FROST: A Distributed Neurocomputational Model of Working Memory Maintenance

F. Gregory Ashby[1], Shawn W. Ell[2], Vivian V. Valentin[1], and Michael B. Casale[1]

Abstract

■ Many studies suggest that the sustained activation underlying working memory (WM) maintenance is mediated by a distributed network that includes the prefrontal cortex and other structures (e.g., posterior parietal cortex, thalamus, globus pallidus, and the caudate nucleus). A computational model of WM, called FROST (short for FROntal–Striatal–Thalamic), is proposed in which the representation of items and spatial positions is encoded in the lateral prefrontal cortex. During delay intervals, activation in these prefrontal cells is sustained via parallel, prefrontal cortical–thalamic loops. Activation reverberates in these loops because prefrontal cortical excitation of the head of the caudate nucleus leads to disinhibition of the thalamus (via the globus pallidus). FROST successfully accounts for a wide variety of WM data, including single-cell recording data and human behavioral data. ■

INTRODUCTION

Working memory (WM) is the ability to maintain and manipulate limited amounts of information during brief periods of cognitive activity. It is heavily used in a variety of cognitive tasks, and for this reason, it is often associated with planning and problem solving. An influential account was proposed by Baddeley (1986), who argued that WM is a multicomponent system composed of a central executive attentional system and subsidiary visuospatial and phonological storage systems (the visuospatial sketch pad and phonological loop).

In addition to countless cognitive studies of WM (for a recent review, see Cowan, 2000), many studies have also investigated the neural basis of WM. These include single-cell recordings (e.g., Fuster & Alexander, 1971), neuroimaging studies (e.g., Carpenter, Just, & Reichle, 2000), and neurocomputational modeling (e.g., Durstewitz, Seamans, & Sejnowski, 2000a). This explosion of new knowledge about the neural basis of WM has inspired the development of many neurobiologically based computational models of this vital cognitive function.

This article proposes and tests a new computational model of WM constructed from components that correspond to groups of similar cells in specific brain regions (e.g., cortical columns or hypercolumns). One of the novel contributions of the model is that it is simple and flexible enough to account for a wide variety of data. This article focuses on testing the model against single-unit recording data and classic behavioral data, but the same model can also be used to account for neuroimaging data (Ashby & Valentin, in press). Other models of WM have simulated single-unit recording and neuroimaging data (Deco, Rolls, & Horwitz, 2004; Tagamets & Horwitz, 2000), but none of these have been tested against human behavioral data.

Another important strength of the present model is its focus on the role of a frontal–striatal–thalamic network in WM maintenance. This is among the first models to implement the hypothesis that *all* brain areas within this network are involved in the maintenance of information. Almost all other anatomically similar neurocomputational models assign the maintenance function to the lateral prefrontal cortex (pFC) and propose that subcortical structures are involved only in the updating of information (e.g., Frank, Loughry, & O'Reilly, 2001).

The present model does not attempt to explain how information in WM is manipulated; it only aims to clarify the most basic mechanism of WM—maintenance. Toward this end, at the single-cell level, we focus on the sustained changes in single-cell firing rates that are commonly observed in delayed-response tasks. This is also rather unique because many neurocomputational models focus on the on-line manipulation of information involved in more complicated extensions of the delayed-response task (Frank et al., 2001; Monchi & Taylor, 1999; O'Reilly, Braver, & Cohen, 1999; Beiser & Houk, 1998).

Despite the limited scope of the resulting model, there is still a huge amount of behavioral and neuro-

[1]University of California, Santa Barbara, [2]University of California, Berkeley

© 2005 Massachusetts Institute of Technology

physiological data that can be used to test its predictions. Once the model is validated against these data, it will be possible to generalize it to the many other more complex manipulation functions and WM-related tasks. By focusing on simple maintenance phenomena, however, we believe the present WM maintenance model is the most comprehensive interpretation of the neurophysiological evidence from single-unit recording, neuroimaging, lesion, and patient studies.

We begin by reviewing the neural basis of WM and then introduce our basic model of WM maintenance and test it against single-cell recording data. Next, we consider the ability of the model to account for the most classic of all WM phenomena—namely, its limited capacity, which was immortalized by Miller's (1956) famous ''magical number seven, plus or minus two.'' We also consider the ability of the model to account for the effects of attention on WM span, as well as individual differences. We close with a discussion of the relationship of the model we propose to other neurobiologically plausible models of WM and with some general conclusions.

Neural Circuits of Working Memory

There is overwhelming evidence linking WM to the pFC (Jonides, Smith, et al., 1993; Fuster, 1989; Goldman-Rakic, 1987). Even so, the question remains of how the pFC is able to maintain a sensory representation over an extended period. One popular proposal has been that delay-related activity is maintained in the lateral pFC via recurrent, excitatory, cortico-cortical connections within the pFC (Durstewitz, Seamans, & Sejnowski, 2000b). However, many well-established results challenge the view that WM is mediated entirely within the pFC. For example, a number of similar studies have been reported in which monkeys performed some sort of spatial delayed-response task while single-cell recordings were made in one of a variety of different brain regions. In a typical application, a monkey watches through a window while an experimenter hides a food reward in a covered location. The window is covered for some seconds and then opened, at which time the monkey is free to retrieve the reward. Using a paradigm such as this, a number of studies have found cells that show sustained activity during the delay period not only in the pFC (delayed response: Fuster, 1973), but also in the posterior parietal cortex (PPC; delayed match-to-sample: Constantinidis & Steinmetz, 1996), the medial dorsal nucleus of the thalamus (MDN; delayed response: Fuster & Alexander, 1973), the caudate nucleus (CD; delayed saccade: Schultz & Romo, 1992; Hikosaka, Sakamoto, & Sadanari, 1989), and the globus pallidus (GP; memory-guided sequential pointing: Mushiake & Strick, 1995). To anticipate, we propose that the sustained activation in these structures all contribute to WM maintenance.

Figure 1 shows typical examples of the sustained activity that has been found in each of these areas. The first classic recordings from the pFC provided an influential clue that the pFC helps mediate WM (e.g., Fuster & Alexander, 1971). In subsequent years, delay-related activity was found in many other brain areas. Figure 1 shows that a number of different profiles have been found, but in each case the cells in question change their firing rate during the delay period. In most cases, the cells in Figure 1 show an increased activation, relative to baseline, during the entire delay period. The

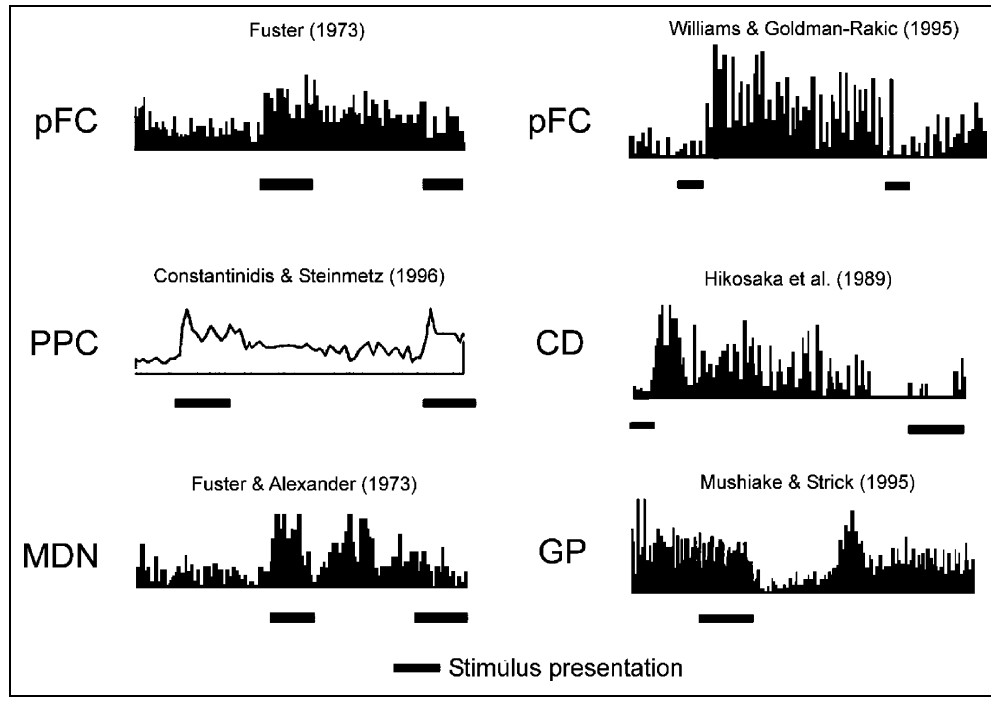

Figure 1. Single-cell recording data collected from monkeys in a variety of delayed-response tasks.

exceptions are the cells in the GP, whose firing rates decrease during the delay period. Note that there are also regional differences in which event triggers the initial change in firing rate. Cells in the PPC and the MDN, and some cells in the pFC increase their firing rate at the moment the target is first presented. In contrast, cells in the CD and the GP, and other pFC cells do not respond to the initial presentation of the target, but instead, their firing rate changes at the beginning of the delay period.

Finding cells in the thalamus and basal ganglia with sustained delay-related activity that is correlated with the activity of ''working memory units'' in pFC challenges the view that WM maintenance function is largely mediated within the cortex, but it is hardly definitive. Even so, a variety of other evidence supports the hypothesis that the thalamus and basal ganglia play critical roles in a widely distributed WM circuit. First, lesions to the MDN have been reported to impair WM in both animals and humans (Van der Werf, Witter, Uylings, & Jolles, 2000). Second, there are several reports that patients with basal ganglia lesions have WM deficits (Janahashi et al., 2002). Third, a number of studies have reported WM deficits in early Parkinson's disease patients, especially in spatial WM tasks (Lewis et al., 2002; Postle, Jonides, Smith, Corkin, & Growdon, 1997). Finally, increased thalamic and basal ganglia activation has been reported in some neuroimaging studies (Callicott et al., 1999; Jonides, Schumacher, et al., 1997). Any one of these studies, by itself, is hardly definitive, but together they provide a challenge to the view that WM is exclusively a cortical phenomenon.

