

The Cognitive Neuroscience of Working Memory

Mark D'Esposito¹ and Bradley R. Postle²

¹Helen Wills Neuroscience Institute and Department of Psychology, University of California, Berkeley, California 94720; email: despo@berkeley.edu

²Departments of Psychology and Psychiatry, University of Wisconsin, Madison, Madison, Wisconsin 53719; email: postle@wisc.edu

Annu. Rev. Psychol. 2015. 66:115–42

First published online as a Review in Advance on September 19, 2014

The *Annual Review of Psychology* is online at psych.annualreviews.org

This article's doi:
10.1146/annurev-psych-010814-015031

Copyright © 2015 by Annual Reviews.
All rights reserved

Keywords

working memory, short-term memory, cognitive control, top-down, prefrontal cortex, connectivity, dopamine

Abstract

For more than 50 years, psychologists and neuroscientists have recognized the importance of a working memory to coordinate processing when multiple goals are active and to guide behavior with information that is not present in the immediate environment. In recent years, psychological theory and cognitive neuroscience data have converged on the idea that information is encoded into working memory by allocating attention to internal representations, whether semantic long-term memory (e.g., letters, digits, words), sensory, or motoric. Thus, information-based multivariate analyses of human functional MRI data typically find evidence for the temporary representation of stimuli in regions that also process this information in non-working memory contexts. The prefrontal cortex (PFC), on the other hand, exerts control over behavior by biasing the salience of mnemonic representations and adjudicating among competing, context-dependent rules. The “control of the controller” emerges from a complex interplay between PFC and striatal circuits and ascending dopaminergic neuromodulatory signals.

Contents

INTRODUCTION	116
COGNITIVE MODELS OF WORKING MEMORY	117
The Temporary Activation of LTM Representations	117
Sensorimotor Recruitment	118
Capacity Limits of Visual Working Memory	119
NEURAL MECHANISMS UNDERLYING WORKING MEMORY	120
The Neural Plausibility of State-Based Accounts of Working Memory	120
Working Memory at the Systems Level	121
CONCLUSIONS	133

INTRODUCTION

Introduction of the term working memory into the behavioral literature can be traced back to a passage in the book by Miller et al. (1960), *Plans and the Structure of Behavior*. In it, the authors state,

Without committing ourselves to any specific machinery, therefore, we should like to speak of the memory we use for the execution of our Plans as a kind of quick access, “working memory.” There may be several Plans, or several parts of a single Plan, all stored in working memory at the same time. In particular, when one Plan is interrupted by the requirements of some other Plan, we must be able to remember the interrupted Plan in order to resume its execution when the opportunity arises. When a Plan has been transferred into the working memory we recognize the special status of its incompleted parts by calling them “intentions.” (p. 65)

Soon thereafter, Pribram and colleagues (1964) posited that the neural machinery supporting working memory may include the prefrontal cortex (PFC). They did so on the basis of the deficits that PFC lesions were known to produce on various tests that imposed a delay between the target stimulus and the subsequent target-related response (or, in the case of delayed alternation, between the execution of one action and the execution of a subsequent action that depended on the former).

The most enduring conceptualization of working memory, however, has been that of the multicomponent model, introduced in 1974 by experimental psychologists Alan Baddeley and Graham Hitch (1974). The model was developed to address two factors in the literature of the time. First, Baddeley and Hitch’s assessment that contemporary models of short-term memory (STM) did not capture the fact that mental operations performed on information in conscious awareness can be carried out independent of interaction with, or influence on, long-term memory (LTM); for example, maintenance rehearsal had recently been shown not to enhance encoding into LTM. And second, their own work indicated that performance on each of two tasks under dual-task conditions could approach levels of performance under single-task conditions if the two engaged different domains of information, specifically verbal and visuospatial. Thus, the original version of their model called for two STM buffers (dubbed the “phonological loop” and the “visuospatial sketchpad,” respectively) that could operate independently of each other and independently of LTM, although both under control of a separate system that they dubbed the “central executive” (Baddeley 1986).

From a functional perspective, the multicomponent model of working memory accomplished the buffering and coordinating operations that Miller et al. (1960) had identified as critical if one

is to be able to maintain and successfully carry out multiple behavioral goals simultaneously. In 1986, Baddeley summarized it as “the temporary storage of information that is being processed in any of a range of cognitive tasks” (p. 34). The following year, Goldman-Rakic (1987) echoed these ideas in an influential synthesis of cognitive and neurobiological perspectives, stating that “the evolution of a capacity to guide behavior by [mnemonic] representation of stimuli rather than by the stimuli themselves introduces the possibility that concepts and plans can govern behavior” (p. 378). Thus, “the ability to guide behavior by representations of discriminative stimuli rather than by the discriminative stimuli themselves is a major achievement of evolution” (p. 378). What is captured in each of these seminal writings is that working memory underlies the successful execution of complex behavior, regardless of the cognitive domain or domains being engaged. When working memory fails, so too does the ability to carry out many activities of daily living. It is not surprising that, viewed from this perspective, working memory can be shown to be impaired in many neurological and psychiatric syndromes that are characterized by disordered behavior (Devinsky & D’Esposito 2003). The centrality of working memory to understanding normal, as well as pathological, behavior is presumably reflected in the intensity with which it has been studied: In late 2014, a search of the term working memory in PubMed retrieved 18,224 citations and in Google Scholar 1,580,000 results were returned. Although we cannot hope to do justice to such a vast literature in just one review, we do hope to highlight what we consider to be important developments in working memory research from a cognitive neuroscience perspective.

COGNITIVE MODELS OF WORKING MEMORY

As we write this review, the multicomponent model of working memory is marking its fortieth anniversary, and from roughly 1985 through 2005—what one might consider the first 20 years of the cognitive neuroscience study of working memory—this was the dominant theoretical framework. More recently, however, what might be called state-based models have taken on increased prominence. As a class, these models assume that the allocation of attention to internal representations—whether semantic LTM (e.g., letters, digits, words), sensory, or motoric—underlies the short-term retention of information in working memory. These models conceptualize information being held in working memory as existing in one of several states of activation established by the allocation of attention.

Our brief review of state-based models is organized into two categories: activated LTM models and sensorimotor recruitment models. Although these two types of models have arisen within different literatures, the principal difference between them seems to be simply the class of stimuli for which each has been proposed. That is, activated LTM models have by and large been articulated for, and tested with, symbolic stimuli typically considered to be semantic (e.g., letters, words, digits). Sensorimotor recruitment models, however, have typically been invoked for classes of stimuli considered to be perceptual (e.g., visual colors and orientations, auditory pitches, tactile vibrational frequencies). Despite these surface-level differences, however, both of these classes of state-dependent models of working memory are grounded in the idea that the attentional selection of mental representations brings them into working memory and that the consequences of attentional prioritization explain such properties as capacity limitations, proactive interference from no-longer-relevant items, etc.

The Temporary Activation of LTM Representations

The subset of state-based models that has been most formalized are those pertaining to working memory of information for which a semantic representation exists in LTM. In perhaps the most

well-known of the state-based models, Cowan (1995) describes two distinct states in STM: a small, capacity-limited state referred to as the focus of attention (FoA) and a more expansive state referred to as the activated portion of LTM (known as activated LTM). In this model, the FoA corresponds to approximately four chunks of information that one can hold in working memory at any moment in time using top-down attentional control. When attention subsequently shifts to other information, the items that were previously in the FoA transition into activated LTM. Activated LTM has no capacity limit per se but is susceptible to temporal decay and interference effects. A variant on this two-level model has been proposed by Oberauer (2002, 2009) in the three-embedded-components theory. In this theory, the four-item FoA from Cowan's model is recast as a region of direct access from which a narrower FoA can efficiently select information. Capacity limits, per se, do not exist for either of these two hypothesized states in working memory. Rather, the amount of information that can be retained in the region of direct access and the FoA is limited only by interference from bindings between object features being retained in working memory (Oberauer 2013). A third model, advocated by McElree (1998, 2006), posits two states of memory: a FoA with a strict capacity limit of one item, and LTM, in which all items exist along a graded continuum of "memory strength," with memory strength (which we construe as level of activation) of an item falling off as a function of how recently it was in the FoA and from which all items are equally accessible.

Despite some differences in terminology, these models all posit the following: When we are presented with symbolic information to be remembered (e.g., a list of names or a telephone number), the LTM representations of this information are accessed during the process of perceptual recognition, and they are subsequently maintained in an elevated state of activation, via attention, until this information is no longer needed to achieve some proximal goal (for our purposes, we gloss over whether there may exist one or more distinct states of attentional prioritization and summarize all as a FoA). These models account for extensive behavioral findings that support the existence of different states of representation of information being held in working memory. For example, Oberauer (2001, 2002, 2005) and colleagues have made clever use of the Sternberg effect, whereby reaction time for a recognition judgment about a memory probe increases linearly with the number of items concurrently held in working memory. The Oberauer studies have modified the basic Sternberg memory paradigm by introducing a retrocue during the memory delay that informs the subject that only a subset of the initially presented memory items will be relevant for an upcoming probe. Given sufficient time to process this retrocue, subjects respond more quickly to memory probes of the cued items (i.e., as would be predicted if they were now holding a smaller memory set). The uncued items are not fully forgotten, however, and continue to influence ongoing processing in the form of intrusion costs on response times when they are presented as negative (to-be-rejected) memory probes. This intrusion effect persists for 5 s, long after the uncued items cease to affect response times for the cued items. The uncued items are therefore hypothesized to have been removed from the FoA but to persist in activated LTM (Oberauer 2001). By varying the retrocue-to-memory probe asynchrony, Oberauer (2005) estimated that it takes ~1 s to remove uncued items from the FoA. Questions about whether there exists one or more qualitatively discrete states of activation outside of the FoA remain a topic of active research.

Sensorimotor Recruitment

The basic premise of sensorimotor recruitment models of working memory is that the systems and representations engaged to perceive information can also contribute to the short-term retention of that information. An early, paradigmatic example of such models is that of attention-based rehearsal, whereby a location in space can be held in working memory via the covert allocation of

attention to that location (e.g., Awh & Jonides 2001). For other domains of sensory information, such as visually perceived spatial frequency, contrast, orientation, or motion, behavioral evidence indicates that each is retained in a highly stimulus-specific manner (Magnussen 2000, Magnussen & Greenlee 1999, Zaksas et al. 2001) that is most parsimoniously explained as the persistent activation of the sensory representations themselves. We show in the next section that a growing body of neural evidence supports this contention.

In the literature, the label “sensory recruitment” is much more common than is “sensorimotor recruitment.” We prefer the latter, however, to accommodate the intimate, often inextricable, coupling between sensory attention and motor intention. This distinction is important particularly in the context of spatial working memory, which is disrupted not only by drawing attention to a distracting location (e.g., Awh et al. 1998) but also by concurrently performing task-irrelevant motor sequences, such as eye movements, tapping, etc. (reviewed in Postle et al. 2006). Conversely, motoric activity, such as the trajectory of a saccade, can be altered when one is concurrently holding a location in working memory (Theeuwes et al. 2005). These results support the idea that the coordinates of a to-be-remembered location are immediately incorporated into a salience map that simultaneously holds them in brain systems that represent them as a percept and as a target for action by any of various motor effectors (Postle 2011).

Capacity Limits of Visual Working Memory

A focus of intensive investigation for sensorimotor recruitment models has been the factors that explain capacity limitations. Much of this work has followed from Luck & Vogel’s (1997) experiments with a change detection task in which a target array of colored squares (varying across trials from a single square to 10 or more) is presented for a few hundred milliseconds, followed by a brief (roughly 1-second) blank delay, followed by a probe array containing the same number of items but presenting one item in a different color on half of the trials (a yes/no recognition procedure). By applying a simple algebraic formula to the results, investigators estimated that subjects had a visual STM capacity of between three and four items. They found that an individual’s capacity did not change with the number of features used to individuate objects, up through objects defined by conjunctions of four features. This observation led them to hypothesize that the capacity of visual STM is constrained by a finite number of hypothetical slots, each one capable of storing an object representation, regardless of the complexity of any single object (Vogel et al. 2001).

