sorting. *Arch. Neurol.* 9:90

Mirenowicz J, Schultz W. 1994. Importance of unpredictability for reward responses in primate dopamine neurons. *J. Neurophysiol.* 72:1024–27

Mirenowicz J, Schultz W. 1996. Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* 379:449–51

Mishkin M. 1964. Perseveration of central sets after frontal lesions in monkeys. In *The Frontal Granular Cortex and Behavior*, ed. JM Warren, K Abert, pp. 219–41. New York: McGraw-Hill

Miyashita Y, Chang HS. 1988. Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature* 331:68–70

Montague PR, Dayan P, Sejnowski TJ. 1996. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16:1936–47

Mozer MC. 1991. *The Perception of Multiple Objects: A Connectionist Approach*. Cambridge, MA: MIT Press

Murray EA, Bussey TJ, Wise SP. 2000. Role of prefrontal cortex in a network for arbitrary visuomotor mapping. *Exp. Brain Res.* In press

Newell A, Simon HA. 1972. *Human Problem Solving*. Englewood Cliffs, NJ: Prentice Hall

Nichelli P, Grafman J, Pietrini P, Alway D, Carton JC, Miletich R. 1994. Brain activity in chess playing. *Nature* 369:191

Nystrom LE, Braver TS, Sabb FW, Delgado MR, Noll DC, Cohen JD. 2000. Working memory for letters, shapes and locations: fMRI evidence against stimulus-based

regional organization in human prefrontal cortex. *Neuroimage.* In press

O'Doherty J, Rolls ET, Francis S, Bowtell R, McGlone F, et al. 2000. Sensory-specific satiety-related olfactory activation of the human orbitofrontal cortex. *NeuroReport* 11:893–97

O'Reilly RC, Braver TS, Cohen JD. 1999. A biologically-based computational model of working memory. In *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*, ed. A Miyake, P Shah. New York: Cambridge Univ. Press

O'Reilly RC, McClelland JL. 1994. Hippocampal conjunctive coding, storage, and recall: avoiding a tradeoff. *Hippocampus* 4:661–82

O'Reilly RC, Munakata Y. 2000. *Computational Explorations in Cognitive Neuroscience: Understanding the Mind*. Cambridge: MIT Press

O Scalaidhe SP, Wilson FA, Goldman-Rakic PS. 1997. Areal segregation of face-processing neurons in prefrontal cortex. *Science* 278:1135–38

Owen AM, Downes JJ, Sahakian BJ, Polkey CE, Robbins TW. 1990. Planning and spatial working memory following frontal lobe lesions in man. *Neuropsychologia* 28:1021–34

Owen AM, Evans AC, Petrides M. 1996. Evidence for a two-stage model of spatial working memory processing within the lateral frontal cortex: a positron emission tomography study. *Cerebr. Cortex* 6:31–38

Pandya DN, Barnes CL. 1987. Architecture and connections of the frontal lobe. In *The Frontal Lobes Revisited*, ed. E Perecman, pp. 41–72. New York: IRBN

Pandya DN, Yeterian EH. 1990. Prefrontal cortex in relation to other cortical areas in rhesus monkey—architecture and connections. *Prog. Brain Res.* 85:63–94

Parker A, Gaffan D. 1998. Memory after frontal/temporal disconnection in monkeys: conditional and non-conditional tasks, unilateral and bilateral frontal lesions. *Neuropsychologia* 36:259–71

Passingham R. 1993. *The Frontal Lobes and Voluntary Action*. Oxford, UK: Oxford Univ. Press

Perret E. 1974. The left frontal lobe of man and the suppression of habitual responses in verbal categorical behaviour. *Neuropsychologia* 12:323–30

Petersen SE, van Mier H, Fiez JA, Raichle ME. 1998. The effects of practice on the functional anatomy of task performance. *Proc. Natl. Acad. Sci. USA* 95:853–60

Petrides M. 1982. Motor conditional associative-learning after selective prefrontal lesions in the monkey. *Behav. Brain Res.* 5:407–13

Petrides M. 1985. Deficits in non-spatial conditional associative learning after periarcuate lesions in the monkey. *Behav. Brain Res.* 16:95–101

