Working Memory as an Emergent Property of the Mind and Brain

Bradley R. Postle
Dept. of Psychology, University of Wisconsin-Madison

Abstract
Cognitive neuroscience research on working memory has been largely motivated by a standard model that arose from the melding of psychological theory with neuroscience data. Among the tenets of this standard model are that working memory functions arise from the operation of specialized systems that act as buffers for the storage and manipulation of information, and that frontal cortex (particularly prefrontal cortex) is a critical neural substrate for these specialized systems. However, the standard model has been a victim of its own success, and can no longer accommodate many of the empirical findings of studies that it has motivated. An alternative is proposed: Working memory functions arise through the coordinated recruitment, via attention, of brain systems that have evolved to accomplish sensory-, representation-, and action-related functions. Evidence from behavioral, neuropsychological, electrophysiological, and neuroimaging studies, from monkeys and humans, is considered, as is the question of how to interpret delay-period activity in the prefrontal cortex.

Working memory refers to the retention of information in conscious awareness when this information is not present in the environment, to its manipulation, and to its use in guiding behavior. Working memory has been implicated as a critical contributor to such essential cognitive functions and properties as language comprehension, learning, planning, reasoning, and general fluid intelligence (Baddeley, 1986; Engle, Kane, & Tuholski, 1999; Jonides, 1995). In this review I will argue against the idea that working memory functions are supported by the operation of one or more specialized systems, and instead, that they arise through the coordinated recruitment, via attention, of brain systems that have evolved to accomplish sensory-, representation-, or action-related functions. One implication of this view is that the contributions of prefrontal cortex (PFC) to working memory do not include the temporary storage of information.

The Evolution of a Standard Model of Working Memory
Although the use of the term working memory in the behavioral sciences dates back at least as far as 1960 (Miller, Galanter, & Pribram, 1960; Pribram, Ahumada, Hartog, & Roos, 1964), the early 1970s witnessed two developments that were seminal in shaping contemporary conceptions of working memory. The first was the observation that individual neurons in PFC of the monkey demonstrated sustained activity throughout the delay period of a delayed-response task (Fuster, 1973; Fuster & Alexander, 1971; Niki, 1974). The impact of this result derived, in part, from the fact that it suggested a neural correlate of two potent ideas from physiological psychology -- that of a PFC-based “immediate memory” (Jacobsen, 1936), and that of a “reverberatory” mechanism for “a transient ‘memory’ of [a] stimulus” (Hebb, 1949) (p. 61). The second development, which occurred in the field of human cognitive psychology, was the introduction of the multiple component model of working memory (Baddeley & Hitch, 1974). This model has proven to be enormously influential, spawning a prodigious amount of research that continues unabated to this day. In its initial instantiation it comprised two independent buffers for the storage of verbal and of visuospatial information, and a Central
Executive to control attention and to manage information in the buffers. Prompted by these
two developments, the neuroscientific and the psychological study of working memory each
proceeded along parallel, but largely independent, paths until the late 1980s, when a third
important advance occurred.

The third advance was the proposal by Goldman-Rakic that the sustained delay-period activity
in PFC that was studied by neuroscientists (e.g. Funahashi, Bruce, & Goldman-Rakic, 1989,
1990; Fuster, 1973; Fuster & Alexander, 1971; Fuster, Buer, & Jervey, 1982; Niki &
Watanabe, 1976; Quintana, Yajeya, & Fuster, 1988; Watanabe, 1981) and the storage buffers
of the multiple-component model (Baddeley, 1986; Baddeley & Hitch, 1974) were cross-
species manifestations of the same fundamental mental phenomenon (Goldman-Rakic, 1987,
1990). This conceptual integration of the neuroscientific and psychological traditions of
working memory research has proven to be a remarkably fruitful and influential idea, to the
extent that it is virtually impossible to find a published report on working memory from the
past decade that doesn’t cite the work of Baddeley, or Goldman-Rakic, or both. Thus, the
conceptual integration of the neuroscientific and psychological traditions of working memory
research has given rise to a “standard model” of working memory.