This slots model has been challenged from at least two perspectives and, at the time of this writing, the nature of visual STM capacity limits remains a topic of vigorous debate. One open question is that of the influence of object complexity—contrary to the findings of Vogel et al. (2001), others have found that visual STM capacity declines with increasing object complexity (e.g., Alvarez & Cavanagh 2004). A second challenge arises from the perspective that visual STM capacity may not depend on a finite number of slots but instead on a single attentional resource. Evidence for this latter view is marshaled when the procedure for testing visual STM is changed from recognition to recall. This procedural change allows researchers to estimate the precision of a mnemonic representation by measuring the error in the recall response. With STM for the orientation of one or more line segments, for example, the average error in recalling the orientation of the probed stimulus is larger when subjects are remembering several stimuli simultaneously, in comparison to when they are remembering just one stimulus (Bays & Husain 2008). That is, mnemonic precision (the inverse of recall error) declines monotonically as a function of memory set size, an outcome that one would expect if STM were supported by a limited resource that must be apportioned ever more thinly as the number of items in the memory set increases. Slots models have been modified to allow for variable representational precision within a slot, but one

contentious question that remains is how best to explain capacity limitations: Is the amount of information an individual can retain in STM limited because she has run out of slots (in which case an absolute ceiling in performance is predicted) or because her attentional resources have been spread so thin that any one item's mnemonic fidelity is too poor to be retrievable? For excellent reviews on these issues, see Ma et al. (2014) and Luck & Vogel (2013).

NEURAL MECHANISMS UNDERLYING WORKING MEMORY

One can consider a mechanism—the process by which something takes place—at many levels of detail. Here we first consider, at a relatively abstract level, evidence for the general ideas of activated LTM and of sensorimotor recruitment. Subsequently, we discuss specific systems-level neural mechanisms that may underlie these phenomena. We anticipate one likely conclusion: Numerous neural mechanisms can support the short-term retention of information in working memory, and many likely operate in parallel.

The Neural Plausibility of State-Based Accounts of Working Memory

State-based models of working memory have gained prominence in recent years because cognitive neuroscience research indicates that they accommodate neural data well. They have been particularly successful since investigators began to apply multivariate pattern analysis (MVPA) techniques to human functional neuroimaging data (these techniques have been summarized in many places; e.g., Lewis-Peacock & Postle 2012). To explore temporary activation of LTM, for example, Lewis-Peacock & Postle (2008) employed the following method. First, they scanned subjects with functional magnetic resonance imaging (fMRI) while the subjects made judgments that required them to access information from LTM: the likability of famous individuals; the desirability of visiting famous locations; the recency with which they had used a familiar object. Next, outside the scanner, they taught subjects arbitrary paired associations among items in the stimulus set. Finally, they scanned subjects a second time but while subjects performed delayed recognition of paired associates (i.e., see one item from the LTM memory set at the beginning of the trial and indicate whether the trial-ending probe was paired specifically with that item). These researchers found that multivariate pattern classifiers trained on data from the first scanning session, when subjects were accessing and thinking about information from LTM, could successfully decode the category of information that subjects were holding in working memory in the second scanning session. Such an outcome was possible only if the working memory task and the LTM task drew on the same neural representations.

MVPA has also been used to generate compelling evidence for sensorimotor recruitment models of working memory. Thus, for example, two studies have demonstrated that the primary visual cortex (V1) supports the delay period–spanning representation of the color or orientation of target stimuli on tests of delayed recognition (Harrison & Tong 2009, Serences et al. 2009). This pattern of results has been replicated with other classes of stimuli: The short-term retention of motion can be decoded from the lateral extrastriate cortex, including area MT+, as well as from medial calcarine and extracalcarine cortex (Emrich et al. 2013, Riggall & Postle 2012); the short-term retention of complex visuospatial patterns can be decoded from the occipital and parietal cortices (Christophel et al. 2012); and the short-term retention of familiar objects, faces, houses, scenes, and body stimuli can be decoded from the ventral occipitotemporal cortex (Han et al. 2013, Lee et al. 2013, Nelissen et al. 2013, Sreenivasan et al. 2014b). Finally, in relation to the frontoparietal salience map, Jerde and colleagues (2012) have demonstrated that a classifier trained only on performance of a test of attention to location, or only on performance of oculomotor

delayed response, or only on performance of spatial delayed recognition can recover trial-specific target direction from any of the three trial types. That is, for example, a classifier trained to discriminate leftward from rightward sustained attention can also correctly discriminate leftward from rightward motor preparation and leftward from rightward spatial STM, even though the classifier was never trained on the latter two trials. Thus, the functions that we label attention, intention, and retention may be treated identically by the brain.

Perhaps the most compelling evidence in support of sensorimotor recruitment models of working memory derives from two studies using multivariate approaches to link the precision of the delay-period neural representation of target stimuli in the sensory cortex with behavioral estimates of mnemonic precision, showing that “the relative ‘quality’ of . . . patterns [of activity in sensory cortex] should determine the clarity of an individual’s memory” (Ester et al. 2013, p. 754). In one study, Ester and colleagues (2013) showed that the precision of population tuning curves in areas V1 and V2 estimated from the delay-period signal from these regions predicted the fidelity with which a subject could reconstruct the target orientation at the end of the delay period. In another study, Emrich et al. (2013) varied from trial to trial the number of directions of motion that had to be remembered, and they found a reliable within-subject correlation between the load-related decline in delayed-recall precision and a load-related decline in MVPA decoding performance.

Complementing these fMRI studies are results of studies using transcranial magnetic stimulation (TMS) to alter activity in sensorimotor regions during the delay period of tests of working memory for visually presented stimuli. Hamidi and colleagues (2008, 2009), for example, have shown that delay-period repetitive TMS of the parietal cortex and frontal eye fields selectively affects spatial working memory performance. A more nuanced approach, taken for working memory of visual motion, has leveraged the fact that TMS of visual area MT can produce the percept of a moving phosphene—a flash of light that contains coherent motion within the area of the flash. The direction of motion is reproducibly perceived as toward the periphery, away from the fovea, in the visual field contralateral to the side of stimulation. Silvanto & Cattaneo (2010) demonstrated that this percept is systematically influenced when TMS is delivered while the subject engages STM to determine the direction of motion of a target stimulus. When the target motion is in the same direction as the expected motion of the phosphene, the perception of the moving phosphene is enhanced. However, when the target motion is in the opposite direction, perception of the moving phosphene is reduced. These results indicate that the physiological state of MT varies systematically as a function of the direction of motion being remembered, just as it does, when a stimulus is present, as a function of the direction of motion being perceived.

Working Memory at the Systems Level

Working memory does not derive from a discrete system, as do vision and motor control. Rather, working memory is a property of the brain that supports successful attainment of behavioral goals that are being carried out by any of several systems, including sensory systems, those that underlie semantic and episodic memory, and motor systems. We next review five neural mechanisms that likely underlie working memory function.

Persistent neural activity. The study of the neural underpinnings of working memory took a significant leap forward in 1971 with the publication of two studies featuring extracellular recordings from the PFC in monkeys performing working memory tasks. In one study, Fuster & Alexander (1971) reported that PFC neurons exhibited persistent activity during the stage of a delayed-response task in which the monkey had to actively maintain information that was no longer present yet was relevant for successfully completing the task. In the second study, Kubota

& Niki (1971) reported a comparable finding during the delay period of a delayed alternation task. The ability of neurons to generate persistent activity in the absence of external stimuli is likely of fundamental importance to the neural basis of working memory. Following on these landmark discoveries, many other labs have found such working memory–related neurons within the PFC (e.g., Funahashi et al. 1989, Miller et al. 1996). With the advent of fMRI in the early 1990s, investigators subsequently demonstrated that human PFC also exhibited persistent neural activity that appeared to be coding task-relevant information during working memory tasks (see Courtney et al. 1997 and Zarahn et al. 1997 for the first of such studies). Many characteristics of this activity support the notion that the PFC maintains representations that are critical for guiding behavior. First, neural activity endures throughout the entire length of the delay period until it can be presumably used to guide a response (Funahashi et al. 1989, Fuster & Alexander 1971). Second, it directly relates to behavior. For example, during the performance of an oculomotor delayed-response task, the magnitude of an fMRI signal in the frontal cortex reflects the fidelity of the maintained representation (Curtis et al. 2004). An open question that cannot be answered with fMRI is what mechanism underlies persistent neural activity. Specifically, the relative importance of cortico-cortical loops (long-range recurrent interactions), thalamo-cortical loops, or local cortical mechanisms (such as excitatory reverberation) for the generation of persistent neural activity has yet to be determined (Pesaran et al. 2002, Wang 2001).

In addition to the circuit-level questions summarized above, recent research applying MVPA to fMRI and electroencephalography (EEG) data has raised intriguing questions about the functions supported by persistent neural activity. These questions fall into two categories, one relating to the nature of the persistent delay-period activity that supports the short-term retention of information and the second relating to the very necessity of this activity. The first question arises from dissociations between the elevated activity that is classically observed in frontoparietal regions and subthreshold activity patterns that MVPA detects in sensory processing–related regions. As first described by Harrison & Tong (2009) and by Serences et al. (2009), the delay-period retention of visual stimulus information can be decoded from V1, despite the absence of sustained, elevated signal levels in this region. Subsequent studies by Riggall & Postle (2012) and by Emrich et al. (2013) replicated these findings and also explicitly failed to find evidence for stimulus information in the elevated delay-period activity that was present in frontal and parietal regions. Furthermore, Sreenivasan et al. (2014b) showed that the magnitude of above-threshold delay-period activity does not correlate positively with the feature weightings that underlie MVPA classification. One implication of all these studies is that above-threshold delay-period activity may support functions other than information storage per se. Determination of these functions is the topic of the final section of this review. A second implication is that the neuronal processes that drive the MVPA-decodable activity in sensory areas are operating at a level that is subthreshold from the perspective of traditional univariate statistics. Thus, an important focus of future research will be to understand the nature of this subthreshold delay-period activity. One possible explanation is that it simply reflects reduced spiking at the population level, as would be expected for a sensory area in the absence of a bottom-up drive. A second possibility, mutually compatible with the first, is that MVPA may be detecting regional heterogeneity in oscillations of local field potentials (LFPs). That is, delay-period stimulus representations may be encoded in LFPs that persist in the same networks that exhibit elevated firing when the stimulus is present.