Petrides M. 1990. Nonspatial conditional learning impaired in patients with unilateral frontal but not unilateral temporal lobe excisions. *Neuropsychologia* 28:137–49

Petrides M. 1996. Specialized systems for the processing of mnemonic information within the primate frontal cortex. *Philos. Trans. R. Soc. London Ser. B* 351:1455–61

Petrides M, Pandya DN. 1984. Projections to the frontal cortex from the posterior parietal region in the rhesus monkey. *J. Comp. Neurol.* 228:105–16

Petrides M, Pandya DN. 1999. Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *Eur. J. Neurosci.* 11:1011–36

Phaf RH, Van der Heiden AHC, Hudson PTW. 1990. SLAM: a connectionist model for attention in visual selection tasks. *Cogn. Psychol.* 22:273–341

Piaget J. 1954 (1937). *The Origins of Intelligence in Children*. New York: Basic Books

Porrino LJ, Crane AM, Goldman-Rakic PS. 1981. Direct and indirect pathways from the amygdala to the frontal lobe in rhesus monkeys. *J. Comp. Neurol.* 198:121–36

Posner MI, Snyder CRR. 1975. Attention and cognitive control. In *Information Processing*

and Cognition, ed. RL Solso. Hillsdale, NJ: Erlbaum

Prabhakaran V, Narayanan K, Zhao Z, Gabrieli JD. 2000. Integration of diverse information in working memory within the frontal lobe. *Nat. Neurosci.* 3:85–90

Price JL. 1999. Prefrontal cortical networks related to visceral function and mood. *Ann. NY Acad. Sci.* 877:383–96

Pucak ML, Levitt JB, Lund JS, Lewis DA. 1996. Patterns of intrinsic and associational circuitry in monkey prefrontal cortex. *J. Comp. Neurol.* 376:614–30

Quintana J, Fuster JM. 1992. Mnemonic and predictive functions of cortical neurons in a memory task. *NeuroReport* 3:721–24

Rainer G, Asaad WF, Miller EK. 1998a. Memory fields of neurons in the primate prefrontal cortex. *Proc. Natl. Acad. Sci. USA* 95:15008–13

Rainer G, Asaad WF, Miller EK. 1998b. Selective representation of relevant information by neurons in the primate prefrontal cortex. *Nature* 393:577–79

Rainer G, Rao SC, Miller EK. 1999. Prospective coding for objects in the primate prefrontal cortex. *J. Neurosci.* 19:5493–505

Rao SC, Rainer G, Miller EK. 1997. Integration of what and where in the primate prefrontal cortex. *Science* 276:821–24

Roberts AC, Wallis JD. 2000. Inhibitory control and affective processing in the prefrontal cortex: neuropsychological studies in the common marmoset. *Cerebr. Cortex* 10:252–62

Rolls ET. 2000. The orbitofrontal cortex and reward. *Cereb. Cortex* 10:284–94

Romo R, Brody CD, Hernandez A, Lemus L. 1999. Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature* 399:470–73

Rossi AF, Rotter PS, Desimone R, Ungerleider LG. 1999. Prefrontal lesions produce impairments in feature-cued attention. *Soc. Neurosci. Abstr.* 25:3

Rushworth MF, Nixon PD, Eacott MJ, Passingham RE. 1997. Ventral prefrontal cortex is not essential for working memory. *J. Neurosci.* 17:4829–38

Schacter DL. 1997. The cognitive neuroscience of memory: perspectives from neuroimaging research. *Philos. Trans. R. Soc. London Ser. B* 352:1689–95

Schmahmann JD, Pandya DN. 1997. Anatomic organization of the basilar pontine projections from prefrontal cortices in rhesus monkey. *J. Neurosci.* 17:438–58

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

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.* 13:900–13

Schultz W, Dickinson A. 2000. Neuronal coding of prediction errors. *Annu. Rev. Neurosci.* 23:473–500

Schultz W, Dayan P, Montague PR. 1997. A neural substrate of prediction and reward. *Science* 275:1593–99

Seltzer B, Pandya DN. 1989. Frontal lobe connections of the superior temporal sulcus in the rhesus monkey. *J. Comp. Neurol.* 281:97–113