An alternative hypothesis, pursued in this article, is that the pFC is part of a distributed reverberating WM circuit that includes both cortical and subcortical structures (Goldman-Rakic, 1995). Alexander, DeLong, and Strick (1986) identified a number of parallel loops that link the cortex to the basal ganglia and thalamus. In each case, a group of cells in the cortex projects directly to the neostriatum (comprising the putamen and the CD), which then projects to the (internal segment of the) GP, which projects to the thalamus, which finally projects back to the cortex. Since this seminal discovery, such cortical–striatal–pallidal–thalamic loops have been the focus of much theorizing. For example, it has been proposed that they play a major role in category learning (Ashby, Alfonso-Reese, Turken, & Waldron, 1998), creative problem solving (Ashby, Isen, & Turken, 1999), motor performance (Strick, Dum, & Pickard, 1995), selective attention (Posner & Petersen, 1990), skill learning (Gabrieli, 1995), and the learning of serial order (Keele, Ivry, Mayr, Hazeltine, & Heuer, 2003; Beiser & Houk, 1998).

The lateral pFC is also part of such a loop, with efferent projections to the CD and afferent projections originating in the MDN. Thus, every subcortical region illustrated in Figure 1 lies on the same cortical–striatal–pallidal–thalamic loop.

A Neurocomputational Model of Working Memory Maintenance

Figure 2 sketches a neural circuit that is compatible with the single-cell recording data shown in Figure 1. The

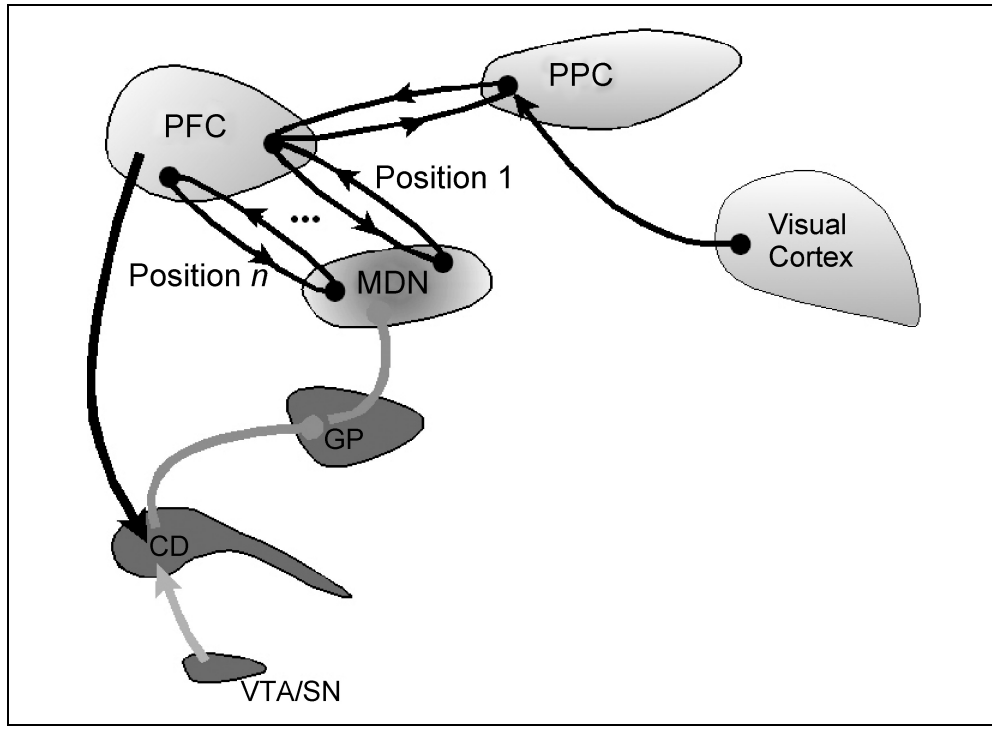

Figure 2. A model of spatial WM.

spatial location of the food reward is encoded in some small set of cells in the PPC. When the animal can see the food, these cells are driven primarily by bottom-up input from lower levels of the visual cortex. These units in the PPC project into WM units in the lateral pFC. The goal of the WM circuit is to keep the pFC cells active during the delay period when the animal can no longer see the reward.

This goal is attained with the help of two excitatory, reverberating loops—one between the PPC and the lateral pFC, and one between the pFC and the MDN. More specifically, the WM cells in the lateral pFC send an excitatory signal to the PPC, which sends a recurrent excitatory signal back to the same pFC neurons. At the same time, a similar cortical–thalamic loop helps to sustain activity in the pFC WM cells.

An inhibitory input to the MDN from the GP disrupts processing in this loop. Spontaneous activity in the GABAergic pallidal cells is high (Wilson, 1990), so without impeding the action of the GP, the GP will prevent the thalamus from closing the cortical–thalamic loop, and the information will quickly decay from WM. The solution is for the CD to inhibit the GP (because the CD cells are GABAergic), thereby preventing the GP from disrupting the reverberating cortical–thalamic WM loop. Cells in the CD have a low spontaneous firing rate and are characterized by bursts of activity when stimulated by the cortex (Bennett & Wilson, 2000). A direct excitatory projection from the pFC causes the CD cells to fire, which inhibit the GP, and ultimately, the cortical–thalamic loop stays active.

The single-cell recording data shown in Figure 1 indicate that the firing rate in the CD increases and the firing rate in the GP decreases only after the visual stimulus is withdrawn. Thus, we propose that the pFC input driving the CD does not originate from the pFC cells that are activated by sensory input and maintain the representation of the stimulus, but rather from pFC cells that become active only when the stimulus disappears with the start of the delay period. It is exactly at this point that attentional demands also increase, and because we hypothesize that the Figure 2 circuit will be important in almost all executive tasks, we interpret this delay-activated pFC unit as providing an attentional signal. As Figure 1 shows, several single-unit recording studies have reported pFC cells whose firing increases at the beginning of the delay rather than during stimulus presentation (Williams & Goldman-Rakic, 1995; Fuster, 1973; Fuster & Alexander, 1971).

Clearly, the Figure 2 model oversimplifies the structure of the cortex. A more realistic model of the pFC is illustrated in Figure 3, which shows the well-known six cortical layers. The best evidence indicates that the pFC–PPC loop shown in Figure 2 projects in and out of Layers 2 and 3, the projection to the CD is from Layer 5, the MDN projection is from Layer 6, and the MDN projects into Layer 4 (e.g., Heimer, 1995). However,

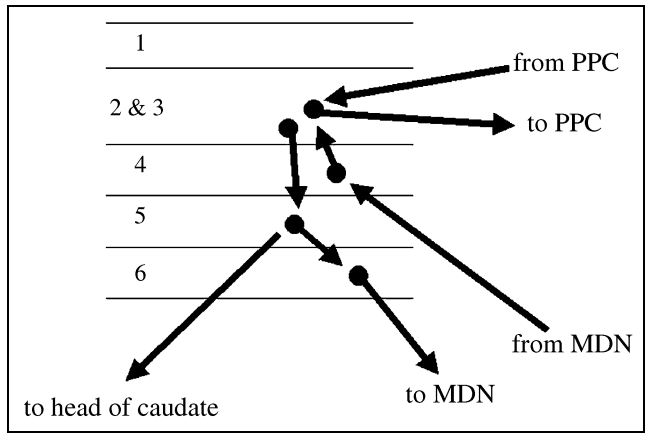

Figure 3. The six layers of the lateral pFC and some excitatory interneurons.

many excitatory interneurons connect these layers. For our purposes, the most important of these are the cells that project from Layer 4 to Layers 2 and 3, and the series of cells that project from Layers 2 and 3 to Layer 6 via Layer 5 (Heimer, 1995). Note that the interneurons that connect the various layers close the proposed cortical–thalamic reverberating circuit. They also provide a mechanism via which cortical input from structures regulating executive attention could drive the pFC input to the CD.

The model described in Figures 2 and 3 assumes WM is mediated by parallel, frontal cortical–striatal–thalamic loops. As a result, we refer to it as the FROST (FROntal–Striatal–Thalamic) model of WM. FROST is at least qualitatively consistent with the single-cell recording data shown in Figure 1 as it assumes that WM is mediated by a widely distributed network. Figure 3 illustrates another important prediction of FROST. Specifically, FROST predicts that multi-unit recordings in the lateral pFC during a WM task should show reverberating (or correlated) activity among units in different pFC layers. As mentioned above, there have been many proposals that delay-related activity is maintained in the lateral pFC via recurrent, excitatory, cortico-cortical connections solely within the pFC (e.g., Tagamets & Horwitz, 2000). FROST therefore offers a possible resolution to the debate between proponents of this position and those who argue for more extended, macro-network models of WM. In FROST, a macro-circuit that includes the PPC, the MDN, and the basal ganglia drives a micro-circuit within the lateral pFC.

Of course, it is one thing to claim that FROST is qualitatively consistent with observed single-cell recording data and quite another to show that it can account *quantitatively* for such data. To derive quantitative predictions from FROST, we developed a computational model that is biophysically complex enough to be consistent with the neural network shown in Figure 2, while still able to account for human behavior.