A second question highlighted by MVPA of fMRI and EEG data is whether persistent activity is even necessary to retain information in working memory. This notion was raised when Lewis-Peacock and colleagues (2012) scanned subjects performing a multistep delayed-recognition task that presented two sample stimuli first then presented a retrocue to indicate which of the two would be relevant for the first memory probe; following this first probe, then, was a second retrocue to

indicate which stimulus would be relevant for the trial-ending second memory probe. With this procedure, an item could be irrelevant for the first memory probe (an unattended memory item) but then be relevant for the second memory probe. Although the authors initially predicted that MVPA evidence for unattended memory items would take on an intermediate level between the item that was in the FoA and baseline, this result is not what they found. Instead, in response to the retrocue, MVPA evidence for the unattended memory item dropped to baseline levels. This item nonetheless remained in working memory, as evidenced by its successful retrieval if cued by the second retrocue. This finding has been replicated in an EEG study, thereby discounting the possibility that the unattended memory item may be transferred to an oscillatory code to which fMRI is insensitive (LaRocque et al. 2013). These findings, therefore, highlight the intriguing possibility that persistent neural activity may not be necessary to maintain representations held in working memory. Indeed, this possibility has also been explored by researchers working at other levels of investigation, including computational modeling, in vitro electrophysiology, and extracellular recordings in the behaving monkey. Computationally, information can likely be sustained over brief intervals via rapid shifts in synaptic weights. In such a scenario, the encoding of the sample stimulus would be accomplished via a transient reconfiguration of synaptic weights in the networks engaged in its initial processing. The contents of working memory could then be read out when the network was activated by a subsequent sweep of activation through this network (Itskov et al. 2011, Mongillo et al. 2008, Sugase-Miyamoto et al. 2008). Empirical evidence consistent with such a mechanism has been recorded from the ventral temporal cortex (Sugase-Miyamoto et al. 2008) and the PFC (Stokes et al. 2013) in monkeys. Which mechanism could support the short-term synaptic facilitation that would be needed to implement such a scheme? Theoretically, Mongillo et al. (2008) proposed residual presynaptic calcium levels. Empirically, Erickson et al. (2010) demonstrated that an associative short-term potentiation is GluR1-dependent in an in vitro preparation. Clearly, the relative contribution of persistent neural activity versus other mechanisms that do not rely on above-baseline activity to sustain working memory representations should be a high priority for future research.

Whether working memory representations are maintained via persistent neural activity, synaptic mechanisms, or some combination of both, these storage mechanisms are consistent with state-based models of working memory, which eliminate the need for currently relevant representations to be transferred to a limited number of dedicated, specialized buffers (D'Esposito 2007, Postle 2006). In neural terms, any population of neurons can serve as a buffer. Moreover, the ability to exhibit persistent neural activity, or a shift in synaptic weights, is likely a property of all neurons, from primary cortex to the multimodal association cortex. In sum, networks of neurons located anywhere in the brain can potentially store information that can be activated in the service of goal-directed behavior.

Hierarchical representations in the prefrontal cortex. What is the nature of the neural code within the PFC? Some have put forth the idea that persistent activity in the PFC represents sensory features of information maintained in working memory (Goldman-Rakic 1995). Indeed, in the systems and cognitive neuroscience literatures, one can see that the variable popularity of stimulus representation models of the PFC have tracked very closely the multicomponent model of working memory. More recently, there has been greater emphasis on the fact that the PFC actually exhibits, at best, coarse selectivity for items and features maintained in working memory (Constantinidis et al. 2001). Furthermore, PFC delay-period activity can represent a broad range of task variables that are not directly related to the to-be-remembered stimuli. For example, lateral PFC neurons recorded from monkeys exhibit differential preferences for task rules (Warden & Miller 2010), contingent motor responses (Romo et al. 1999), and stimulus–response

mappings (Wallis et al. 2001). Studies examining population coding of lateral PFC delay activity have also found information about stimuli (Stokes et al. 2013), rules (Riggall & Postle 2012), and object categories (Meyers et al. 2008) throughout the delay period of working memory tasks. In fact, Rigotti and colleagues (2013) have recently demonstrated that neuronal activity within the PFC is tuned to mixtures of multiple task-related variables, suggesting that PFC representations exhibit high-dimensionality. That is, many dimensions are needed to characterize the distinct (multivariate) patterns that can be taken on by the sampled population of neurons across various experimental conditions. Moreover, this dimensionality is predictive of the animal's behavior: The population of PFC neurons exhibited a decrease in dimensionality on error trials. The authors of the very first reports of persistent activity within the PFC offered interpretations that are in line with many current models. For example, Fuster & Alexander (1971) wrote,

The temporal pattern of firing frequency observed in prefrontal and thalamic units during cue and delay periods suggest [sic] the participation of these units in the acquisition and temporary storage of sensory information which are implicated in delay response performance. Their function, however, does not seem to be the neural coding of information contained in the test cues, at least according to a frequency code, for we have not found any unit showing differential reactions to the two positions of the reward. It is during the transition from cue to delay that apparently the greatest number of prefrontal units discharge at firing levels higher than the intertrial baseline We believe that the excitatory reactions of neurons in MD [nucleus medialis dorsalis] and granular frontal cortex during delayed response trials are specifically related to the focusing of attention by the animal on information that is being or has been placed in temporary memory storage for prospective utilization. (p. 654)

Several human fMRI studies have directly investigated the nature of representations being maintained in PFC as compared with posterior cortical regions. In one study, subjects viewed a sample display of dot motion, then, halfway through the delay period, subjects were informed about whether they would be probed on memory for the speed or for the direction of the sample motion. Delay-period MVPA decoding of stimulus direction was successful only at lateral and medial regions of the occipital cortex that are associated with visual perception. The PFC, however, was seen to represent a more abstract level of task performance: whether a trial was a speed trial or a direction trial (Riggall & Postle 2012). A different study using different stimuli but a similar procedure found analogous results. In it, subjects first viewed a common object and were then informed about whether the memory probe would require a fine-grained perceptual judgment or a category-membership judgment. On perceptual trials, MVPA decoded stimulus identity in the ventral occipitotemporal cortex but not in the PFC. On category trials, MVPA decoded stimulus identity from the PFC but not from the occipitotemporal cortex (Lee et al. 2013). These two findings are consistent with prior studies demonstrating that the lateral PFC preferentially encodes and maintains arbitrary and abstract representations of object category over representations of visual similarity (Chen et al. 2012; Freedman et al. 2001, 2003; Meyers et al. 2008). Further support for the distinction between stimulus-selective lateral PFC representations and sensory representations comes from a second fMRI study that required subjects to remember over a short interval either faces or scenes or both categories of information (Sreenivasan et al. 2014b). The investigators reasoned that if a region supports a sensory representation of working memory stimuli, then the remember faces trials should be incorrectly classified as remember both trials more often than they should be misclassified as remember scenes trials because the sensory representation of faces is more similar to the representation of remembering faces and scenes together than it is to remembering only scenes. Similarly, remember scenes trials should also be disproportionately misclassified as remember both trials if activity patterns encode sensory

representations. The findings from this fMRI study suggested that what is represented by the PFC is higher-order information, such as task rules, goals, or abstract representations of the categories, as compared with what is represented by the extrastriate cortex, which may be more stimulus specific (e.g., the identity of specific faces).

These empirical findings are consistent with the original theoretical notions put forth by Fuster (1990) and Miller & Cohen (2001) that integrated representations of task contingencies and rules are maintained in the PFC, which is critical for the mediation of events separated in time but contingent on one another. This formulation of PFC function places less emphasis on a storage role and instead emphasizes its role in providing top-down control over all other brain regions where information is actually stored (D'Esposito et al. 2000, Petrides 2000, Smith & Jonides 1999). Thus, the sustained activity in the PFC does not reflect the storage of representations per se; it reflects the maintenance of high-level representations that provide top-down signals to guide the flow of activity across brain networks [see also Sreenivasan et al. (2014a) and Postle (2014)]. This idea is explored further in the next section. However, we must consider first the nature of the information represented within the PFC about the functional organization of the PFC as a whole.

The PFC is a heterogeneous region covering a significant amount of territory in the brain. In this review we are focusing on the lateral PFC and not the medial or the orbital PFC regions, which likely have distinct yet complementary functions (Cummings 1993). Any understanding of the nature of the representations stored and maintained in the PFC that are necessary for goal-directed behavior must consider subregional differences in both cellular makeup and connectivity. Numerous neuropsychological, physiological, and imaging studies support the general idea that as one moves rostral (anteriorly) in the frontal cortex, from the premotor cortex to the frontopolar cortex, the processing requirements of these regions for planning and selection of action are of higher order (Burgess et al. 2007, Christoff et al. 2003, Ramnani & Owen 2004). Koechlin and colleagues (2003) have put forth a hypothesis that the frontal cortex may be organized from rostral to caudal in a hierarchical fashion en route to action (see also Fuster 2004 for an earlier formulation of a similar idea). Specifically, Koechlin & Summerfield (2007) propose a cascade model that predicts that competition among alternative action representations is resolved on the basis of mutual information with various contextual information, termed control signals. Using fMRI in healthy subjects, Koechlin and colleagues (2003) found support for their predictions by demonstrating that as contextual information required to select a response was more abstract and relevant over a longer temporal interval, fMRI activation progressed from caudal to more rostral regions of the frontal cortex.

In an fMRI study (Badre & D'Esposito 2007), we aimed to replicate and extend Koechlin's findings regarding the proposed rostral-caudal functional gradient along the frontal cortex. We specifically tested an alternative idea, that this gradient derives from a hierarchy ranked by the abstractness of the representation to be selected. In this study, healthy subjects performed a response selection task that required more abstract action decisions to be made across behavioral conditions. The lowest level of the task performed was called the response task, where subjects learned that a colored square corresponded to a particular finger response. At the next level, known as the feature task, each colored square corresponded to a particular shape, and then subjects chose their motor response if the colored square matched the shape. Thus, at this level, there is not enough information in color alone to determine the correct response. The object shape had to be considered in conjunction with the color to make a response. The only difference from the response task was that the colors now mapped to relevant shapes that cued a correct response rather than mapping directly to the correct response. Thus, an action decision must be based on a more abstract action representation. At the next level, known as the dimension task, subjects learned that a particular color corresponded to a particular dimension of an object (shape

or orientation), and they were required to compare the two objects along a particular dimension and indicate with a motor response whether the objects matched or mismatched along only the relevant dimension. The subject knew which dimension was relevant on the basis of the color of the square bounding the objects. Hence, the design for the mappings was identical to those of the feature and response tasks, except that it was now color mapped to dimension rather than to feature or response. Again, the action decision must be based on more abstract representation. The final and highest level was the context task, during which subjects performed the dimension task but conflict was manipulated by varying the frequency of the sets of color to dimension mappings. In this case, the temporal context was required to select the appropriate context (the color cue) for determining the dimension. Thus, selection of the relevant context was more abstract.

During the lowest-level task, the response task, activation was found in the posterior frontal cortex within the premotor cortex; area 6). At the next higher-level task, the feature task, activation was found anterior to the premotor cortex within the pre-premotor cortex; area 8). On the next higher-level task, the dimension task, activation was noted anterior to this location within the inferior frontal sulcus (IFS) on the border of areas 45 and 9/46. Finally, activation on the highest-level task, the context task, was found in the most anterior location within the frontopolar cortex or area 10. Thus, as action representations became more abstract, activation within the frontal cortex moved anteriorly (or rostrally). This progression of activation from the posterior to the anterior portions of the frontal cortex was not simply due to the task becoming more complex or difficult; we also varied the difficulty within each individual task (e.g., response, feature, dimension, or context) and found that activation within that particular region engaged by each task increased in magnitude with difficulty but did not change its location within the frontal cortex. Whereas Koechlin et al. (2003) emphasize temporal and contextual factors in differentiating regions of the frontal cortex, these results instead suggest that regions of the PFC may be differentiated by the level of abstraction at which the action representations must be selected over competition.