Shadmehr R, Holcomb H. 1997. Neural correlates of motor memory consolidation. *Science* 277:821–24

Shallice T. 1982. Specific impairments of planning. *Philos. Trans. R. Soc. London Ser. B* 298:199–209

Shallice T. 1988. *From Neuropsychology to Mental Structure*. Cambridge, UK: Cambridge Univ. Press

Shallic T, Burgess P. 1996. The domain of supervisory processes and temporal organization of behaviour. *Philos. Trans. R. Soc. London Ser. B* 351:1405–11

Shiffrin RM, Schneider W. 1977. Controlled and automatic information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychol. Rev.* 84:127–90

Smith EE, Jonides J. 1999. Storage and executive processes in the frontal lobes. *Science* 283:1657–61

Squire LR. 1992. Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol. Rev.* 99:195–231

Stroop JR. 1935. Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18:643–62

Stuss DT, Benson DF. 1986. *The Frontal Lobes*. New York: Raven

Swedo SE, Shapiro MB, Grady CL, Cheslow DL, Leonard HL, et al. 1989. Cerebral glucose metabolism in childhood-onset OCD. *Arch. Gen. Psychiatr.* 46:518–23

Tomita H, Ohbayashi M, Nakahara K, Hasegawa I, Miyashita Y. 1999. Top-down signal from prefrontal cortex in executive control of memory retrieval. *Nature* 401:699–703

Treisman A, Gelade G. 1980. A feature integration theory of attention. *Cogn. Psychol.* 12:97–136

Tremblay L, Hollerman JR, Schultz W. 1998. Modifications of reward expectation-related neuronal activity during learning in primate striatum. *J. Neurophysiol.* 80:964–77

Tremblay L, Schultz W. 1999. Relative reward preference in primate orbitofrontal cortex. *Nature* 398:704–8

Tzelgov J, Henik A, Berger J. 1992. Controlling Stroop effects by manipulating expectations for color words. *Mem. Cogn.* 20:727–35

Usher M, Cohen JD. 1997. *Interference-based capacity limitations in active memory*. Presented at Abstr. Psychonom. Soc., Philadelphia

Van Hoesen GW, Pandya DN, Butters N. 1972. Cortical afferents to the entorhinal cortex of the rhesus monkey. *Science* 175:1471–73

Vendrell P, Junque C, Pujol J, Jurado MA, Molet J, Grafman J. 1995. The role of prefrontal regions in the Stroop task. *Neuropsychologia* 33:341–52

Wagner AD. 1999. Working memory contributions to human learning and remembering. *Neuron* 22:19–22

Wallis JD, Anderson KC, Miller EK. 2000. Neuronal representation of abstract rules in the orbital and lateral prefrontal cortices (PFC). *Soc. Neurosci. Abstr.* In press

Wang XJ. 1999. Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. *J. Neurosci.* 19:9587–603

Watanabe M. 1990. Prefrontal unit activity during associative learning in the monkey. *Exp. Brain Res.* 80:296–309

Watanabe M. 1992. Frontal units of the monkey coding the associative significance of visual and auditory stimuli. *Exp. Brain Res.* 89:233–47

Watanabe M. 1996. Reward expectancy in primate prefrontal neurons. *Nature* 382:629–32

White IM, Wise SP. 1999. Rule-dependent neuronal activity in the prefrontal cortex. *Exp. Brain Res.* 126:315–35

Wilson FAW, O Scalaidhe SP, Goldman-Rakic PS. 1993. Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science* 260:1955–58

Wise SP, Murray EA, Gerfen CR. 1996. The frontal-basal ganglia system in primates. *Crit. Rev. Neurobiol.* 10:317–56

Yamaguchi S, Knight RT. 1991. Anterior and posterior association cortex contributions to the somatosensory P300. *J. Neurosci.* 11:2039–54

Zipser D, Kehoe B, Littlewort G, Fuster J. 1993. A spiking network model of short-term active memory.13:3406–20

Zola-Morgan S, Squire LR. 1993. Neuroanatomy of memory. *Annu. Rev. Neurosci.* 16:547–63