The computational version of FROST assumes that WM is mediated by the functional interconnections between the brain regions shown in Figure 2, so these interconnections are captured in the mathematical equations that model the network. In addition, FROST models two key biophysical properties that are shared by all neurons: (1) saturation—every neuron has a maximum firing rate, and (2) decay—if all inputs to a neuron cease then the activation in that cell will decay to some baseline firing level. The Methods section develops a set of differential equations that accomplishes these goals—that is, a different equation describes activation in each of the Figure 2 brain regions, these equations model the functional interconnections shown in Figure 2, and the equations also model the biophysical properties of saturation and decay.[1] Each of these equations models a functional unit that corresponds to a group of similar cells in a specific brain region (e.g., cortical columns or hypercolumns, caudate domains). The principle advantage of this approach, and one of the truly unique contributions of the FROST model of WM, is its extreme flexibility with respect to the types of data that it can be tested against. For example, in addition to accounting simultaneously for single-cell recording data and human behavioral data (some of these applications are described below), we have also had some success in fitting FROST to functional neuroimaging data (i.e., to the fMRI BOLD signal; Ashby & Valentin, in press). To our knowledge, FROST is the only current WM model that can be tested against such diverse types of data.

RESULTS

Simulation of Single-Cell Recording Data

The basic FROST model (see Equations 1–6 in the Methods) describes continuous changes (over time) in activation in each brain region shown in Figure 2. These can be converted into spike trains using a standard integrate-and-fire model.[2] With this addition, we might ask whether the units in the model behave in a qualitatively similar fashion as real neurons. To answer this question, we compared the behavior of the FROST units to the monkey single-cell recording data shown in Figure 1. Given the significant intercell variability, our goal was not to provide a precise quantitative fit to these data. Instead, our goal was simply to ask whether the units in FROST show delay-related response profiles that are similar to those seen in real cells.

With this goal in mind, we crudely searched through different numerical values of the unknown constants in Equations 1–6 for values that produced reasonable behavior.[3] As is clear from the Methods section, most of these parameters are measures of synaptic strength. Results are shown in Figure 4. Note that FROST successfully displays sustained activity in the PPC, the pFC, the MDN, and the CD, and it also displays a sustained depression in the GP. To our knowledge, FROST is the only existing model able to account for the CD and GP data shown in Figure 4 (e.g., most other models with subcortical components predict phasic changes in CD and GP activation). Note also that FROST successfully models the two different types of pFC cells shown in Figure 1. Because we expect learning (e.g., changes in synaptic strength) to be minimally important in the kinds of WM tasks considered in this article, all other applications of FROST reported in this article used the same numerical constants that produced the simulated single-cell data shown in Figure 4. These values are given in Table 1.

Model Fits to Behavioral Data: The Magical Number 4, 5, 6, or 7

As mentioned above, a major advantage of FROST, compared to many other computational models of WM, is that it can be tested against a wide variety of data types. Figure 4 shows that it can account for single-cell recording data. Our next step was to examine its ability to account for human behavioral data. As currently specified, FROST makes no behavioral predictions. Instead, it only predicts neural activations in specific brain regions in certain WM tasks. However, the model could be used to generate behavioral predictions by making assumptions about how activation in a certain region (e.g., pFC) could lead to a behavioral response. This is the approach we take in the remainder of this article.

To this point, FROST has been developed primarily in the context of spatial WM. However, in addition to the PPC, all other sensory association areas also project to the pFC (Heimer, 1995), so there is no reason that other types of WM could not operate in much the same way. For example, Figure 5 shows how FROST might work in a more generic WM task (e.g., digit span). Presentation of a stimulus will cause bottom-up activation of some sensory association area that encodes a high-level representation of the stimulus. This region, in turn, will be reciprocally connected to a WM unit in the pFC, and from there the circuitry is essentially identical to the version of FROST shown in Figure 2. It is important to note, however, that the specific pFC WM cells activated in a spatial delayed-response task and in say, a digit span task will be different. This is because there are separate topographic pFC representations of activity in the PPC and the auditory cortex. Posterior cortical units of the type shown in Figure 5 could be interpreted as encoding the long-term memory of the object or event reverberating in WM. Many theorists have argued that WM follows many of the same rules as long-term memory (e.g., Nairne, 1991). FROST is consistent with many of these arguments because it postulates a mechanism via which long-term memories are loaded into WM.

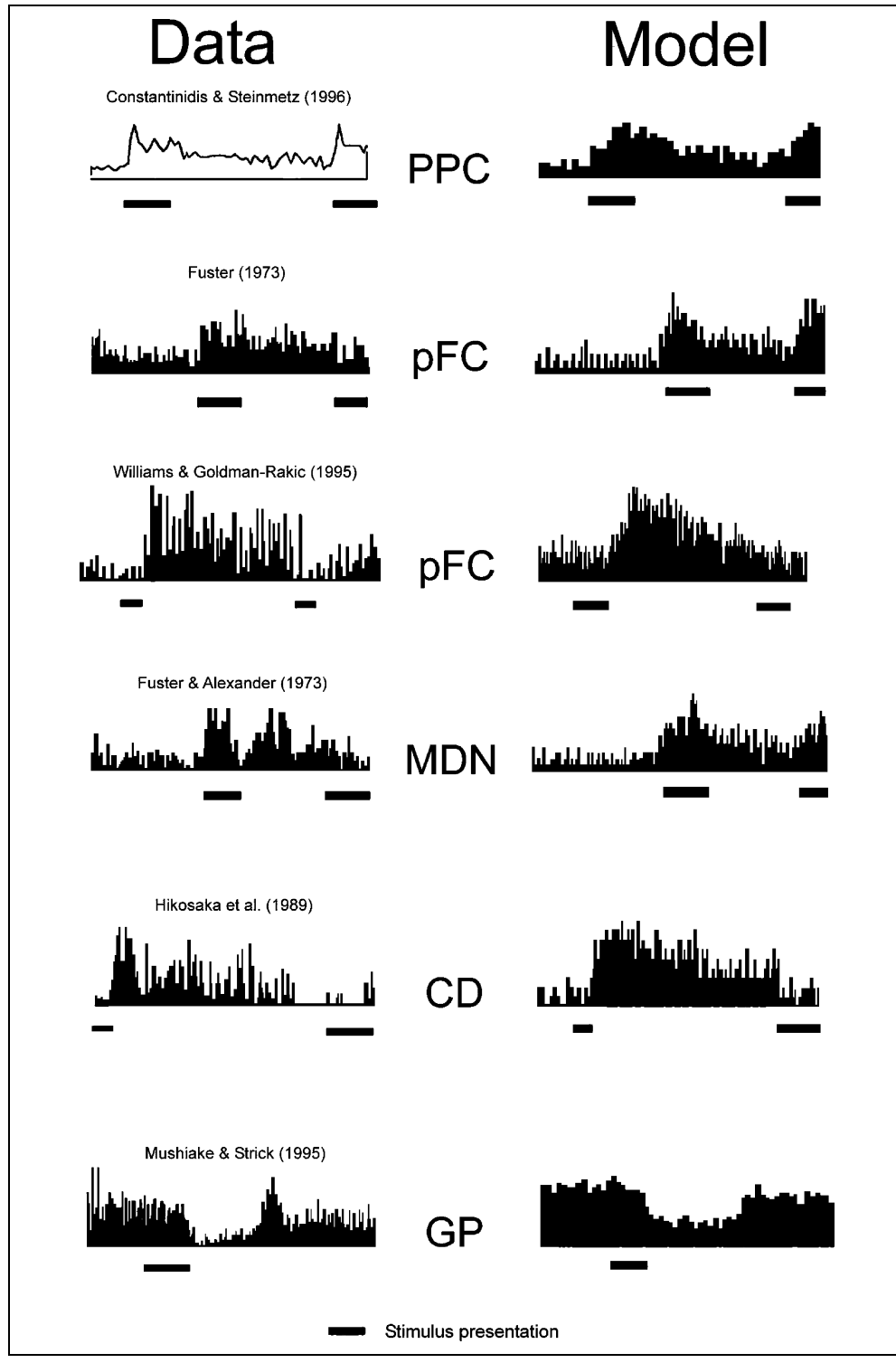

Figure 4. Single-cell recording data collected from monkeys in a variety of delayed-response tasks (left column) and single-cell recording data simulated from FROST.

Using this more general version of FROST allows us to explore many other phenomena. Perhaps the most widely known and extensively replicated phenomenon of WM is its extremely limited storage capacity. This feature was made famous by Miller's (1956) claim that the limit is ''seven, plus or minus two.'' Although a fixed limit is widely accepted, debate continues as to whether the limit is as high as seven, or as low as four (Cowan, 2000). Figure 6 summarizes data from three very different WM experiments—classic memory span (Guilford & Dallenbach, 1925), span of apprehension or object enumeration (Mandler & Shebo, 1982), and absolute identification of pure tones (Pollack, 1952).

In a classic memory span task, in which a participant is read a list of unrelated items and then immediately asked to recall that list, a natural application of FROST

Table 1. Parameter Estimates Used to Simulate Single-cell Recording Data Shown in Figure 4

Equation	*Parameter*	*Value*
1	α_P	0.4
	β_P	0.001726
	γ_P	0.01398
2	α_F	0.0055
	β_F	0.0030
	δ_F	0.0025
3	α_T	6.2421
	β_T	22.161
	γ_T	60.087
4	α_G	2.2414
	β_G	31.445
	G_B	0.5
5	α_C	1.882
	β_C	6.2889
	γ_C	99.914
6	α_A	0.0015
	β_A	0.0015
	γ_A	0.0050
Integrate-and-fire model	σ^2	0.8
	V_0	1

assumes that each successive item activates a unique unit in some posterior sensory association area (assuming all items are different), and that each of these units projects to a separate pFC WM unit. FROST sets no strict limit on the number of pFC WM units that can be active simultaneously. However, activation in these units will decay during the time when the observer is retrieving other items from WM. Thus, FROST predicts an upper limit on WM span because of decay to later items while the observer is reporting the earlier items. In addition, lateral inhibition between WM units will increase with memory span. Other models have attributed the magic number to such decay and interference phenomena rather than to a fixed limit on the number of storage slots per se (e.g., Baddeley, 1992).