Thus, human fMRI studies support the notion that there is a functional gradient along the anterior-to-posterior axis of the frontal lateral cortex. A similar functional gradient relating to motivational aspects of cognitive control has been identified along the medial PFC axis (Kounieher et al. 2009, Venkatraman et al. 2009), and functional connectivity between the medial and lateral PFC has been observed (Blumenfeld et al. 2013). The reader should note that consensus has not been reached regarding the specific details of the functional gradient observed in the PFC (see Badre 2008 for review). Nevertheless, an important component of emerging models of lateral PFC organization is the presence of a hierarchy. A processing hierarchy within the frontal cortex requires that anterior regions influence the processing in posterior regions more than posterior regions influence anterior regions. How can one obtain direct evidence to support this claim? Essential clues (albeit indirect ones) regarding a hierarchical rostro-caudal organization of the frontal lobe can be derived from its anatomical organization. If there were a hierarchical arrangement, anatomical connectivity among PFC subregions would likely display a pattern where area 10, at the highest level, would have projections back down to area 6 at the lowest level. However, area 6 would not necessarily project back up to area 10. Such a pattern does appear to exist, at least in rhesus monkeys (Badre & D'Esposito 2009). Barbas & Pandya (1991) have also noted that different frontal regions have different degrees of differentiation at the columnar level. More differentiated regions are more laminated (e.g., aggregation of cells into cortical layers). Caudal areas with well-developed laminar differentiation (such as area 8 or caudal 46) have restricted connections mostly to neighboring regions. In contrast, rostral areas that have less laminar differentiation (such as area 10) have widespread connections to other areas. In this scheme, less differentiated areas such as those in the rostral PFC (areas 10, 9, 46), which have more diffuse projections, are well situated to constitute the top of a hierarchy. In contrast, more differentiated areas such as those in

the caudal PFC (area 9/46, 8) have more intrinsic connections and are well situated to be lower in a hierarchy.

Further indirect evidence for a hierarchical organization within the lateral PFC derives from functional neuroimaging studies that examine effective connectivity or the causal influence that one brain region may have on another. For example, in the aforementioned Koechlin et al. (2003) study, structural equation modeling of the imaging data showed that activation in rostral frontal regions accounted for variance in activation in caudal frontal regions but not vice versa. Direct evidence for a hierarchical organization within the lateral PFC requires lesion data. That is, a rostral to caudal flow of control processing within the frontal lobes predicts that performance on tasks requiring higher-order control should be impaired by disruptions to lower-order processors, even when the higher-order processors are intact. However, when a higher-order control processor is disrupted, performance should be unaffected on tasks that require only lower-order control. This hypothesized asymmetric deficit pattern cannot be directly tested with neurophysiological methods such as fMRI, EEG, and single-unit recording. Rather, it requires a lesion method that leads to isolated disruption of specific processors along the proposed hierarchical gradient.

Additionally, using the cognitive tasks we implemented in the fMRI study, we have carried out a behavioral study of patients with focal frontal lesions to test the hypothesis that the frontal cortex has a hierarchy (Badre et al. 2009). Specifically, we tested whether a lesion to the pre-motor cortex region of the frontal cortex (area 8), assumed to damage a second-level processor, would impair performance on the feature task as well as on the dimension and context tasks but would not affect performance on the response task. The reasoning was that disruption of the second level of a hierarchy should interfere with processing at higher levels (feature, dimension, and context tasks at the third and fourth levels) but not at lower levels (response task at the first level). By contrast, a more anterior IFS lesion (areas 45; 9/46), which would damage a third-level processor, should impair performance on the dimension task (third level) as well as on the context task (fourth level), but not performance on the feature (second level) or response (first level) tasks. Such a pattern of behavioral results in patients with focal frontal lesions would be direct evidence for a hierarchical organization of frontal lobe function. We predicted that because of the asymmetric dependencies predicted by a hierarchy, deficits in higher-level tasks will be more likely across patients, regardless of their lesion site, than will deficits in lower-level tasks. Thus, the presence of an impairment at any level should increase the likelihood of an impairment at all higher levels but should not increase the odds of an impairment at a lower level. We observed that the probability of a deficit on any task was 62% across patients. Critically, however, the probability of a deficit at any level, given a deficit at a lower level, was 91% across patients, a significant change over the probability of a deficit on any task. By contrast, the probability of a deficit at any level given a deficit at a higher level was only 76%, a weak change over the prior probability of a deficit on any task. This asymmetry provides initial support for the hierarchical dependencies among behavioral deficits at the different task levels as well as for the aggregation account of the group data. Recently, this pattern of findings supporting a frontal hierarchy has been replicated in another group of patients with focal frontal lesions (Azuar et al. 2014).

Hierarchical organization of rules and goals has many advantages. For example, increasingly abstract representations of rules and goals could serve as different top-down signals that could bias particular but different action pathways over competitors, allowing for flexible goal-directed behavior. Take the example of the seemingly simple act of hitting a golf ball. Hitting the ball in the proper direction requires the flag to remain on the green temporarily—a relatively concrete representation. If the golf ball is in a fairway bunker, the golfer must temporarily retain the more abstract representation of the golf rule stating that the golf club cannot touch the sand before hitting the ball or a penalty will be assessed. Finally, throughout this act of hitting the ball,

the golfer might also benefit from maintaining an even more abstract representation of the knowledge that golf provides exercise and is a healthy behavior. In this way, simultaneous maintenance of hierarchically organized representations within the PFC can provide independent, yet likely interactive top-down bias signals that may (or may not) lead to a successful goal-directed behavior.

Top-down signaling. The PFC has long been implicated as a source of top-down signals that can influence processing in other cortical and subcortical brain regions (Braver et al. 2008, Duncan 2001, Fuster 2008, Shallice 1982). One type of PFC top-down signal likely provides direct feedback to posterior cortical regions that process incoming sensory input from a particular modality (e.g., visual or auditory). For example, when a person is looking into a crowd of people, the visual scene presented to the retina may include a vast array of visual information. However, if someone is searching for a friend, some top-down mechanism must exist that allows one to suppress irrelevant visual information while enhancing task-relevant information, allowing for an efficient yet effective search. In this way, the maintenance and representation of the goal (e.g., find your friend) by the PFC serve as bias signals. As Miller & Cohen (2001) have stated, “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” (p. 167). As described earlier in this review, given that the PFC represents rules and goals at multiple levels of abstraction, it is in an ideal position to influence processing in downstream brain regions that receive its anatomical projections.

We have used fMRI and event-related potentials (ERP) in humans to investigate such top-down mechanisms (Gazzaley et al. 2005). In this study, during each trial of a working memory task participants observed sequences of two faces and two natural scenes presented in a randomized order. In separate blocks of trials, subjects were required to remember faces and ignore scenes, remember scenes and ignore faces, or passively view faces and scenes without attempting to remember them. Because each trial had equivalent bottom-up visual information (i.e., faces and scenes), we could directly determine if top-down signals were engaged. Moreover, the inclusion of a passive baseline allowed for the dissociation of possible enhancement and suppression mechanisms. With both fMRI and ERP, we obtained activity measures from areas of the visual association cortex specialized in face and scene processing. For fMRI, we used an independent functional localizer to identify both stimulus-selective face regions [within the fusiform face area (FFA); Kanwisher et al. 1997] and scene regions [within the parahippocampal place area (PPA); Epstein & Kanwisher 1998]. For ERP, we utilized a face-selective ERP, the N₁₇₀, a component localized to posterior occipital electrodes, which reflect visual association cortex activity with face specificity (Bentin et al. 1996). Our fMRI and ERP data revealed top-down modulation of both activity magnitude and processing speed that occurred above and below the perceptual baseline, depending on task instruction. That is, during the encoding period of the delay task, FFA activity was enhanced, and the N₁₇₀ occurred earlier, when faces had to be remembered as compared with a condition where they were passively viewed. Likewise, FFA activity was suppressed, and the N₁₇₀ occurred later, when faces had to be ignored compared with a condition where they were passively viewed. These results suggest that there are at least two types of top-down signals: One serves to enhance task-relevant information, and the other serves to suppress task-relevant information. By generating contrast via enhancing and suppressing activity magnitude and processing speed, top-down signals can bias the likelihood of successful representation of relevant information in a competitive system (Corbetta et al. 1990, Hillyard et al. 1973, Moran & Desimone 1985).

With fMRI or any type of neurophysiological method applied to animals or humans, there is no direct way to determine the source of top-down signals. Thus, to obtain evidence that the PFC

is the source of top-down signals that modulate the visual association cortex, the physiological responses of the visual association cortex must be examined after disruption of PFC function (Miller & D'Esposito 2005). The first attempt at such an approach was performed by Fuster and colleagues (1985), who studied the effect of PFC inactivation in monkeys by cooling on spiking activity in inferotemporal cortex neurons during a delayed-match-to-sample color task. During the delay interval in this task, when persistent stimulus-specific activity in inferotemporal cortical neurons is observed, PFC inactivation caused attenuated spiking activity and a loss of stimulus specificity of inferotemporal cortical neurons. These two alterations of inferotemporal cortex activity strongly implicated the PFC as a source of top-down signals necessary for maintaining robust sensory representations in the absence of bottom-up sensory activity.

Many years passed before any other attempt was made with animals or humans to follow up this landmark finding by Fuster. In fact, the combined lesion/electrophysiological approach is rarely implemented even today. Translating this approach to humans, Chao & Knight (1998) investigated patients with lateral PFC lesions during delayed match-to-sample tasks. They found that when distracting stimuli are presented during the delay period, the amplitude of the ERP recorded from posterior electrodes was markedly increased in patients with frontal lesions compared with controls. Investigators interpreted that these results demonstrated disinhibition of sensory processing, which supports a role for the PFC in suppressing the representation of task-irrelevant stimuli. Recently, we investigated the causal role of the PFC in the modulation of evoked activity in the human extrastriate cortex during the encoding of faces and scenes (Miller et al. 2011). We employed two experimental approaches to disrupt PFC function: TMS of the PFC in healthy subjects and focal PFC lesions in stroke patients. We then investigated the effect of disrupted PFC function on the selectivity of category representations (faces or scenes) in the temporal cortex. Different object categories, such as faces and scenes, are represented by spatially distributed yet overlapping assemblies in the extrastriate visual cortex (Haxby et al. 2001). Thus, we reasoned that disruption of PFC function would lead to higher spatial correlations between scene- and face-evoked activity in the extrastriate cortex, suggesting a decrease in category selectivity. Consistent with our predictions, following disruption of PFC function (i.e., TMS session versus baseline, or lesion versus intact hemisphere in stroke patients), stimulus-evoked activity in the extrastriate cortex exhibited less distinct category selectivity to faces and scenes (more spatial overlap). In a follow-up study (Lee & D'Esposito 2012), we further demonstrated that the decreased tuning of the extrastriate cortex response coincided with decrements in working memory performance. This work extended the findings of Fuster and colleagues (1985) from monkeys to humans and suggests that the PFC may sharpen the representations of different object categories in the extrastriate cortex by increasing the distinctiveness of their distributed neural representations. These findings are also consistent with other recent combined TMS/fMRI and TMS/EEG studies demonstrating decreased attentional modulation of stimulus-selective visual regions following PFC disruption (Feredoes et al. 2011, Higo et al. 2011, Zanto et al. 2011). Together, such causal evidence clearly supports the notion that the PFC is the source of top-down signals that act via both gain and selectivity mechanisms.

A key to understanding the role of the PFC in cognition likely rests in its connectivity with other regions (Yeterian et al. 2012). Any top-down signal from a particular PFC region, representing a particular goal, could have a different influence and behavioral consequence depending on which brain regions receive these signals. For example, PFC top-down signals could enhance internal representations of relevant sensory stimuli in the extrastriate cortex or anticipated motor plans in the premotor cortex. It is likely that multiple top-down signals are engaged in a parallel fashion during the evolution of any goal-directed behavior. Moreover, studies propose that other cortical regions, such as the parietal cortex and the hippocampus, also provide top-down signals

during cognition (Eichenbaum 2013, Ruff 2013). Consideration of the mechanisms by which multiple higher-order brain regions can influence lower-order brain regions highlights the enormous complexity of the human brain and how much further we must travel to understand it.