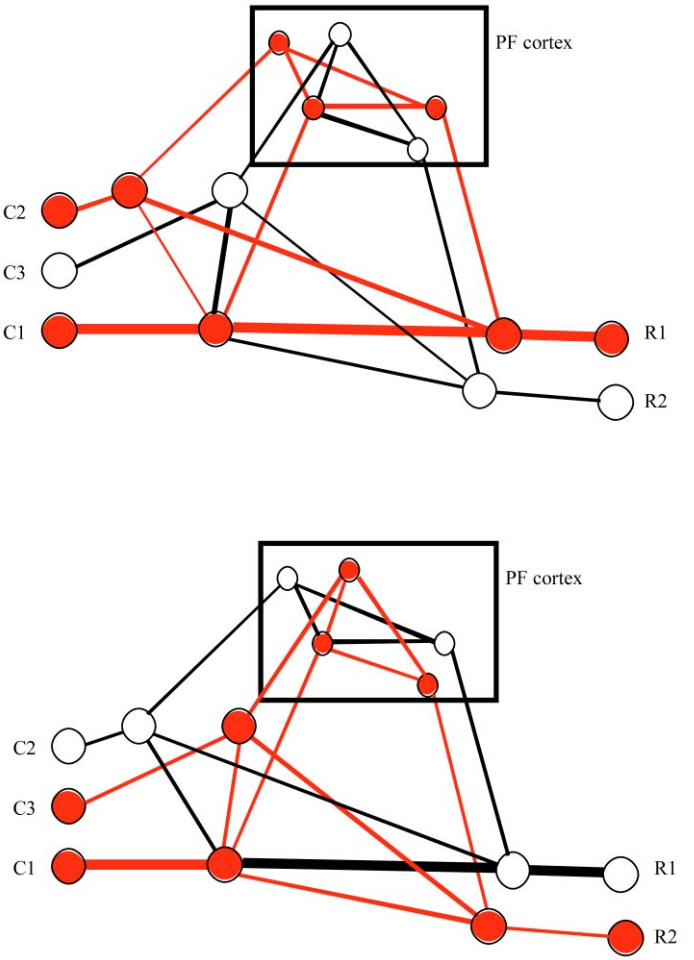


Figure 2 Schematic diagram illustrating our suggested role for the PF cortex in cognitive control. Shown are processing units representing cues such as sensory inputs, current motivational state, memories, etc. (C1, C2, and C3), and those representing two voluntary actions (e.g., "responses", R1 and R2). Also shown are internal or "hidden " units that represent more central stages of processing. The PF cortex is not heavily connected with primary sensory or motor cortices but instead connected with higher-level "association " and premotor cortices. Hence, we illustrate connections between the PFC and the hidden units. Reward signals foster the formation of a task model, a neural representation that reflects the learned associations between task-relevant information. A subset of the information (e.g., C1 and C2) can then evoke the entire model, including information about the appropriate response (e.g., R1). Excitatory signals from the PF cortex feeds back to other brain systems to enable task-relevant neural pathways. Thick lines indicate well-established pathways mediating a prepotent behavior. Red indicates active units or pathways.

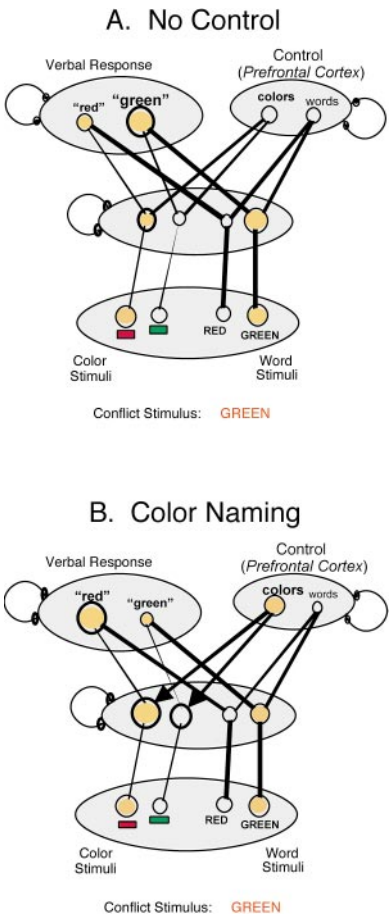


Figure 4 Schematic of the Stroop model. Circles represent processing units, corresponding to a population of neurons assumed to code a given piece of information. Lines represent connections between units, with heavier ones indicating stronger connections. Looped connections with small black circles indicate mutual inhibition among units within that layer (e.g., between the "red" and "green" output units). Adapted from Cohen, Dunbar & McClelland (1990).