To fit FROST to these data, we made the following extra assumptions. First, we assumed a separate WM loop for each item in the study list. Second, we assumed that study time was long enough to ensure perfect encoding. We also assumed that retrieval proceeded in a serial fashion at the rate of two items per second. Finally, we assumed that an item is correctly retrieved if the activation in the pFC is greater than some threshold at the time of the retrieval (see Equation 7 in the Methods section). For the basic FROST circuit, we used the same parameter estimates that were estimated in the Figure 4 simulations of the single-cell recording data. The resulting model fits are also shown in Figure 6. As can be seen, FROST provides an excellent account of these behavioral data, accounting for 99.7% of the variance in the mean memory span data.

Many other variations on this same experimental task have been reported that provide additional challenges for FROST. We briefly consider just one of these. There is a large literature showing that individuals differ substantially in their WM span (e.g., Miyake, 2001). One such study, which was reported by Cowan, Nugent, Elliott, Ponomarev, and Saults (1999), is summarized in Figure 7. In a pretest, the memory span of each participant was estimated by determining the largest number of items they could recall without error. This resulted in four groups, with spans that ranged from six to nine. The performance of each group in a subsequent test (with spoken digits) is shown in the top of Figure 7 (the ''attend'' data). In a second condition, these same participants performed either a picture-naming or rhyme-matching task at the same time that a sequence of digit lists was spoken. Participants were instructed to ignore the digits, although every once in awhile, they were asked to recall these unattended lists. These ''ignore'' data are shown in the bottom half of Figure 7.

To apply FROST to these data, it is natural to assume that the primary difference between the attend and ignore conditions will be in the strength of the attentional signal (determined by the parameter Δ from Equation 6 in the Methods section). Modeling individual differences, however, is not so straightforward. There are a number of possibilities. High-span individuals might have less noise, an increased gain, slower decay, or less lateral inhibition. In FROST, any of these would increase WM span. Of course, it is quite likely that there are multiple causes of individual differences in WM span. Certainly, a complete model of such differences is beyond the scope of this article.

One clue for how to model individual differences comes from experiments in which participants must later recall information that was previously irrelevant. For example, in one popular procedure, participants memorize and are tested on a list of items, and then later, on a second list. As might be expected, high-span individuals perform better than low-span individuals on both of these tests. Following the second test, however, participants are then retested on the first list. Somewhat surprisingly, low-span individuals outperform high spans on this retest (Engle, 2001). One popular hypothesis is that to perform well on the second list, participants must inhibit the first list, and that this type of inhibition is easier for high-span individuals. Inhibiting the first list helps high spans memorize the second list, but it hurts

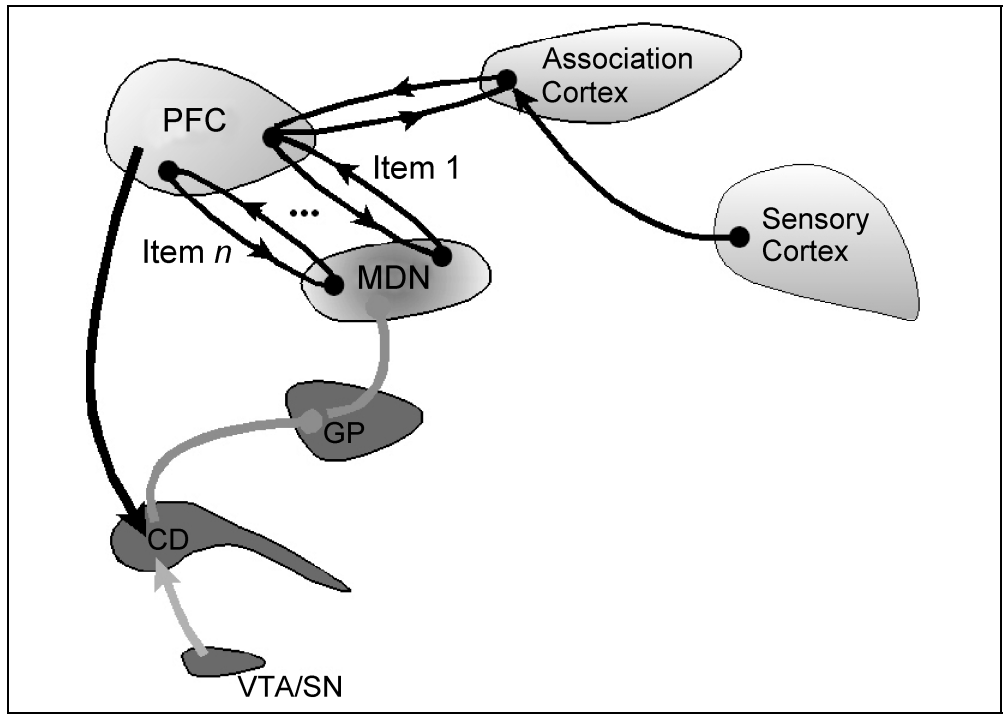

Figure 5. A general model of WM.

their performance on the first list retest (Conway & Engle, 1994). A wide variety of evidence now supports this inhibition hypothesis. The Cowan et al. (1999) Figure 7 data are from a similar paradigm—during a picture-naming or rhyme-matching task, a list of digits was unexpectedly cued for recall. The inhibition hypothesis predicts that high spans should be better able to ignore the digits while performing the primary task (i.e., picture-naming or rhyme-matching), which should impair their recall of the digits. The neural mechanisms mediating this hypothesized inhibition have not been articulated. In FROST, a natural way to instantiate the inhibition hypothesis is to assume that high-span individuals have increased attentional control.

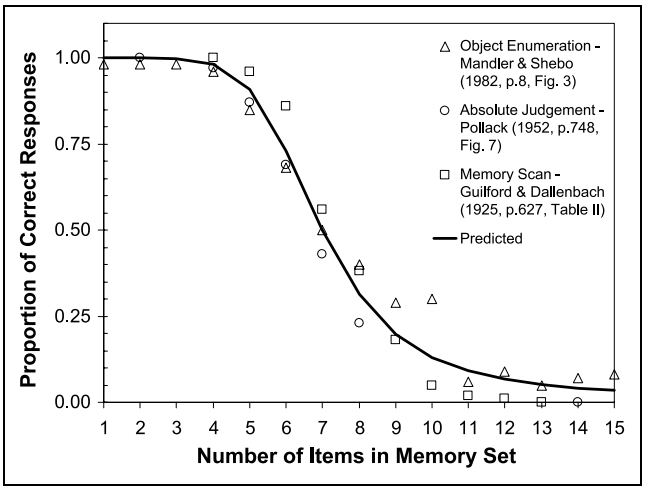

Figure 6. The proportion of correct responses as a function of memory load in three different working memory tasks. Pollack's data have been corrected. Also shown are fits of FROST (% of variance accounted for = 99.7%, SSE = 0.008; parameter estimates: threshold = 0.06, σ_ϵ = 0.027, γ_F = 0.002).

Because of massive cortical–striatal convergence, the resolution of medium spiny cells in the CD is much poorer than in the cortical cells from which they receive input (Wilson, 1995). For this reason, a single attentional signal might suffice when memorizing a list of similar items. However, if the items are dissimilar, then they will be encoded by spatially separated WM units, which will be innervated by separate thalamic units, and therefore, will require separate attentional signals. In the Cowan et al. (1999) experiment in which high-span individuals generally recalled less of the ignored material, very different items were used for the primary task (pictures) and the ignored list (spoken digits). Thus, in such paradigms, the two lists would each have their own attentional signal. Now assume that high-span individuals have improved attentional control. During the attend conditions, the magnitude of the attentional signal would be higher for the high-span than for the low-span groups, leading to better high-span performance. However, the high spans would also have better control of the attentional signal for the ignore items than the low spans. Thus, if participants were rewarded for ignoring the ignore items, then this model of individual differences would predict that the high spans would have a weaker attentional signal for the ignore items than the low spans, thereby paradoxically guaranteeing that the low-span individuals would have better memory

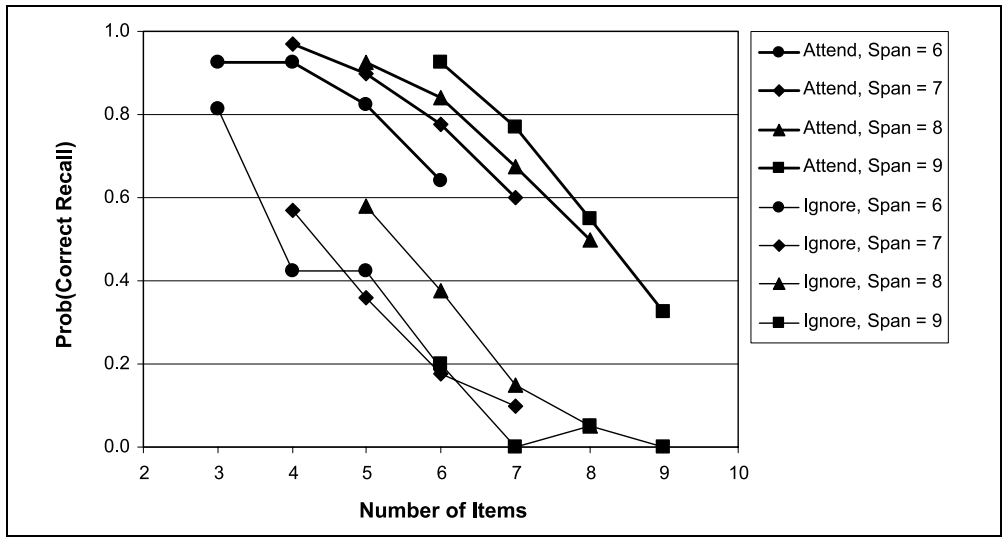

Figure 7. Recall after attended and ignored speech for four different groups of participants who differed in their WM capacity (from Cowan et al., 1999).

for the ignored items. In the case of the Cowan et al. experiment, this latter prediction must be weakened somewhat because participants were not rewarded for ignoring the ignore items.