Long-range connectivity. Another mechanism critical for working memory is the synchronization of activity among distributed brain regions. Because the available methodology is limited in both animals and humans, few studies to date have been able to assess if and how neurons and brain regions communicate and interact to support working memory. We developed a multivariate method designed specifically to characterize functional connectivity in event-related fMRI data that can measure interregional correlations during the individual stages of a cognitive task (Rissman et al. 2004). Using this method, we specifically sought to characterize the network of brain regions associated with maintaining a representation of face stimuli over a short delay interval. With this approach (Gazzaley et al. 2004), we found significant functional connectivity between the FFA and the PFC and the parietal cortex during the delay period of the task, which supports the notion that higher-order association cortices interact with posterior sensory regions to facilitate the active maintenance of a sensory percept. We have also found that posterior language-related areas involved in maintaining words in the absence of visual input also exhibit increased functional connectivity with the PFC (Fiebach et al. 2006).

Distributed synchronized activity could occur via synaptic reverberations in recurrent circuits (Durstewitz et al. 2000a, Wang 1999) or synchronous oscillations between neuronal populations (Buzsáki & Draguhn 2004, Fries 2005, Singer 2009). In humans, EEG, magnetoencephalographic (MEG), and electrocorticographic (ECoG) recordings have been used to investigate which particular frequencies of oscillations may be related to working memory. Activity in low and high frequencies in the theta (4–7 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–200 Hz) ranges was modulated during working memory tasks (for a comprehensive review of 26 studies, see Roux & Uhlhaas 2014). Roux & Uhlhaas (2014) have proposed a different functional role for each of these frequency bands. They propose that gamma-band oscillations are specifically involved in the active maintenance of working memory information, theta-band oscillations are specifically involved in the temporal organization of working memory items, and alpha-band oscillations are involved in the inhibition of task-irrelevant information. These notions are based on studies that have demonstrated amplitude modulation of neural oscillations presumably emanating from particular brain regions involved in working memory. For example, during a delayed match-to-sample task while recording human EEG, investigators observed that occipital gamma and frontal beta oscillations were sustained across the retention interval. Moreover, as this delay interval lengthened, these oscillations decreased in parallel with decreased performance on the task (Tallon-Baudry et al. 1999). In a recent study, Anderson et al. (2014) showed that the spatial distribution of power in the alpha frequency band (8–12 Hz) tracked both the content and the quality of the representations stored in visual working memory. These empirical findings support the notion that neural oscillations are critical for working memory maintenance processes.

Long-range synchronization of these oscillations between brain regions likely also plays an important role in working memory function (Crespo-Garcia et al. 2013, Sauseng et al. 2005). For example, in a human MEG study, synchronized oscillations in the alpha, beta, and gamma bands were observed between frontoparietal and visual areas during the retention interval of a delayed match-to-sample visual working memory task. Moreover, these observed synchronized oscillations were sustained and stable throughout the delay period of the task, were memory load dependent, and were correlated with an individual's working memory capacity (Palva et al. 2010). Monkey physiology data have also provided considerable insight into the possible mechanisms underlying communication between brain regions during working memory. For example, in one study (Liebe

et al. 2012), neural interactions between visual area V4 and the lateral PFC were investigated during the performance of a visual delayed match-to-sample task. During the retention interval of the task, these two areas exhibited synchronization of local field potentials in the theta frequencies. Moreover, neuronal spiking activity in these two regions was phase-locked to these observed theta oscillations. Most importantly, the strength of this intercortical locking predicted the animal's performance, that is, higher for subsequently correctly remembered stimuli and session-to-session variability in memory performance. The authors concluded that these findings reflect a mechanism for effective communication between brain regions involved in the temporary maintenance of relevant visual information, an idea corroborated by other researchers (Fell & Axmacher 2011, Fries 2005). An intriguing recent finding suggests a critical role for the thalamus in regulating information transmission across cortical regions, at least at the local level (Saalmann et al. 2012).

Brain stem neuromodulators. In many models of cognition, neuromodulators such as dopamine, serotonin, norepinephrine, or acetylcholine play a limited role, if any role at all. Yet, given that brain stem neuromodulatory neurons send projections to all areas of the brain, their influence on cognitive function is without question. Abundant evidence from both animal and human studies indicates that dopaminergic modulation of frontostriatal circuitry in particular is critical for working memory function (Cools & D'Esposito 2009).

Dopaminergic neurons in the human brain are organized into several major subsystems (mesocortical, mesolimbic, and nigrostriatal). The mesocortical and mesolimbic dopaminergic systems originate in the ventral tegmental area of the midbrain and project to the frontal cortex, the anterior cingulate, the nucleus accumbens, and the anterior temporal structures such as the amygdala, hippocampus, and entorhinal cortex (Bannon & Roth 1983). Across the cerebral cortex, the concentration of dopamine is highest within the frontal cortex (Brown et al. 1979, Williams & Goldman-Rakic 1993). However, there is also a strong dopaminergic input into the hippocampus (Samson et al. 1990), and abundant evidence from both animal and human studies shows that dopamine is involved in hippocampal-dependent LTM (for a review of this topic, see Shohamy & Adcock, 2010).

The functional importance of dopamine to working memory and PFC function has been demonstrated in several ways. First, in monkeys, depletion of PFC dopamine or pharmacological blockade of dopamine receptors induces working memory deficits (Brozoski et al. 1979, Sawaguchi & Goldman-Rakic 1991). These deficits are as severe as those in monkeys with PFC lesions and are not observed in monkeys in which other neurotransmitters, such as serotonin, are depleted. Furthermore, dopaminergic agonists administered to monkeys with dopamine depletion reverse working memory deficits (Arnsten et al. 1994, Brozoski et al. 1979). Likewise, numerous studies have shown that administration of dopamine receptor agonists to healthy young human subjects improves working memory performance (Kimberg et al. 1997, Kimberg & D'Esposito 2003, Luciana & Collins 1997, Muller et al. 1998). An important feature of the dopaminergic system is that it exhibits an inverted U-shaped dose-response curve: Specific doses of dopaminergic drugs produce optimal performance on working memory tasks (Arnsten 1997, Kimberg et al. 1997; reviewed in detail in Cools & D'Esposito 2011). These observations illustrate that more is not better; rather, an optimal brain dopamine concentration is necessary for optimal working memory function.

Different classes of dopamine receptors exist in varying concentrations throughout the brain. D₂ dopamine receptors are present in much lower concentrations in the cortex than are D₁ receptors and are found mostly within the striatum (Camps et al. 1989). However, D₂ receptors are at their highest concentrations in the PFC (Goldman-Rakic et al. 1990). Moreover, dopamine release in the brain can be either transient (phasic) or sustained (tonic). Grace (2000) has proposed that these two dopamine mechanisms of action are functionally distinct and antagonistic.

Specifically, studies propose that tonic dopamine release is mediated by D_1 receptors, whereas D_2 receptor-mediated effects are phasic. Likewise, during a working memory task in monkeys, a dopamine D_2 receptor agonist selectively modulated the phasic component of the task, yet it had little effect on the persistent mnemonic-related activity, which was instead modulated by a D_1 receptor agonist (Sawaguchi 2001, Wang et al. 2004). Thus, these two dopamine receptors likely have complementary functions, which serve to modulate active memory representations stored within the PFC (Cohen et al. 2002). The dual-state theory of PFC dopamine function put forward by Durstewitz & Seamans (2008) states that a D_1 -dominated state favors robust online maintenance of information, whereas a D_2 -dominated state is beneficial for flexible and fast switching among representational states.

Regarding working memory function, Braver & Cohen (1999) proposed that tonic dopamine effects may increase the stability of maintained representations, whereas phasic dopamine effects may serve as gating signals to indicate when new inputs should be encoded and maintained or when currently maintained representations should be updated. In this way, two separate mechanisms underlie cognitive flexibility and stability that nevertheless must work together: Dopamine would promote stability or flexibility of maintained representations depending on the neural site of modulation (Cools & Robbins 2004). Specifically, dopamine receptor stimulation in the PFC would promote stability by increasing distractor resistance (Durstewitz et al. 2000b). Conversely, dopamine receptor stimulation in the striatum would promote flexibility by allowing newly relevant representations to update (Bilder et al. 2004, Frank et al. 2001). In the context of real-world situations, demands for cognitive flexibility and stability are reciprocal: If we are too flexible, we are likely to become distracted; if we are too stable, we become inflexible and unresponsive to new information.

We have tested this dopaminergic model of working memory with a human pharmacological fMRI study (Cools et al. 2007). Healthy young subjects underwent fMRI scanning on two occasions, once after intake of the dopaminergic agonist bromocriptine and once after placebo (in a double-blind, crossover design). During scanning, subjects performed a working memory task that allowed researchers to study working memory updating and maintenance processes separately. Specifically, subjects had to encode, maintain, and retrieve visual stimuli over a short delay. Two faces and two scenes were always presented during the encoding period, and subjects were instructed to remember either the faces or the scenes. During the retention period, another stimulus was presented, which subjects were instructed to ignore. This distractor was either a scrambled image or a novel face or scene. The critical measure of working memory updating was the behavioral switch cost, which was calculated by subtracting performance (error rates and reaction times measured at probe) on trials where subjects switched to a new instruction as compared with using the existing instruction. The critical measure of working memory maintenance was the behavioral distractor cost, which was calculated by subtracting performance (measured at probe) after scrambled as compared with nonscrambled distractors. We predicted that bromocriptine would modulate PFC activity during the epoch of the task following distraction but that the striatum would be modulated during the instruction epoch. This is exactly what we observed, which corroborates the hypothesis that working memory maintenance and updating processes are modulated by differential dopaminergic stimulation of the PFC and the striatum, respectively. This finding suggests that high levels of dopamine within the PFC (and lower levels in the striatum) optimize the maintenance of task-relevant representations, whereas high levels of dopamine within the striatum (and low levels in the PFC) optimize the flexible updating of information (for a more detailed review of dopaminergic functions, see Cools & D'Esposito 2009). The functional opponency between stability and flexibility of working memory representations maps well onto the neurochemical reciprocity between dopamine in the PFC and that in the striatum: Increases

and decreases in PFC dopamine lead to decreases and increases in striatal dopamine, respectively (Akil et al. 2003, Meyer-Lindenberg et al. 2005, Pycocock et al. 1980).

A working memory gate provides a computationally efficient mechanism: The gate allows information necessary for goal-directed behavior to be updated when it is open, but when it is closed it keeps irrelevant information out and allows current information to be sustained (Badre 2012, Frank & O'Reilly 2006). Using high-resolution MRI of the midbrain, D'Ardenne et al. (2008) demonstrated activation in a region likely comprising the substantia nigra and ventral tegmental area during trials on a task that required working memory updating. Midbrain activity also correlated with PFC activity as well as with behavior. These findings support the idea that dopamine acts as a gating signal to the PFC when updating of maintained representations is required. Recently, Badre and Frank have provided computational and empirical evidence for the possible mechanisms underlying working memory gating (Badre & Frank 2012, Chatham & Badre 2013, Frank & Badre 2012). Specifically, as a refinement of the original O'Reilly & Frank model, which proposed that the striatum can deliver selective gating inputs into the PFC, Frank and Badre propose two types of striatal gating signals. The first type provides gating of inputs to be maintained by the frontal cortex (input gating), and the second type of gating signal determines which of these maintained representations will have an influence on particular selected actions (output gating). Selective gating (rather than a global mechanism arising from midbrain dopaminergic input that would update everything) allows for some information to be maintained by the PFC while other information is updated. The idea of selective striatal gating also allows for a hierarchy within frontostriatal circuitry such that contextual representations in the rostral frontal cortex can influence striatal gating of contextual representations in the caudal frontal cortex. An MRI study using diffusion tractography has demonstrated that the proper wiring is in place for such a mechanism in that there is a rostral-caudal correspondence in the connectivity pattern between frontal and striatal regions (Verstynen et al. 2012).