The utility of integrating psychological and neuroscientific approaches to working memory
was first seen in the proposal by Goldman-Rakic and her colleagues that the “what/where”
organization of the visual system might also apply to visual working memory. Results from
monkey electrophysiology (Wilson, O’Scalaidhe, & Goldman-Rakic, 1993), experimental
psychology (Della Sala, Gray, Baddeley, Allamano, & Wilson, 1999; Hecker & Mapperson,
1997; Smith et al., 1995; Tresch, Sinnamon, & Seamon, 1993), neuroimaging (e.g., Smith et
al., 1995), human electrophysiology (Mechlking & Muller, 1996), and human
neuropsychology (Owen, Iddon, Hodges, Summers, & Robbins, 1997; Postle, Jonides, Smith,
Corkin, & Growdon, 1997) confirmed the validity of this idea. The multiple component model
was adjusted accordingly, with the visuospatial sketchpad divided into “visual cache” and
“inner scribe” components for representing object and spatiotemporal information,
respectively (Baddeley & Logie, 1999; Logie, 1995).

Two tenets of the standard model

The standard model of working memory has two central tenets, one theoretical and one
neuroscientific. The theoretical tenet holds that working memory storage functions are the
product of the operation of specialized systems that serve as buffers for the storage and
manipulation of information: the phonological store/articulatory loop; the visuospatial
sketchpad; and the more recently proposed episodic buffer. That this continues to be an
influential view is seen, for example, in its centrality to many contemporary accounts of visual
short-term memory (VSTM, e.g., Vogel, Woodman, & Luck, 2001; Woodman & Vogel,
2005). The neuroscientific tenet holds that PFC is a critical neural substrate for the storage
buffers of working memory. Consistent with this view is the well-established fact that damage
to PFC disrupts working memory performance (e.g., Funahashi, Bruce, & Goldman-Rakic,
1993a; Goldman & Rosvold, 1970; Gross, 1963; Jacobsen, 1936; Petrides & Milner, 1981,

1The foundational theoretical expositions of the multiple component model emphasize the commonalities between the phonological
buffer and the speech production apparatus (e.g., Baddeley, Thompson, & Buchanan, 1975), and between the visuospatial sketchpad and
the visual system (e.g., Logie, 1986). Nonetheless, its principal proponents also argue explicitly that working memory comprises
specialized systems (e.g., Baddeley, 2003; Logie & Della Sala, 2003).

2For expository simplicity, use of the term “PFC” in this review encompasses the region of Superior Frontal Sulcus immediately anterior
to the Frontal Eye Fields that has been characterized as “specialized” for spatial working memory (Courtney, Petit, Maisog, Ungerleider,
Most influential at popularizing the neuroscientific tenet of the standard model have been single unit electrophysiological recordings from awake behaving monkeys. In particular, one series of experiments demonstrated that the delay-period activity of units in the Principal Sulcus of the PFC was tuned for specific regions of space in the contralateral visual field (Funahashi et al., 1989, 1990; Funahashi, Chafee, & Goldman-Rakic, 1993b) (i.e., these neurons demonstrated “memory fields”), and a subsequent study found that, on a test of working memory for object identity, delay-period responsive neurons were localized to a region of PFC, the Inferior Convexity, that is ventrolateral to the principle sulcus (Wilson et al., 1993). The model articulated by Wilson, O’Scalaidhe, and Goldman-Rakic in this latter paper is as close to an “official” position statement of the standard model as exists, declaring that “prefrontal cortex contains separate processing mechanisms for remembering “what” and “where” an object is” (p. 1955). Thus, the standard model advances the intuitively appealing and parsimonious idea that working memory for different domains of information is accomplished by PFC modules that receive direct projections from specific posterior perceptual information processing areas (e.g., dorsolateral PFC carries out working memory processing of information from the posterior “dorsal stream,” and ventrolateral PFC performs the same function for information processed by the posterior “ventral stream” of the visual system)\(^3\). The idea that PFC is the neural substrate for the storage buffers of the standard model continues to enjoy broad support (e.g., Davachi, Romanski, Chafee, & Goldman-Rakic, 2004; Goldman-Rakic & Leung, 2002; Hamker, 2005; Haxby, Petit, Ungerleider, & Courtney, 2000; Leung, Gore, & Goldman-Rakic, 2002; Leung, Seelig, & Gore, 2004; Logie & Della Salla, 2003; Mottaghy, Gangitano, Sparing, Krause, & Pascual-Leone, 2002; Munk et al., 2002; Pessoa, Gutierrez, Bandettini, & Ungerleider, 2002; Sala, Rama, & Courtney, 2003; Slotnick, in press; Tek et al., 2002). Indeed, a recent theoretical review by Courtney (2004) offers a contemporary counterpoint to this review, from the perspective of the standard model.