A. No control. Activation of conflicting inputs in the two pathways produces a response associated with the word, due to the stronger connections in the word reading pathway.

B. Presentation of a conflict stimulus. The color unit is activated (indicated by the orange fill), representing the current intent to name the color. This passes activation to the intermediate units in the color naming pathway (indicated by arrows), which primes those units (indicated by larger size), and biases processing in favor of activity flowing along this pathway. This biasing effect favors activation of the response unit corresponding to the color input, even though the connection weights in this pathway are weaker than in the word pathway.

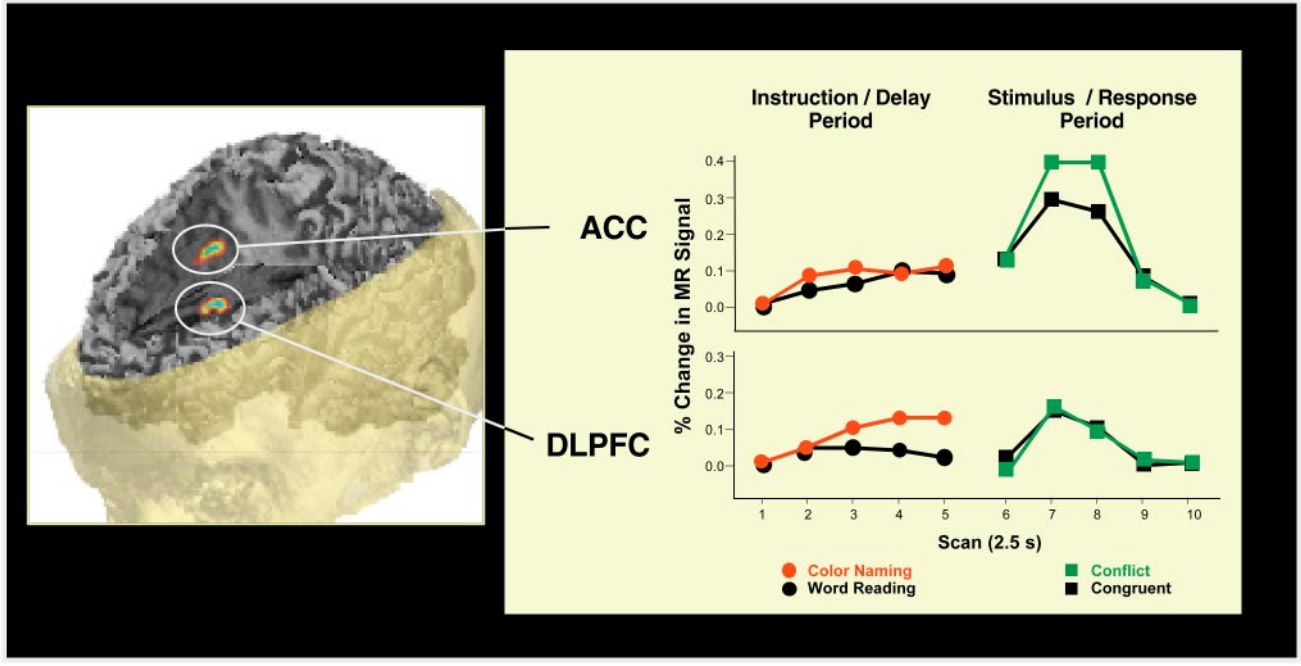


Figure 5 Time course of fMRI activity in dorsolateral prefrontal cortex (DLPFC) and anterior cingulate cortex (ACC) during two phases of a trial in the instructed Stroop task. During the instruction and preparatory period, there is significantly greater activation of DLPFC for color naming than word reading, but no difference in ACC. During the stimulus and response phase, there is greater activation of ACC for conflict than congruent stimuli, but no difference between these for DLPFC. Adapted from MacDonald et al (2000).