CONCLUSIONS

Working memory is a construct that has motivated research in many domains—cognitive, neuroscientific, clinical—for the past 50 years. The cumulative results from this half-century of research have reinforced the centrality, articulated in seminal writings from the 1960s, 1970s, and 1980s, of working memory in behavior control. The past decade has witnessed many exciting advances in our understanding of the mechanisms that underlie working memory, and these have necessarily prompted the near-continuous updating of working memory models. At a broader level, however, one could make the case that our current neural systems-level models were foreshadowed by a core feature of the Baddeley & Hitch (1974) multiple-component model: the important distinction between stimulus representation, on the one hand, and the control of behavior with those representations, on the other. Baddeley has always purported that his construal of the central executive of the multiple-component model was something akin to Shallice's supervisory attentional system, that is, a control system that was not in any sense specialized for or dedicated to working memory operations but one that could use and/or manipulate the contents of working memory storage to more effectively guide behavior. The prefrontal, basal ganglia, thalamic, and brain stem systems reviewed here can be construed as a neural substrate for this central executive. We believe that a conceptual error at the root of some of the systems- and cognitive-neuroscience research from the 1980s–2000s derived from a misattribution of PFC activity to the functioning of one of the storage buffers from the multicomponent model rather than to the central executive. The research we have reviewed here makes it clear that the functions of the PFC (and related systems) are too flexible and operate on a level too abstract to be construed as simply performing a buffering role.

The past ten years have also witnessed considerable progress in our understanding of how the function of buffering is accomplished in the primate brain. In digital computers, this function is carried out by random access memory (RAM), circuitry that is physically distinct from hard drive storage and is specialized for and dedicated to this role. The analogy to computer architecture may have, at least implicitly, influenced previous thinking about biological working memory. What recent research has established, however, is that there are no dedicated RAM circuits in the primate brain. Rather, the operation of holding information in working memory occurs within the same circuits that process that information in nonmnemonic contexts. For symbolic information, this concept has been demonstrated in models of activated semantic LTM, whereas for sensorimotor information, it is demonstrated in sensorimotor recruitment models.

In this review, we have emphasized the fundamental importance of working memory for cognitive control. We believe that any understanding of the basic mechanisms of working memory leads directly to a further understanding of the most complex aspects of human cognition. The frontal cortex continues to be a primary area of focus in attempts to uncover the neural mechanisms that support component processes necessary for cognitive control. The frontal cortex is hierarchically organized and provides critical bias signals that sculpt goal-directed behavior. Much work is still needed regarding the nature of these signals and the mechanisms by which the frontal cortex maintains relevant information and communicates with other brain regions. Moreover, ascending brain stem neuromodulatory systems, such as the dopaminergic system, likely influence most of the cognitive processes mentioned above. A consideration of all these mechanisms together, rather than in isolation, should provide a clearer picture of the neural bases of cognitive control.

SUMMARY POINTS

1. An enduring principle of the multiple-component model of working memory (Baddeley & Hitch 1974) is that the short-term retention of information (also known as working memory storage) and the control of how that information is used to guide behavior are subserved by distinct processes. With regard to the former, however, earlier ideas of specialized buffers have been largely superseded by state-based models.
2. Although state-based models of working memory storage are often categorized as activated LTM models or sensorimotor recruitment models, all are grounded in the idea that the attentional selection of mental representations brings them into working memory and that the consequences of attentional prioritization explain such properties as capacity limitations, proactive interference from no-longer-relevant items, etc.
3. Recent research applying MVPA to fMRI and EEG data has provided compelling neural evidence for state-based models of working memory storage.
4. Some recent findings from computational modeling, extracellular electrophysiology, fMRI, and EEG suggest that working memory storage may depend on the transient reorganization of synaptic weights rather than on sustained, elevated activity.
5. The PFC likely represents higher-order information, such as task rules, goals, or abstract representations of categories, as compared with feature- and stimulus-specific representations in the posterior cortex. Moreover, a critical mechanism for working memory function is the synchronization of PFC activity with activity in other brain regions.

6. One dimension of functional organization of the PFC is a hierarchical caudal to rostral gradient of the level of abstraction of the rules and goals that guide behavior.
7. Top-down control signals emanating from the PFC likely take at least two forms: signals that modulate gain by either enhancing task-relevant information or suppressing task-irrelevant information, and signals that can modulate the selectivity of information represented in posterior cortical regions.
8. Dopamine plays a critical role in working memory function. The complex interplay of midbrain dopamine in prefrontal and striatal circuits underlies tonic maintenance and phasic gating functions that govern the balance between cognitive flexibility and stability.

FUTURE ISSUES

1. How is the FoA organized? Does it have a strict capacity limit of one item, or can it contain multiple items? Are there multiple distinct levels within the FoA (or levels of activation within working memory), or is everything outside a unitary FoA in the same state of LTM?
2. Which class of models better accounts for capacity limitations in visual STM: slots models, single-resource models, a hybrid of the two, or some as-yet-to-be-described alternative?
3. Because recent MVPA studies have dissociated working memory storage from sustained, elevated delay-period activity, what functions do the latter subserve?
4. Is it possible, as suggested by recent experiments, that all delay-period activity that is decodable with MVPA, even activity that is below univariate statistical thresholds, corresponds to the FoA rather than to the storage of information per se? If so, is the latter accomplished via the transient reorganization of synaptic weights?
5. Is the high dimensionality that has been ascribed to ensembles of PFC neurons a property that is unique to that region, or is the property also characteristic of other brain regions?
6. What are the different functional roles of particular frequencies of oscillations [e.g., theta (4–7 Hz), alpha (8–13 Hz), and gamma (30–200 Hz)] for working memory?
7. Does dopamine play a similar role in both input and output working memory gating signals?
8. In addition to dopamine, what roles do other neurotransmitters and hormones play in working memory function?

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

In this article, we have drawn from the original ideas and empirical work of many of our trainees whom we sincerely acknowledge, including, but not limited to, David Badre, Brad

Buchsbaum, Roshan Cools, Clay Curtis, Adam Gazzaley, Joshua LaRocque, Derek Nee, Jarrod Lewis-Peacock, Jesse Rissman, Kartik Sreenivasan, Charan Ranganath, Adam Riggall, and Bart Rypma. We also acknowledge the generous funding we have received over the years from the National Institutes of Health.

LITERATURE CITED

- Akil M, Kolachana BS, Rothmond DA, Hyde TM, Weinberger DR, Kleinman JE. 2003. Catechol-O-methyltransferase genotype and dopamine regulation in the human brain. *J. Neurosci.* 23:2008–13
- Alvarez GA, Cavanagh P. 2004. The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychol. Sci.* 15:106–11
- Anderson DE, Serences JT, Vogel EK, Awh E. 2014. Induced alpha rhythms track the content and quality of visual working memory representations with high temporal precision. *J. Neurosci.* 34:7587–99
- Arnsten A. 1997. Catecholamine regulation of the prefrontal cortex. *J. Psychopharmacol.* 11:151–62
- Arnsten KT, Cai JX, Murphy BL, Goldman-Rakic PS. 1994. Dopamine D1 receptor mechanisms in the cognitive performance of young adult and aged monkeys. *Psychopharmacology* 116:143–51
- Awh E, Jonides J. 2001. Overlapping mechanisms of attention and spatial working memory. *Trends Cogn. Sci.* 5:119–26
- Awh E, Jonides J, Reuter-Lorenz PA. 1998. Rehearsal in spatial working memory. *J. Exp. Psychol.: Hum. Percept. Perform.* 24:780–90
- Azuar C, Reyes P, Slachevsky A, Volle E, Kinkingnehun S, et al. 2014. Testing the model of caudo-rostral organization of cognitive control in the human with frontal lesions. *NeuroImage* 84:1053–60
- Baddeley A. 1986. *Working Memory*. New York: Oxford Univ. Press
- Baddeley A, Hitch GJ. 1974. Working memory. In *Psychology of Learning and Motivation*, ed. G Bower, Vol. 8, pp. 47–89. New York: Academic Press**
- Badre D. 2008. Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends Cogn. Sci.* 12:193–200
- Badre D. 2012. Opening the gate to working memory. *Proc. Natl. Acad. Sci. USA* 109:19878–79
- Badre D, D’Esposito M. 2007. Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *J. Cogn. Neurosci.* 19:2082–99
- Badre D, D’Esposito M. 2009. Is the rostro-caudal axis of the frontal lobe hierarchical? *Nat. Rev. Neurosci.* 10:659–69**
- Badre D, Frank MJ. 2012. Mechanisms of hierarchical reinforcement learning in cortico-striatal circuits 2: evidence from fMRI. *Cereb. Cortex* 22:527–36
- Badre D, Hoffman J, Cooney JW, D’Esposito M. 2009. Hierarchical cognitive control deficits following damage to the human frontal lobe. *Nat. Neurosci.* 12:515–22
- Bannon MJ, Roth RH. 1983. Pharmacology of mesocortical dopamine neurons. *Pharmacol. Rev.* 35:53–68
- Barbas H, Pandya DN. 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, H Eisenberg, AL Benton, pp. 35–58. Oxford, UK: Oxford Univ. Press
- Bays PM, Husain M. 2008. Dynamic shifts of limited working memory resources in human vision. *Science* 321:851–54
- Bentin S, Allison T, Puce A, Perez E, McCarthy G. 1996. Electrophysiological studies of face perception in humans. *J. Cogn. Neurosci.* 8:551–65
- Bilder RM, Volavka J, Lachman HM, Grace AA. 2004. The catechol-O-methyltransferase polymorphism: relations to the tonic-phasic dopamine hypothesis and neuropsychiatric phenotypes. *Neuropsychopharmacology* 29:1943–61
- Blumenfeld RS, Nomura EM, Gratton C, D’Esposito M. 2013. Lateral prefrontal cortex is organized into parallel dorsal and ventral streams along the rostro-caudal axis. *Cereb. Cortex* 23:2457–66
- Braver TS, Cohen JD. 1999. Dopamine, cognitive control, and schizophrenia: the gating model. *Prog. Brain Res.* 121:327–49

The paper that introduces the highly influential multiple-component model of working memory.

A synthesis of the evidence that the rostral-caudal functional gradient observed along the frontal cortex is hierarchical.