As happens with many important ideas in science, however, the remarkable influence of the standard model has accelerated the generation of empirical results that now demand its revision. As a result, the two central tenets of the standard model are becoming increasingly difficult to sustain.

**Logical Problems with the Theoretical Framework of the Standard Model**

As reviewed above, the cognitive architecture of the multiple component model (and, therefore, of the standard model) was updated in the 1990s to accommodate the what vs. where distinction in visual working memory. Since that time, however, it has not kept up with an ever increasing number of empirical dissociations that would seem to require of it the declaration of an ever increasing number of working memory subsystems. For example,

- in the spatial domain (i.e., the “inner scribe”), there is evidence for a dissociation between egocentric and allocentric spatial working memory (Postle & D’Esposito, 2003; Woodin & Allport, 1999); and within egocentric working memory there is evidence for even further dissociations according to the motor effector system engaged by the task (e.g., hand centered vs. eye centered vs. foot centered spatial working memory, Cheffi, Allport, & Woodin, 1999);
- working memory for each of several low-level visual attributes, including spatial frequency, contrast, orientation, and motion, is retained in a highly stimulus-specific,

\(^3\)In parallel with the articulation of the standard model by Goldman-Rakic and her colleagues, Petrides has advocated the view of the functional organization of PFC that emphasizes organization by process (i.e., maintenance and monitoring) rather than by domain (e.g., Owen, Evans, & Petrides, 1996; Petrides, 1989, 1991, 1994b). Although these two models have often been portrayed as conflicting alternatives, it has been argued and demonstrated that they are not necessarily mutually exclusive (e.g., Johnson, Raye, Mitchell, Greene, & Anderson, 2003).
and therefore dissociable, manner (Magnussen, 2000; Magnussen & Greenlee, 1999; Zaksas, Bisley, & Pasternak, 2001);

- neuroimaging studies of working memory for visually perceived objects indicate that working memory for manipulable vs. nonmanipulable objects is neurally dissociable (Mecklinger, Gruenewald, Besson, Magnie, & Von Cramon, 2002), as is working memory for faces vs. houses (Ranganath, Cohen, Dam, & D’Esposito, 2004a) and for faces vs. outdoor scenes (Ranganath, DeGutis, & D’Esposito, 2004b);

- within verbal working memory, there are neuropsychological and neuroimaging dissociations suggesting separability of working memory for phonological vs. semantic vs. syntactic information (Martin, Hamilton, Lipszyc, & Potts, 2004; Shivde & Thompson-Schill, 2004);

- also within verbal working memory, there is evidence for a distinction between the “content” of a signal (i.e., the specific words used in a sentence) and its “carrier” (i.e., the identity of the voice speaking the sentence, Stevens, 2004).