For these reasons, we modeled individual differences in span by assuming that the parameter Δ in Equation 6 (see Methods) increases with WM span under conditions of high attention and decreases with WM span under conditions of low attention. To model the 32 data points of Cowan et al. (1999), therefore, FROST had 10 free parameters (a threshold, a noise variance, and a Δ for each curve in Figure 7, with the provision that Δ increases with span in the attend conditions and decreases with span in the ignore conditions). The results are shown in Figure 8. Note that FROST accurately captures the effects of attention and individual differences on WM span (accounting for 95.7% of the variance in the Cowan et al. data).

Lesions to FROST

FROST predicts that WM ability is distributed across a variety of neural structures. Nevertheless, the model predicts that the only structures strictly necessary for WM maintenance are the pFC WM units. FROST predicts that a variety of other neural structures (i.e., those shown in Figure 2) should facilitate WM performance, but lesions to these structures should still preserve some rudimentary WM ability. For example, FROST predicts that subcortical lesions to the head of the CD, the GP, or the MDN of the thalamus should reduce (or eliminate) the efficacy of the cortical–thalamic loops, but some pFC activity should still persist during the delay period of a delayed-response task because of reverberating activity in the cortical–cortical loops. Disrupting activity in the PPC during the delay period should have similar effects—that is, delay-related activity in the pFC will be

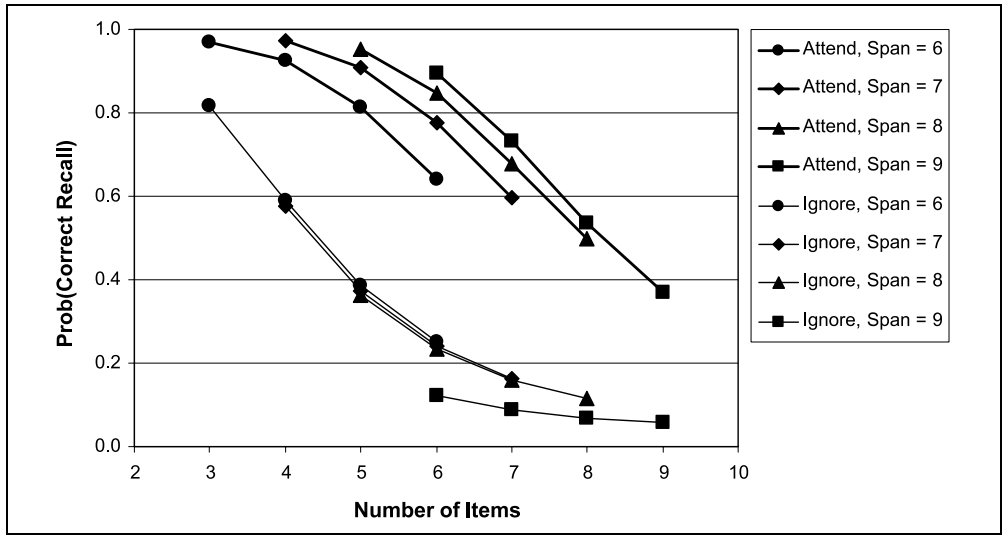

Figure 8. Fits of FROST to the data of Cowan et al. (1999) [% of variance accounted for = 95.7%, SSE = 0.139; parameter estimates: threshold = 0.048, σ_ϵ = 0.026, Δ in attend condition = 0.120 (span 6), 0.213 (span 7), 0.590 (span 8), 4.00 (span 9), Δ in ignore condition = 0.040 (span 6), 0.038 (span 7), 0.037 (span 8), 0.019 (span 9)].

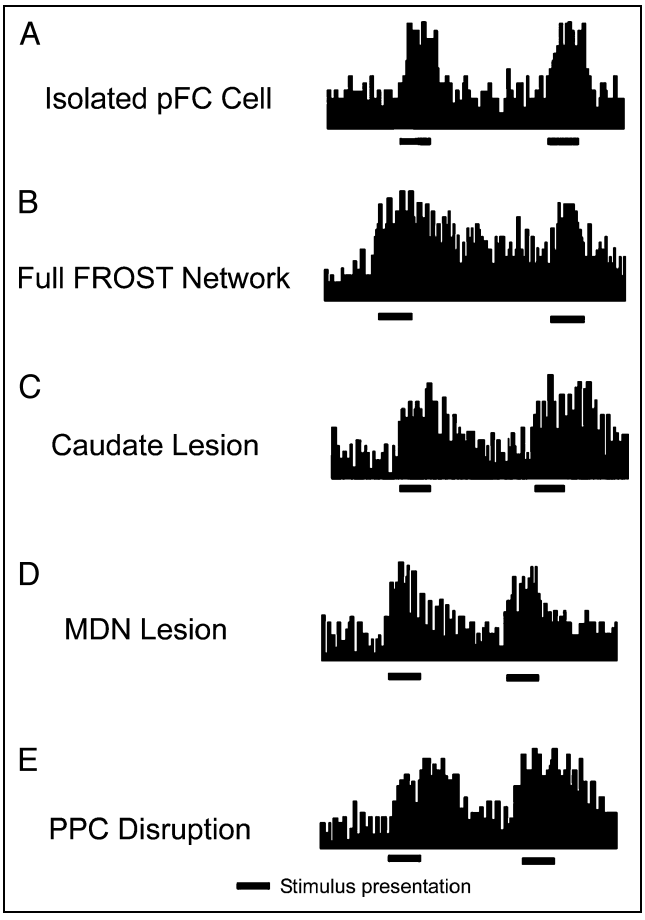

Figure 9. Activity of a pFC WM unit in FROST when (A) the cell only receives input from the PPC and sends no outputs, (B) the full network is intact, (C) the head of the CD has been lesioned, (D) the MDN of the thalamus has been lesioned, and (E) activity in the PPC was disrupted after the working memory was initiated.

reduced but not eliminated, because of reverberating activity in the cortical–thalamic loops.

Figure 9 illustrates these robust predictions of FROST. Each panel in the figure shows simulated single-cell firing data from a pFC WM unit during a delayed-response task. For reference, Figure 9A shows the performance of an isolated pFC cell. This is a cell that is completely disconnected from the network. Its only input is from the visual cortex, and it has no outputs. Activation increases during stimulus presentation, but the cell shows no sustained activity during the delay. In contrast, Figure 9B shows the performance of this same cell when the full network is intact. Figure 9C shows the performance when the CD has been lesioned, and Figure 9D shows the performance when the MDN has been lesioned. Note that, as anticipated, both lesions attenuate the delay related activity in the pFC WM units, but neither lesion completely abolishes this activity (as compared to Figure 9A).

The posterior cortical regions of FROST are necessary for the initial perceptual representation of the position or item to be held in memory. However, after the WM loops are initialized, these posterior cortical regions facilitate, but are not necessary for, WM maintenance. For example, the bottom panel of Figure 9 shows the effects on pFC activation of disrupting activity in the PPC after the delay period begins. Note that, as with thalamic lesions, such disruptions impair, but do not abolish, pFC delay-related activity. Some experimental data support this prediction of FROST. For example, Oliveri et al. (2001) showed that disrupting posterior cortical activation in humans (via transcranial magnetic stimulation) during the delay period of a WM task had only minor effects on performance. Similarly, several studies have demonstrated that pFC activity is robust to interference from distractor stimuli (Constantinidis & Steinmetz, 1996; Miller, Erickson, & Desimone, 1996).

Figure 10 shows that the reduced pFC activations shown in Figure 9 lead to impaired performance in classic behavioral tests of WM span. Note that FROST predicts that MDN or basal ganglia lesions should have little effect on performance when the memory load is small, but with higher memory loads such lesions should produce substantial deficits.

In summary, FROST predicts that lesions to the CD or the MDN should impair, but not completely abolish, WM. Of course, most lesions that occur because of stroke or surgery would not eliminate the entire CD or the MDN. In the case of such partial lesions, FROST predicts WM deficits, but not as severe as those shown in Figure 10. In fact, the FROST predictions that lesions to the CD or MDN will impair, but not abolish, WM have considerable empirical support (Janahashi et al., 2002; Van der Werf et al., 2000; Gabrieli, Singh, Stebbins, & Goetz, 1996; Hunt & Aggleton, 1991). As mentioned above, such results, which are predicted by FROST, represent a major challenge to purely cortical models of WM.

RELATIONS TO OTHER WORKING MEMORY MODELS

WM has been a topic of intense research for many years. As a result, there are many WM models in the literature. Even so, FROST is perhaps the only computational model to account simultaneously for single-cell recording data and such a diverse amount of human behavioral data. Nevertheless, because it is based on widely known neuroanatomy, there are a number of similar biologically plausible models.