- Braver TS, Gray JR, Burgess GC. 2008. Explaining the many varieties of working memory variation: dual mechanisms of cognitive control. In *Variation in Working Memory*, ed. ARA Conway, C Jarrold, MJ Kane, A Miyake, JN Towse, pp. 76–106. Oxford, UK: Oxford Univ. Press
- Brown RM, Crane AM, Goldman PS. 1979. Regional distribution of monoamines in the cerebral cortex and subcortical structures of the rhesus monkey: concentrations and in vivo synthesis rates. *Brain Res.* 168:133–50
- Brozoski TJ, Brown RM, Rosvold HE, Goldman PS. 1979. Cognitive deficit caused by regional depletion of dopamine in prefrontal cortex of rhesus monkey. *Science* 205:929–32
- Burgess PW, Dumontheil I, Gilbert SJ. 2007. The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends Cogn. Sci.* 11:290–98
- Buzsáki G, Draguhn A. 2004. Neuronal oscillations in cortical networks. *Science* 304:1926–29
- Camps M, Cortés R, Gueye B, Probst A, Palacios JM. 1989. Dopamine receptors in human brain: autoradiographic distribution of D₂ sites. *Neuroscience* 28:275–90
- Chao LL, Knight RT. 1998. Contribution of human prefrontal cortex to delay performance. *J. Cogn. Neurosci.* 10:167–77
- Chatham CH, Badre D. 2013. Working memory management and predicted utility. *Front. Behav. Neurosci.* 7:83
- Chen AJ, Britton M, Turner GR, Vytlačil J, Thompson TW, D’Esposito M. 2012. Goal-directed attention alters the tuning of object-based representations in extrastriate cortex. *Front. Hum. Neurosci.* 6:187
- Christoff K, Ream JM, Geddes LP, Gabrieli JD. 2003. Evaluating self-generated information: anterior prefrontal contributions to human cognition. *Behav. Neurosci.* 117:1161–68
- Christophel TB, Hebart MN, Haynes JD. 2012. Decoding the contents of visual short-term memory from human visual and parietal cortex. *J. Neurosci.* 32:12983–89
- Cohen JD, Braver TS, Brown JW. 2002. Computational perspectives on dopamine function in prefrontal cortex. *Curr. Opin. Neurobiol.* 12:223–29
- Constantinidis C, Franowicz MN, Goldman-Rakic PS. 2001. The sensory nature of mnemonic representation in the primate prefrontal cortex. *Nat. Neurosci.* 4:311–16
- Cools R, D’Esposito M. 2009. Dopaminergic modulation of flexible control in humans. In *Dopamine Handbook*, ed. A Bjorklund, SB Dunnett, LL Iversen, SD Iversen, pp. 249–60. Oxford, UK: Oxford Univ. Press
- Cools R, D’Esposito M. 2011. Inverted-U-shaped dopamine actions on human working memory and cognitive control. *Biol. Psychiatry* 69:e113–25**
- Cools R, Robbins TW. 2004. Chemistry of the adaptive mind. *Philos. Transact. A Math. Phys. Eng. Sci.* 362:2871–88
- Cools R, Sheridan M, Jacobs E, D’Esposito M. 2007. Impulsive personality predicts dopamine-dependent changes in frontostriatal activity during component processes of working memory. *J. Neurosci.* 27:5506–14
- Corbetta M, Miezin FM, Dobmeyer S, Shulman GL, Petersen SE. 1990. Attentional modulation of neural processing of shape, color, and velocity in humans. *Science* 248:1556–59
- Courtney SM, Ungerleider LG, Keil K, Haxby JV. 1997. Transient and sustained activity in a distributed neural system for human working memory. *Nature* 386:608–11
- Cowan N. 1995. *Attention and Memory: An Integrated Framework*. New York: Oxford Univ. Press
- Crespo-García M, Pinal D, Cantero JL, Díaz F, Zurrón M, Atienza M. 2013. Working memory processes are mediated by local and long-range synchronization of alpha oscillations. *J. Cogn. Neurosci.* 25:1343–57
- Cummings JL. 1993. Frontal-subcortical circuits and human behavior. *Arch. Neurol.* 50:873–80
- Curtis CE, Rao VY, D’Esposito M. 2004. Maintenance of spatial and motor codes during oculomotor delayed response tasks. *J. Neurosci.* 24:3944–52
- D’Ardenne K, McClure SM, Nystrom LE, Cohen JD. 2008. BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science* 319:1264–67
- D’Esposito M. 2007. From cognitive to neural models of working memory. *Philos. Trans. R. Soc. B* 362:761–72
- D’Esposito M, Postle B, Rypma B. 2000. Prefrontal cortical contributions to working memory: evidence from event-related fMRI studies. *Exp. Brain Res.* 133:3–11
- Devinsky O, D’Esposito M. 2003. *Neurology of Cognitive and Behavioral Disorders*. New York: Oxford Univ. Press

A review of evidence from studies with experimental animals, healthy humans, and patients with Parkinson’s disease, which demonstrate that optimum levels of dopamine are necessary for successful cognitive control.

A powerful demonstration, with MVPA encoding models, that the precision of neural representations in sensory cortex determines the precision of STM.

A study using fMRI and ERP that provides converging evidence that both the magnitude of neural activity and the speed of neural processing are modulated by top-down influences.

- Duncan J. 2001. An adaptive coding model of neural function in prefrontal cortex. *Nat. Rev. Neurosci.* 2:820–29
- Durstewitz D, Seamans JK. 2008. The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-*O*-methyltransferase genotypes and schizophrenia. *Biol. Psychiatry* 64:739–49
- Durstewitz D, Seamans JK, Sejnowski TJ. 2000a. Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. *J. Neurophysiol.* 83:1733–50
- Durstewitz D, Seamans JK, Sejnowski TJ. 2000b. Neurocomputational models of working memory. *Nat. Neurosci.* 3(Suppl.):1184–91
- Eichenbaum H. 2013. Memory on time. *Trends Cogn. Sci.* 17:81–88
- Emrich SM, Riggall AC, Larocque JJ, Postle BR. 2013. Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. *J. Neurosci.* 33:6516–23
- Epstein R, Kanwisher N. 1998. A cortical representation of the local visual environment. *Nature* 392:598–601
- Erickson MA, Maramba LA, Lisman J. 2010. A single brief burst induces GluR1-dependent associative short-term potentiation: a potential mechanism for short-term memory. *J. Cogn. Neurosci.* 22:2530–40
- Ester EF, Anderson DE, Serences JT, Awh E. 2013. A neural measure of precision in visual working memory. *J. Cogn. Neurosci.* 25:754–61**
- Fell J, Axmacher N. 2011. The role of phase synchronization in memory processes. *Nat. Rev. Neurosci.* 12:105–18
- Feredoes E, Heinen K, Weiskopf N, Ruff C, Driver J. 2011. Causal evidence for frontal involvement in memory target maintenance by posterior brain areas during distracter interference of visual working memory. *Proc. Natl. Acad. Sci. USA* 108:17510–15
- Fiebach CJ, Rissman J, D'Esposito M. 2006. Modulation of inferotemporal cortex activation during verbal working memory maintenance. *Neuron* 51:251–61
- Frank MJ, Badre D. 2012. Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. *Cereb. Cortex* 22:509–26
- Frank MJ, Loughry B, O'Reilly RC. 2001. Interactions between frontal cortex and basal ganglia in working memory: a computational model. *Cogn. Affect. Behav. Neurosci.* 1:137–60
- Frank MJ, O'Reilly RC. 2006. A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. *Behav. Neurosci.* 120:497–517
- Freedman DJ, Riesenhuber M, Poggio T, Miller EK. 2001. Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* 291:312–16
- Freedman DJ, Riesenhuber M, Poggio T, Miller EK. 2003. A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *J. Neurosci.* 23:5235–46
- Fries P. 2005. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* 9:474–80
- Funahashi S, Bruce CJ, Goldman-Rakic PS. 1989. Mnemonic coding of visual space in the monkey's dorso-lateral prefrontal cortex. *J. Neurophysiol.* 61:331–49
- Fuster JM. 1990. Prefrontal cortex and the bridging of temporal gaps in the perception-action cycle. *Ann. N. Y. Acad. Sci.* 608:318–29; discussion 330–36
- Fuster JM. 2004. Upper processing stages of the perception-action cycle. *Trends Cogn. Sci.* 8:143–45
- Fuster JM. 2008. *The Prefrontal Cortex*. Oxford, UK: Elsevier
- Fuster JM, Alexander GE. 1971. Neuron activity related to short-term memory. *Science* 173:652–54
- Fuster JM, Bauer RH, Jervey JP. 1985. Functional interactions between inferotemporal and prefrontal cortex in a cognitive task. *Brain Res.* 330:299–307
- Gazzaley A, Cooney JW, McEvoy K, Knight RT, D'Esposito M. 2005. Top-down enhancement and suppression of the magnitude and speed of neural activity. *J. Cogn. Neurosci.* 17:507–17**
- Gazzaley A, Rissman J, D'Esposito M. 2004. Functional connectivity during working memory maintenance. *Cogn. Affect. Behav. Neurosci.* 4:580–99
- Goldman-Rakic PS. 1987. Circuitry of the prefrontal cortex and the regulation of behavior by representational knowledge. In *Handbook of Physiology: The Nervous System*, ed. F Plum, VB Mountcastle, pp. 373–417. Bethesda, MD: Am. Physiol. Soc.
- Goldman-Rakic PS. 1995. Cellular basis of working memory. *Neuron* 14:477–85
- Goldman-Rakic PS, Lidow MS, Gallagher DW. 1990. Overlap of dopaminergic, adrenergic, and serotonergic receptors and complementarity of their subtypes in primate prefrontal cortex. *J. Neurosci.* 10:2125–38

- Grace AA. 2000. The tonic/phasic model of dopamine system regulation and its implications for understanding alcohol and psychostimulant craving. *Addiction* 95(Suppl. 2):S119–28
- Hamidi M, Tononi G, Postle BR. 2008. Evaluating frontal and parietal contributions to spatial working memory with repetitive transcranial magnetic stimulation. *Brain Res.* 1230:202–10
- Hamidi M, Tononi G, Postle BR. 2009. Evaluating the role of prefrontal and parietal cortices in memory-guided response with repetitive transcranial magnetic stimulation. *Neuropsychologia* 47:295–302
- Han X, Berg AC, Oh H, Samaras D, Leung HC. 2013. Multi-voxel pattern analysis of selective representation of visual working memory in ventral temporal and occipital regions. *NeuroImage* 73:8–15
- Harrison SA, Tong F. 2009. Decoding reveals the contents of visual working memory in early visual areas. *Nature* 458:632–35
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293:2425–30
- Higo T, Mars RB, Boorman ED, Buch ER, Rushworth MFS. 2011. Distributed and causal influence of frontal operculum in task control. *Proc. Natl. Acad. Sci. USA* 108:4230–35
- Hillyard SA, Hink RF, Schwent VL, Picton TW. 1973. Electrical signs of selective attention in the human brain. *Science* 182:177–80
- Itskov V, Hansel D, Tsodyks M. 2011. Short-term facilitation may stabilize parametric working memory trace. *Front. Comput. Neurosci.* 5:40
- Jerde TA, Merriam EP, Riggall AC, Hedges JH, Curtis CE. 2012. Prioritized maps of space in human frontoparietal cortex. *J. Neurosci.* 32:17382–90
- Kanwisher N, McDermott J, Chun MM. 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17:4302–11
- Kimberg DY, D'Esposito M. 2003. Cognitive effects of the dopamine receptor agonist pergolide. *Neuropsychologia* 41:1020–27
- Kimberg DY, D'Esposito M, Farah MJ. 1997. Effects of bromocriptine on human subjects depend on working memory capacity. *NeuroReport* 8:3581–85
- Koechlin E, Ody C, Kouneiher F. 2003. The architecture of cognitive control in the human prefrontal cortex. *Science* 302:1181–85
- Koechlin E, Summerfield C. 2007. An information theoretical approach to prefrontal executive function. *Trends Cogn. Sci.* 11:229–35
- Kouneiher F, Charron S, Koechlin E. 2009. Motivation and cognitive control in the human prefrontal cortex. *Nat. Neurosci.* 12:939–45
- Kubota K, Niki H. 1971. Prefrontal cortical unit activity and delayed alternation performance in monkeys. *J. Neurophysiol.* 34:337–47
- LaRocque JJ, Lewis-Peacock JA, Drysdale AT, Oberauer K, Postle BR. 2013. Decoding attended information in short-term memory: an EEG study. *J. Cogn. Neurosci.* 25:127–42
- Lee SH, Kravitz DJ, Baker CI. 2013. Goal-dependent dissociation of visual and prefrontal cortices during working memory. *Nat. Neurosci.* 16:997–99
- Lee TG, D'Esposito M. 2012. The dynamic nature of top-down signals originating from prefrontal cortex: a combined fMRI-TMS study. *J. Neurosci.* 32:15458–66
- Lewis-Peacock JA, Drysdale AT, Oberauer K, Postle BR. 2012. Neural evidence for a distinction between short-term memory and the focus of attention. *J. Cogn. Neurosci.* 24:61–79
- Lewis-Peacock JA, Postle BR. 2008. Temporary activation of long-term memory supports working memory. *J. Neurosci.* 28:8765–71
- Lewis-Peacock JA, Postle BR. 2012. Decoding the internal focus of attention. *Neuropsychologia* 50:470–78
- Liebe S, Hoerzer GM, Logothetis NK, Rainer G. 2012. Theta coupling between V4 and prefrontal cortex predicts visual short-term memory performance. *Nat. Neurosci.* 15:456–62, S1–2
- Luciana M, Collins PF. 1997. Dopaminergic modulation of working memory for spatial but not object cues in normal humans. *J. Cogn. Neurosci.* 9:330–47
- Luck SJ, Vogel EK. 1997. The capacity of visual working memory for features and conjunctions. *Nature* 390:279–81
- Luck SJ, Vogel EK. 2013. Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends Cogn. Sci.* 17:391–400**

An authoritative
summary of evidence
supporting slots models
of STM capacity limits.