FROST differs from other biologically plausible models of WM in that it hypothesizes that the pFC, CD, GP, and MDN are all important for WM maintenance. More specifically, in agreement with existing data from single-unit recording studies, FROST predicts sustained activation during the delay period in each of the above regions, and suggests a mechanism by which these regions contribute to WM maintenance. In contrast,

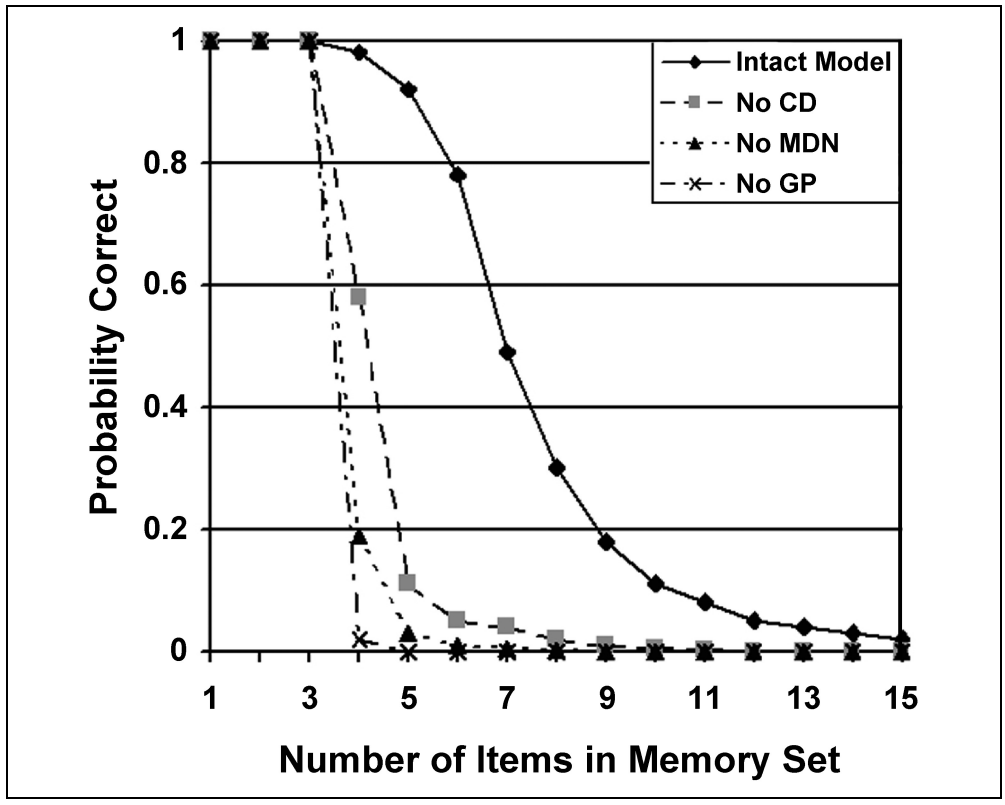

Figure 10. Predicted performance in a classic memory span test (e.g., digit span) following lesions to the CD, GP, or MDN. Parameter values were identical to those in generating the FROST prediction shown in Figure 6.

most models assume that WM maintenance is mediated primarily by sustained, delay-related activity entirely within the pFC (Deco et al., 2004; Frank et al., 2001; Braver & Cohen, 2000; Durstewitz et al., 2000a; Tagamets & Horwitz, 2000; O'Reilly et al., 1999). Furthermore, although many pFC models of maintenance provide a more detailed account of pFC function than the current instantiation of FROST (see Durstewitz et al., 2000b for a review), a more detailed version of FROST, which replaces the single pFC unit of Equation 2 with the more detailed layer-rich model shown in Figure 3, could endow FROST with many of the attractive features of the pFC models, yet still account for the data implicating subcortical structures.

The emphasis on the pFC in models of WM is, at least in part, due to the observation that, unlike cells in posterior cortical regions, single-unit activity in the pFC is extremely resistant to interference (Miller et al., 1996). In fact, sustained activation in other cortical regions has been assumed to be driven by sensory input (Durstewitz et al., 2000b). However, this account is at odds with single-unit recording data indicating mnemonic functions of the basal ganglia (Mushiake & Strick, 1995; Hikosaka et al., 1989) and data suggesting a functional role of the thalamus and basal ganglia in WM (Janahashi et al., 2002; Van der Werf et al., 2000; Gabrieli et al., 1996; Hunt & Aggleton, 1991). FROST offers a powerful resolution to this controversy. According to FROST, delay activity is most robust within the pFC because the pFC is the only region receiving recurrent excitation from three sources—the sensory association cortex, the thalamus, and the other layers within the pFC.

The models that are most similar to FROST include some, but not all, of FROST's recurrent, excitatory loops. One class of models includes the recurrent loop between sensory association areas of the cortex and the pFC (Raffone & Wolters, 2001). Another class includes the pFC–thalamic recurrent loop (Monchi, Taylor, & Dagher, 2000; Taylor & Taylor, 2000; Beiser & Houk, 1998). Beiser and Houk (1998) assume the pFC–thalamic recurrent loop facilitates maintenance, and that the basal ganglia gates input into WM (via the GP or midbrain dopamine neurons), but they do not assume that sustained changes in firing within the basal ganglia are responsible for WM maintenance.

Another popular class of models assumes a similar gating role for the basal ganglia either by disinhibition of the thalamus (Frank et al., 2001) or by dopamine input to the pFC (Braver & Cohen, 2000; O'Reilly et al., 1999). However, these models also assume that WM maintenance is mediated primarily within the pFC rather than being distributed. In the Frank et al. model, for example, cortico-cortical and cortico-thalamic loops play a relatively minor role, participating mainly in the passive maintenance of information. The type of active WM maintenance that is assumed to be robust to interference from distracting stimuli is mediated by recurrent, excitatory connections across different cortical layers within individual pFC neurons. Similar to the model of

Beiser and Houk (1998), the basal ganglia is assumed to be important for gating input to WM via phasic disinhibition of the thalamus, and consequently, switching on this intracellular maintenance mechanism in the pFC. According to Frank et al., the neurophysiological evidence for such an intracellular mechanism is tentative; thus, the biophysically simpler mechanism proposed by FROST for maintaining pFC activity presents a plausible alternative—namely, that robust maintenance in the pFC is a consequence of the activation of *many* recurrent loops between *many* brain areas with sustained-activation profiles rather than being a property of a *single* cell, area, or loop.

The model that is most similar to FROST was proposed by Monchi et al. (2000) and Taylor and Taylor (2000). Both models assign a similar role to the CD. However, in the Monchi, Taylor et al. model, the caudate becomes active upon stimulus presentation, whereas in FROST the caudate becomes active only after the stimulus is removed. The single-cell recording data shown in Figure 1 (i.e., Mushiake & Strick, 1995; Hikosaka, Sakamoto, & Sadanari, 1989) favor FROST on this point.

To summarize, FROST assumes that maintenance is achieved primarily by recurrent cortico-thalamic projections that are initiated by recurrent cortico-cortical projections from sensory areas that depend on the stimulus modality (e.g., PPC) and by cortico-subcortical projections from frontal areas to the CD. As in other models that postulate a role for the basal ganglia, disinhibition of the thalamus is critical for the initiation of sustained activation. However, we do not assume that this disinhibition is phasic in nature. As previously discussed, the regions implicated in FROST all display sustained activation during the delay period of WM tasks. Thus, sustained depression of the GP (via inhibitory projections from the CD) results in a sustained increase in the firing rate of cells in the MDN and, in turn, sustained activation of recurrent cortico-thalamic loops. We conclude that, although FROST is based on the same neuroanatomy as other models of WM, it assigns a unique functional role to the basal ganglia and thalamus.

Conclusions

This article proposed and tested a neurocomputational model that assumes WM is mediated by a widely distributed neural network that includes multiple reverberating circuits. The model assigns a unique role to the basal ganglia. Specifically, FROST is the first model to assume that the head of the caudate and the GP contribute to WM maintenance via sustained activation that is primarily under attentional control (rather than, e.g., under stimulus control). Another unique feature of FROST is that it can be tested against both single-cell recording and behavioral data. For example, FROST predicts that, in any standard WM task, sustained increases in firing rate that begin with stimulus presentation and end with the relevant response should be observed in the sensory association cortex, the MDN of the thalamus, and some areas of the pFC. In addition, other units in the pFC and cells in the head of the caudate should show sustained increases that begin with stimulus offset (i.e., and the onset of WM demands) and end with the response. Finally, some cells in the GP should show a sustained *decrease* in firing rate that also begins with stimulus offset and ends with the response. At the behavioral level, FROST accounts for the most classic WM phenomena and by adding straightforward assumptions relating pFC activation to behavior, it can be tested against almost any data in which performance is primarily mediated by WM maintenance.

In this article, we focused on the ability of FROST to account for single-cell recording data and human behavioral data. But the model can also be fit to other types of data. For example, we have had some initial success fitting FROST to the fMRI Blood Oxygen Level Dependent (BOLD) signal (Ashby & Valentin, in press). Recent evidence suggests that the BOLD signal is driven by local field potentials (Logothetis, 2003), which are closely related to the direct solutions of the differential equations that define FROST. FROST can be fit to fMRI data by supplementing it with a model of the transformation from local field potentials to the fMRI BOLD signal. We have experimented with both linear and nonlinear models of this transformation (Ashby & Valentin, in press).

In its current form, FROST is a model of WM maintenance. WM phenomena frequently also require active manipulation of stored information. Thus, before FROST could be considered a complete model of WM, it would need to be generalized to account for manipulation as well as maintenance.

METHODS

Detailed Description of FROST

A simple model of the activation in PPC cell i at time t, denoted by $P_i(t)$, which mimics the input–output relations shown in Figure 2, is defined by the differential equation:

$$\frac{dP_i(t)}{dt} = \alpha_P U_V(t)[\alpha_P - P_i(t)] + \beta_P F_i(t)[1 - P_i(t)] - \gamma_P P_i(t), \quad (1)$$

In this equation, $F_i(t)$ is the activation in pFC cell i at time t, and the step function $U_V(t)$ represents input from lower visual areas — that is, $U_V(t)$ has a numerical value of 1 when the animal can see the target, and a value of 0 when there is no visual access to the target. The constants α_P and β_P measure the strength of the synapses

between the PPC and lower visual areas, and between the PPC and the pFC, respectively. Finally, γ_P is a measure of how quickly activation decays in the absence of any input. A simple interpretation of this equation is that the derivative on the left represents the firing rate of a PPC cell. The right side of Equation 1 therefore specifies that this firing rate increases with the magnitude of the input from lower visual areas and with the input from pFC. The term $[1 - P_i(t)]$ induces an asymptotic activation level of 1,[4] and the last term on the right causes activation to decay with time.

Using this same approach, it is straightforward to construct similar equations that describe the firing rates of the other cell types shown in Figure 2. For pFC WM unit i, activation at time t is described by:

$$\frac{dF_i(t)}{dt} = [\alpha_F T_i(t) + \beta_F P_i(t)][1 - F_i(t)] - \gamma_F \left[\sum_{j \neq i} F_j(t)\right] F_i(t) - \delta_F F_i(t), \quad (2)$$

where α_F, β_F, γ_F, and δ_F are constants, and $T_i(t)$ is the activation in thalamic unit i at time t. The first term formalizes the idea that the pFC units in Figure 2 receive excitatory inputs from the thalamus and the PPC. The second term is a computationally convenient way to model the lateral inhibition that is known to exist among pFC WM cells. Neuroanatomical studies indicate that this inhibition is mediated by GABAergic interneurons (Melchitzky & Lewis, 2003), so the second term on the right side of Equation 2 oversimplifies the neuroanatomy. It simply assumes that the amount of lateral inhibition received by pFC unit i is proportional to the total activation of all other pFC WM units. The $F_i(t)$ multiplier prevents this lateral inhibition from driving the overall activation below zero. Finally, as in Equation 1, the last term in Equation 2 ensures that in the absence of all input, activation will decay back to zero.

Activation in thalamic unit i at time t is described by:

$$\frac{dT_i(t)}{dt} = \alpha_T F_i(t)[1 - T_i(t)] - \beta_T G(t) T_i(t) - \gamma_T T_i(t), \quad (3)$$

where α_T, β_T, and γ_T are again constants, and $G(t)$ represents the input activation from the GP at time t. Note that activation in the GP decreases the firing rate of its thalamic target, reflecting its inhibitory nature (i.e., these are GABA cells).

Activation in GP at time t is described by:

$$\frac{dG(t)}{dt} = -\alpha_G C(t) G(t) - \beta_G [G(t) - G_B], \quad (4)$$

where α_G and β_G are constants, and $C(t)$ represents the inhibitory input activation from the CD at time t. The constant G_B is the baseline firing rate of cells in the GP, so the last term in Equation 4 ensures that, in the absence of input, the firing rate of pallidal cells decays to G_B rather than to zero, as in the other cell types.

Finally, activation in the CD at time t is described by:

$$\frac{dC(t)}{dt} = \alpha_C F_A(t)[1 - C(t)] - \beta_C C(t), \quad (5)$$

where α_C and β_C are again constants. The term $F_A(t)$ denotes pFC input at time t. As mentioned above, the single-cell recording data shown in Figure 1 suggest that the delay-related activity of the CD, and also the downstream cells in the GP, are driven by the onset of the delay interval, rather than by target presentation. This supports the hypothesis that the CD is driven by some pFC unit other than the WM cells modeled in Equation 2. In fact, as mentioned above, many single-unit recording studies have identified cells in the lateral pFC that increase their firing only at the onset of the delay period. An example is shown in Figure 1. Note that the attentional demands on the subject rise significantly when the delay begins, so this other pFC unit could be interpreted as providing an attentional signal (hence, the subscript A). We model $F_A(t)$ in exactly the same way as the pFC WM units of Equation 2, except we assume the cortical input to these pFC attentional units is from some frontal-based attentional network, rather than from sensory association areas (e.g., PPC). A complete model of this attentional network is beyond the scope of this article. Instead, we model the input from the attentional network in exactly the same way that we modeled the PPC input from visual cortex (i.e., as a boxcar function). As a result, we model $F_A(t)$ as:

$$\frac{dF_A(t)}{dt} = [\alpha_A U_A(t) + \beta_A T_A(t)][1 - F_A(t)] - \gamma_A F_A(t), \quad (6)$$

for some constants α_A, β_A, and γ_A. The boxcar function $U_A(t)$ equals Δ (where $0 \leq \Delta \leq 1$) for the first second during which the animal must remember, and is set to zero otherwise.[5] The numerical value of Δ is assumed to increase with attentional demands. Thus, under conditions of high attention, the model assumes that the pFC input to the CD is greater than under conditions of low attention. The function $T_A(t)$ represents the thalamic input to this pFC unit.[6]

Fitting the Single-Cell Recording Data

Although the Figure 1 data were collected using several different methodologies, they all required monkeys to remember the spatial location of a target for some brief period. The major difference across experiments was in

the motor response required of the animals (e.g., arm movement vs. eye movement). FROST does not include motor circuits (which are downstream of the FROST architecture), so these differences are irrelevant to our simulations. Our goal in these simulations was to determine whether FROST is consistent with the qualitative results of these experiments. We made no attempt to fit the data quantitatively. For this reason, all simulations were done with the same temporal parameters. Specifically, stimulus presentation was set to 1 sec and the delay was set to 6 sec. Given these constraints, parameter estimates were crudely adjusted until the results shown in Figure 4 were obtained.

Fitting the Magical Number 4, 5, 6, or 7 Data

We assumed that item i is correctly retrieved if activation in pFC unit i is greater than some threshold T at retrieval time t_R — that is, that

$$\begin{aligned}&\text{Probability (Correct on item } i\text{)}\\&\quad = \text{Probability}[F_i(t_R) + \epsilon > T],\end{aligned} \tag{7}$$

where $F_i(t_R)$ is the activation in pFC unit i at retrieval time t_R, and ϵ is noise (normally distributed with mean 0 and variance σ_ϵ^2).

For the basic FROST circuit, we used the same parameter estimates that were used in the Figure 4 simulations of the single-cell recording data. Thus, there are only three parameters that must be estimated during the data fitting process: the threshold T, the error variance σ_ϵ^2, and the strength of pFC lateral inhibition (i.e., γ_F from Equation 2). Unlike the other parameters in Equations 1–5, γ_F could not be estimated from the single-cell recording data because these recordings were all made during tasks when the memory load was only a single item. With only one active WM loop, there is no opportunity for lateral inhibition. The parameters were estimated using an iterative least squares algorithm.

Acknowledgments

This research was supported in part by National Science Foundation Grant BCS99-75037 and by Public Health Service Grant MH3760. We thank Kelle Karp and Emily Waterhouse for their help in manuscript preparation.

Reprint requests should be sent to F. Gregory Ashby, Department of Psychology, University of California, Santa Barbara, CA 93106, or via e-mail: ashby@psych.ucsb.edu.

Notes

1. A more detailed computational model could be built that would also capture the neuroanatomical complexities shown in Figure 3, but for the applications considered in this article, the predictions would be virtually identical.

2. The integrate-and-fire model assumes a spike is generated whenever the integrated activation exceeds a threshold V_0. White noise is added to the activation and after a spike occurs, the integral is reset to zero. The integrate-and-fire model has two free parameters (V_0 and the noise variance) (see Koch, 1999 or Ashby & Valentin, in press for details).

3. The magnitude of lateral inhibition in the PFC, γ_F in Equation 2, will be estimated later. This is because the Figure 1 data were all collected in tasks when the memory load was only a single item. With only one active working memory loop, there is no opportunity for lateral inhibition. The parameters of Equation 6 were estimated from the recordings made in the caudate nucleus (because the PFC attentional signal is the primary caudate input).

4. The value of α_P is used as the first asymptote, instead of 1, only to cause activation induced by stimulus presentation to increase quickly enough so that asymptote is reached even for relatively brief presentation times (i.e., the numerical value of α_P can be set to a larger value when the asymptote is set to α_P rather than 1). Another method of achieving this same goal would be to set $U_V(t)$ to a value greater than 1 when the stimulus is present.

5. The 1-sec duration during which this function equals 1 is arbitrary. Changing this value has little effect on any predictions derived in this article.

6. For simplicity, we assumed that the thalamic input to the PFC attentional signal is the same as the thalamic input to the PFC WM unit (i.e., and therefore modeled by Equation 3). Whether the PFC attentional signal and WM units have the same or different thalamic inputs makes no difference to any predictions derived in this article.

REFERENCES

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience, 9,* 357–381.

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review, 105,* 442–481.

Ashby, F. G., Isen, A. M., & Turken, A. U. (1999). A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review, 106,* 529–550.

Ashby, F. G., & Valentin, V. V. (in press). Computational cognitive neuroscience: Building and testing biologically plausible computational models of neuroscience, neuroimaging, and behavioral data. In M. J. Wenger (Ed.), *2004 Notre Dame series on quantitative methodology.* Hillsdale, NJ: Erlbaum.