A comprehensive review of psychophysical and neural evidence for single-resource models of STM capacity limits.

An influential model of how the PFC implements top-down control for the flexible control of behavior.

A summary of the novel insights provided by MVPA, including the possibility that elevated activity reflects the focus of attention rather than working memory storage per se.

- Ma WJ, Husain M, Bays PM. 2014. Changing concepts of working memory. *Nat. Neurosci.* 17:347–56
- Magnussen S. 2000. Low-level memory processes in vision. *Trends Neurosci.* 23:247–51
- Magnussen S, Greenlee MW. 1999. The psychophysics of perceptual memory. *Psychol. Res.* 62:81–92
- McElree B. 1998. Attended and non-attended states in working memory: accessing categorized structures. *J. Mem. Lang.* 38:225–52
- McElree B. 2006. Accessing recent events. *Psychol. Learn. Motiv.* 46:155–200
- Meyer-Lindenberg A, Kohn PD, Kolachana B, Kippenhan S, McInerney-Leo A, et al. 2005. Midbrain dopamine and prefrontal function in humans: interaction and modulation by COMT genotype. *Nat. Neurosci.* 8:594–96
- Meyers EM, Freedman DJ, Kreiman G, Miller EK, Poggio T. 2008. Dynamic population coding of category information in inferior temporal and prefrontal cortex. *J. Neurophysiol.* 100:1407–19
- Miller BT, D'Esposito M. 2005. Searching for “the top” in top-down control. *Neuron* 48:535–38
- Miller BT, Vytalil J, Fegen D, Pradhan S, D'Esposito M. 2011. The prefrontal cortex modulates category selectivity in human extrastriate cortex. *J. Cogn. Neurosci.* 23:1–10
- Miller EK, Cohen JD. 2001. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24:167–202
- 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 GA, Galanter E, Pribham KH. 1960. *Plans and the Structure of Behavior*. New York: Holt, Rinehart and Winston
- Mongillo G, Barak O, Tsodyks M. 2008. Synaptic theory of working memory. *Science* 319:1543–46
- Moran J, Desimone R. 1985. Selective attention gates visual processing in the extrastriate cortex. *Science* 229:782–84
- Muller U, Von Cramon DY, Pollmann S. 1998. D1- versus D2-receptor modulation of visuospatial working memory in humans. *J. Neurosci.* 18:2720–28
- Nelissen N, Stokes M, Nobre AC, Rushworth MF. 2013. Frontal and parietal cortical interactions with distributed visual representations during selective attention and action selection. *J. Neurosci.* 33:16443–58
- Oberauer K. 2001. Removing irrelevant information from working memory: a cognitive aging study with the modified Sternberg task. *J. Exp. Psychol.: Learn. Mem. Cogn.* 27:948–57
- Oberauer K. 2002. Access to information in working memory: exploring the focus of attention. *J. Exp. Psychol.: Learn. Mem. Cogn.* 28:411–21
- Oberauer K. 2005. Control of the contents of working memory—a comparison of two paradigms and two age groups. *J. Exp. Psychol.: Learn. Mem. Cogn.* 31:714–28
- Oberauer K. 2009. Design for a working memory. *Psychol. Learn. Motiv.* 51:45–100
- Oberauer K. 2013. The focus of attention in working memory—from metaphors to mechanisms. *Front. Hum. Neurosci.* 7:673
- Palva JM, Monto S, Kulashkhar S, Palva S. 2010. Neuronal synchrony reveals working memory networks and predicts individual memory capacity. *Proc. Natl. Acad. Sci. USA* 107:7580–85
- Pesaran B, Pezaris JS, Sahani M, Mitra PP, Andersen RA. 2002. Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nat. Neurosci.* 5:805–11
- Petrides M. 2000. Dissociable roles of mid-dorsolateral prefrontal and anterior inferotemporal cortex in visual working memory. *J. Neurosci.* 20:7496–503
- Postle BR. 2006. Working memory as an emergent property of the mind and brain. *Neuroscience* 139:23–38
- Postle BR. 2011. What underlies the ability to guide action with spatial information that is no longer present in the environment? In *Spatial Working Memory*, ed. A Vandierendonck, A Szmalec, pp. 87–101. Hove, UK: Psychology
- Postle BR. 2015. The cognitive neuroscience of visual short-term memory. *Curr. Opin. Behav. Sci.* 1:40–46
- Postle BR, Idzikowski C, Sala SD, Logie RH, Baddeley AD. 2006. The selective disruption of spatial working memory by eye movements. *Q. J. Exp. Psychol. (Hove)* 59:100–20

- Pribram KH, Ahumada A, Hartog J, Roos L. 1964. A progress report on the neurological processes disturbed by frontal lesions in primates. In *The Frontal Cortex and Behavior*, ed. JM Warren, K Akert, pp. 28–55. New York: McGraw-Hill
- Pycock CJ, Kerwin RW, Carter CJ. 1980. Effect of lesion of cortical dopamine terminals on subcortical dopamine receptors in rats. *Nature* 286:74–76
- Ramnani N, Owen AM. 2004. Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. *Nat. Rev. Neurosci.* 5:184–94
- Riggall AC, Postle BR. 2012. The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. *J. Neurosci.* 32:12990–98
- Rigotti M, Barak O, Warden MR, Wang X-J, Daw ND, et al. 2013. The importance of mixed selectivity in complex cognitive tasks. *Nature* 497:585–90
- Rissman J, Gazzaley A, D'Esposito M. 2004. Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage* 23:752–63
- Romo R, Brody CD, Hernández A, Lemus L. 1999. Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature* 399:470–73
- Roux F, Uhlhaas PJ. 2014. Working memory and neural oscillations: alpha–gamma versus theta–gamma codes for distinct WM information? *Trends Cogn. Sci.* 18:16–25
- Ruff CC. 2013. Sensory processing: Who's in (top-down) control? *Ann. N. Y. Acad. Sci.* 1296:88–107
- Saalmann YB, Pinsk MA, Wang L, Li X, Kastner S. 2012. The pulvinar regulates information transmission between cortical areas based on attention demands. *Science* 337:753–56
- Samson Y, Wu JJ, Friedman AH, Davis JN. 1990. Catecholaminergic innervation of the hippocampus in the cynomolgus monkey. *J. Comp. Neurol.* 298:250–63
- Sauseng P, Klimesch W, Schabus M, Doppelmayr M. 2005. Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory. *Int. J. Psychophysiol.* 57:97–103
- Sawaguchi T. 2001. The effects of dopamine and its antagonists on directional delay-period activity of prefrontal neurons in monkeys during an oculomotor delayed-response task. *Neurosci. Res.* 41:115–28
- Sawaguchi T, Goldman-Rakic PS. 1991. D1 dopamine receptors in prefrontal cortex: involvement in working memory. *Science* 251:947–50
- Serences JT, Ester EF, Vogel EK, Awh E. 2009. Stimulus-specific delay activity in human primary visual cortex. *Psychol. Sci.* 20:207–14
- Shallice T. 1982. Specific impairments of planning. *Philos. Trans. R. Soc. B* 298:199–209
- Shohamy D, Adcock RA. 2010. Dopamine and adaptive memory. *Trends Cogn. Sci.* 14:464–72
- Silvanto J, Cattaneo Z. 2010. Transcranial magnetic stimulation reveals the content of visual short-term memory in the visual cortex. *NeuroImage* 50:1683–89
- Singer W. 2009. Distributed processing and temporal codes in neuronal networks. *Cogn. Neurodyn.* 3:189–96
- Smith EE, Jonides J. 1999. Storage and executive processes of the frontal lobes. *Science* 283:1657–61
- Sreenivasan KK, Curtis CE, D'Esposito M. 2014a. Revisiting the role of persistent neural activity during working memory. *Trends Cogn. Sci.* 18:82–89**
- Sreenivasan KK, Vytlačil J, D'Esposito M. 2014b. Distributed and dynamic storage of working memory stimulus information in extrastriate cortex. *J. Cogn. Neurosci.* 26:1141–53
- Stokes MG, Kusunoki M, Sigala N, Nili H, Gaffan D, Duncan J. 2013. Dynamic coding for cognitive control in prefrontal cortex. *Neuron* 78:364–75
- Sugase-Miyamoto Y, Liu Z, Wiener MC, Optican LM, Richmond BJ. 2008. Short-term memory trace in rapidly adapting synapses of inferior temporal cortex. *PLOS Comput. Biol.* 4:e1000073
- Tallon-Baudry C, Kreiter A, Bertrand O. 1999. Sustained and transient oscillatory responses in the gamma and beta bands in a visual short-term memory task in humans. *Vis. Neurosci.* 16:449–59
- Theeuwes J, Olivers CN, Chizk CL. 2005. Remembering a location makes the eyes curve away. *Psychol. Sci.* 16:196–99
- Venkatraman V, Rosati AG, Taren AA, Huettel SA. 2009. Resolving response, decision, and strategic control: evidence for a functional topography in dorsomedial prefrontal cortex. *J. Neurosci.* 29:13158–64
- Verstynen TD, Badre D, Jarbo K, Schneider W. 2012. Microstructural organizational patterns in the human corticostriatal system. *J. Neurophysiol.* 107:2984–95

A review of recent neural evidence for sensorimotor-recruitment models and for non-activity-based mechanisms for working memory storage.

- Vogel EK, Woodman GF, Luck SJ. 2001. Storage of features, conjunctions and objects in visual working memory. *J. Exp. Psychol.: Hum. Percept. Perform.* 27:92–114
- Wallis JD, Anderson KC, Miller EK. 2001. Single neurons in prefrontal cortex encode abstract rules. *Nature* 411:953–56
- Wang M, Vijayraghavan S, Goldman-Rakic PS. 2004. Selective D2 receptor actions on the functional circuitry of working memory. *Science* 303:853–56
- Wang XJ. 1999. Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. *J. Neurosci.* 19:9587–603
- Wang XJ. 2001. Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci.* 24:455–63
- Warden MR, Miller EK. 2010. Task-dependent changes in short-term memory in the prefrontal cortex. *J. Neurosci.* 30:15801–10
- Williams SM, Goldman-Rakic PS. 1993. Characterization of the dopaminergic innervation of the primate frontal cortex using a dopamine-specific antibody. *Cereb. Cortex* 3:199–222
- Yeterian EH, Pandya DN, Tomaiuolo F, Petrides M. 2012. The cortical connectivity of the prefrontal cortex in the monkey brain. *Cortex* 48:58–81
- Zaksas D, Bisley JW, Pasternak T. 2001. Motion information is spatially localized in a visual working-memory task. *J. Neurophysiol.* 86:912–21
- Zanto TP, Rubens MT, Thangavel A, Gazzaley A. 2011. Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. *Nat. Neurosci.* 14:656–61
- Zarahn E, Aguirre G, D'Esposito M. 1997. A trial-based experimental design for fMRI. *NeuroImage* 6:122–38