Baddeley, A. D. (1986). *Working memory.* New York: Oxford University Press.

Baddeley, A. D. (1992). Working memory. *Science, 255,* 556–559.

Beiser, D. G., & Houk, J. C. (1998). Model of cortical–basal ganglionic processing: Encoding the serial order of sensory events. *Journal of Neurophysiology, 79,* 3168–3188.

Bennett, B. D., & Wilson, C. J. (2000). Synaptology and physiology of neostriatal neurones. In R. Miller & J. R. Wickens (Eds.), *Brain dynamics and the striatal complex* (pp. 111–140). Amsterdam: Harwood.

Braver, T. S., & Cohen, J. D. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working memory. In S. Monsell & J. Driver (Eds.), *Attention and performance: XVIII. Control of cognitive processes* (pp. 713–737). Cambridge: MIT Press.

Callicott, J. H., Mattay, V. S., Bertolino, A., Finn, K., Coppola, R., Frank, J. A., Goldberg, T., & Weinberger, D. R. (1999). Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. *Cerebral Cortex, 9,* 20–46.

Carpenter, P. A., Just, M. A., & Reichle, E. D. (2000). Working memory and executive functions: Evidence from neuroimaging. *Current Opinion in Neurobiology, 10,* 195–199.

Constantinidis, C., & Steinmetz, M. A. (1996). Neuronal activity in posterior parietal area 7a during the delay periods of a spatial memory task. *Journal of Neurophysiology, 76,* 1352–1354.

Conway, A. R. A., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General, 123,* 354–373.

Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87–185.

Cowan, N., Nugent, L. D., Elliott, E. M., Ponomarev, I., & Saults, J. S. (1999). The role of attention in the development of short-term memory: Age differences in the verbal span of apprehension. *Child Development, 70,* 1082–1097.

Deco, G. Rolls, E. T., & Horwitz, B. (2004). ''What'' and ''where'' in visual working memory: A computational neurodynamical perspective for integrating fMRI and single-neuron data. *Journal of Cognitive Neuroscience, 16,* 683–701.

Durstewitz, D., Seamans, J. K., & Sejnowski, T. J. (2000a). Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. *Journal of Neurophysiology, 83,* 1733–1750.

Durstewitz, D., Seamans, J. K., & Sejnowski, T. J. (2000b). Neurocomputational models of working memory. *Nature Neuroscience Supplement, 3,* 1184–1191.

Engle, R. W. (2001). What is working memory capacity? In H. L. Roediger III, J. S. Nairne, J. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 297–314). Washington, DC: American Psychological Association.

Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: A computational model. *Cognitive, Affective, and Behavioral Neuroscience, 1,* 137–160.

Fuster, J. M. (1973). Unit activity in prefrontal cortex during delayed-response performance: Neuronal correlates of transient memory. *Journal of Neurophysiology, 36,* 61–78.

Fuster, J. M. (1989). *The prefrontal cortex* (2nd ed.). Philadelphia: Lippincott-Raven.

Fuster, J. M., & Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science, 173,* 652–654.

Fuster, J. M., & Alexander, G. E. (1973). Firing changes in cells of the nucleus medialis dorsalis associated with delayed response behavior. *Brain Research, 61,* 79–91.

Gabrieli, J. D. (1995). Contribution of the basal ganglia to skill learning and working memory in humans. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 277–294). Cambridge: MIT Press.

Gabrieli, J. D., Singh, J., Stebbins, G. T., & Goetz, C. G. (1996). Reduced working memory span in Parkinson's disease: Evidence of the role of a frontostriatal system in working and strategic memory. *Neuropsychology, 10,* 322–332.

Goldman-Rakic, P. S. (1987). Circuitry of the prefrontal cortex and the regulation of behavior by representational knowledge. In F. Plum & V. Mountcastle (Eds.), *Handbook of physiology,* (pp. 373–417). Bethesda, MD: American Physiological Society.

Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron, 14,* 477–485.

Guilford, J. P., & Dallenbach, K. M. (1925). The determination of memory span by the method of constant stimuli. *American Journal of Psychology, 36,* 621–628.

Heimer, L. (1995). *The human brain and spinal cord* (2nd ed.). New York: Springer-Verlag.

Hikosaka, O., Sakamoto, M., & Sadanari, U. (1989). Functional properties of monkey caudate neurons: III. Activities related to expectation of target and reward. *Journal of Neurophysiology, 61,* 814–831.

Hunt, P. R., & Aggleton, J. P. (1991). Medial dorsal thalamic lesions and working memory in the rat. *Behavioral and Neural Biology, 55,* 227–246.

Janahashi, M., Rowe, J., Saleem, T., Brown, R., Limousin-Dowsey, P., Rothwell, J., Thomas, D., & Quinn, N. (2002). Striatal contribution to cognition: Working memory and executive function in Parkinson's Disease before and after unilateral posteroventral pallidotomy. *Journal of Cognitive Neuroscience, 142,* 298–310.

Jonides, J., Schumacher, E. H., Smith, E. E., Lauber, E. J., Awh, E., Minoshima, S., & Coeppe, R. A. (1997). Verbal working memory load affects regional brain activation as measured by PET. *Journal of Cognitive Neuroscience, 9,* 462–475.

Jonides, J., Smith, E. E., Koeppe, R. A., Awh, E., Minoshima, S., & Mintun, M. (1993). Spatial working memory in humans as revealed by PET. *Nature, 363,* 623–625.

Keele, S. W., Ivry, R., Mayr, U., Hazeltine, E., & Heuer, H. (2003). The cognitive and neural architecture of sequence representation. *Psychological Review, 110,* 316–339.

Koch, C. (1999). *Biophysics of computation: Information processing in single neurons.* New York: Oxford.

Lewis, S. J. G., Cools, R., Robbins, T. W., Dove, A., Barker, R. A., & Owen, A. M. (2002). Using executive heterogeneity to explore the nature of working memory deficits in Parkinson's disease. *Neuropsychologia, 41,* 645–654.

Logothetis, N. K. (2003). The underpinnings of the BOLD functional magnetic resonance imaging signal. *Journal of Neuroscience, 23,* 3963–3971.

Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General, 111,* 1–22.

Melchitzky, D., & Lewis, D. (2003). Pyramidal neuron local axon terminals in monkey prefrontal cortex: Differential targeting of subclasses of GABA neurons. *Cerebral Cortex, 13,* 452–460.

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

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

Miyake, A. (2001). Individual differences in working memory: Introduction to the special section. *Journal of Experimental Psychology: General, 130,* 163–168.

Monchi, O., & Taylor, J. G. (1999). A hard wired model of coupled frontal working memories for various tasks. *Information Sciences, 113,* 221–243.

Monchi, O., Taylor, J. G., & Dagher, A. (2000). A neural model of working memory processes in normal subjects, Parkinson's disease, and schizophrenia for fMRI design and predictions. *Neural Networks, 13,* 953–973.

Mushiake, H., & Strick, P. L. (1995). Pallidal neuron activity

during sequential arm movements. *Journal of Neurophysiology, 74,* 2754–2758.
Nairne, J. S. (1991). Positional uncertainty in long-term memory. *Memory and Cognition, 19,* 232–240.
Oliveri, M., Turriziani, P., Carlesimo, G. A., Koch, G., Tomaiuolo, F., Panella M., & Caltagirone, C. (2001). Parieto-frontal interactions in visual–object and visual–spatial working memory: Evidence from transcranial magnetic stimulation. *Cerebral Cortex, 11,* 606–618.
O'Reilly, R. C., Braver, T. S., & Cohen, J. D. (1999). A biologically based computational model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 375–411). New York: Cambridge University Press.
Pollack, I. (1952). The information of elementary auditory displays. *Journal of the Acoustical Society of America, 24,* 745–749.
Posner, M. I., & Petersen, S. E. (1990). Attention systems in the human brain. *Annual Review of Neuroscience, 13,* 25–42.
Postle, B. R., Jonides, J., Smith, E., Corkin, S., & Growdon, J. H. (1997). Spatial, but not object, delayed response is impaired in early Parkinson's disease. *Neuropsychology, 11,* 1–9.
Raffone A., & Wolters, G. (2001). A cortical mechanism for binding in visual working memory. *Journal of Cognitive Neuroscience, 13,* 766–785.
Schultz, W., & Romo, R. (1992). Role of primate basal ganglia and frontal cortex in the internal generation of movements: I. Preparatory activity in the anterior striatum. *Experimental Brain Research, 91,* 363–384.
Strick, P. L., Dum, R. P., & Pickard, N. (1995). Macro-organization of the circuits connecting the basal ganglia with the cortical motor areas. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia.* Cambridge: MIT Press.
Tagamets, M. A., & Horwitz, B. (2000). A model of working memory: Bridging the gap between electrophysiology and human brain imaging. *Neural Networks, 13,* 941–952.
Taylor, J. G., & Taylor, N. R. (2000). Analysis of recurrent cortico-basal ganglia–thalamic loops for working memory. *Biological Cybernetics, 82,* 415–432.
Van der Werf, Y. D., Witter, M. P., Uylings, H. B. M., & Jolles, J. (2000). Neuropsychology of infarctions in the thalamus: A review. *Neuropsychologia, 38,* 613–627.
Williams, G. V., & Goldman-Rakic, P. S. (1995). Modulation of memory fields by dopamine D1 receptors in prefrontal cortex. *Nature, 376,* 572–575.
Wilson, C. J. (1990). *The synaptic organization of the brain: Basal ganglia* (pp. 279–316). Oxford: Oxford University Press.
Wilson, C. J. (1995). The contribution of cortical neurons to the firing pattern of striatal spiny neurons. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 29–50). Cambridge: Bradford.