- within audition, working memory for pitch, loudness, and location are dissociable from each other (Anourova et al., 1999; Clement, Demany, & Semal, 1999; Deutsch, 1972)

- working memory for tactile stimuli is a robust phenomenon (Harris, Miniussi, Harris, & Diamond, 2002; Hernandez, Salinas, Garcia, & Romo, 1997; Sinclair & Burton, 1996)

- working memory for olfactory stimuli is also a robust phenomenon (Dade, 2001).

This summary has doubtless overlooked some, and there are doubtless many more such empirical demonstrations of dissociable working memory performance to come. From a memory systems perspective, such as that of the standard model, the proliferation of systems and subsystems demanded by these empirical findings creates fundamental problems. At the theoretical level, it requires an increasingly complex taxonomy of boxes and arrows linking each independent storage buffer (or subbuffer) to the Central Executive. At the neurobiological level, it demands an increasingly complex scheme of compartmentalization of PFC into ever more numerous topographically discrete modules, each responsible for the processing in working memory of information of a particular domain (e.g., identity of voice, or location with respect to the right foot)\(^4\). Followed to its logical extreme, the cognitive architecture of the standard model would eventually depict a working memory system organized into hundreds (if not thousands or more) of domain-specific buffers, each responsible for the working memory processing of a different kind of information; and the neural instantiation of the standard model would eventually depict a PFC organized into hundreds (if not thousands or more) of domain-specific “grandmother modules”, each responsible for the working memory processing of a different kind of information. Such an extreme state of affairs subverts a model that once made a strong intuitive appeal to parsimony into an unwieldy organizational scheme that redundantly duplicates every representational system in the mind and brain.

Independent of the logical considerations reviewed above, another challenge to the idea of specialized, “high-level” capacity-limited working-memory buffers comes from the perspective of signal detection theory. Based on the results of a series of studies of change detection in VSTM, Wilken and Ma (2004) argue that the apparent capacity limitation in VSTM, which has been interpreted as evidence for a limited number of “slots” in a visual cache-like store, is more simply explained as a byproduct of noise in retained perceptual representations, which increases monotonically as a function of memory set size. (Similar ideas

\(^4\)Examples of the beginning of this process are seen in Fig. 48.5 of Davachi et al. (2004) and Fig. 1 of Courtney (2004).
have also been proposed by Davis and Holmes (2005), and entertained by Magnussen et al. (1998) and by Vogel et al. (2001).

Empirical Challenges to the Neurobiological Tenet of the Standard Model

Since the publication of the Wilson et al. (1993) paper, evidence for segregation by domain of PFC working memory activity has been viewed as a cornerstone of evidence for the standard model. As a result, researchers leery of the problem of mushrooming working memory systems tend to take a skeptical view of the segregation-by-domain model of PFC, because it reifies the multiple systems view. For this reason, the debate about the working memory functions of PFC has implications beyond the brain-mapping question of where different working memory functions are performed, and the question of the organization of working memory function in PFC has become a conceptual battleground between adherents to the standard model and advocates of alternative views (e.g., Duncan & Owen, 2000; Fuster, 2002; Goldman-Rakic, 2000; Goldman-Rakic & Leung, 2002; Haxby et al., 2000; Miller, 2000; Muller, Machado, & Knight, 2002; Owen et al., 1999; Passingham & Rowe, 2002; Petrides, 2000b; Postle & D’Esposito, 2000; Postle, Druzdgal, & D’Esposito, 2003; Sala et al., 2003; Slotnick, in press; Ungerleider, Courtney, & Haxby, 1998).

Monkey electrophysiology

In mounting perhaps the most direct challenge to the standard model, Miller and colleagues have noted prior evidence for the integration of delay-period representation of spatial and nonspatial information within individual PFC neurons (Watanabe, 1981), and the possibility that the results of Wilson and colleagues (1993) are attributable to the idiosyncrasies of the tasks and to the conditions under which their monkeys were trained (Bichot, Schall, & Thompson, 1996; Braver & Cohen, 1995), rather than to a fundamental organizational principle of the PFC. Miller and colleagues evaluated this alternative by training their monkeys to perform delayed object matching and spatial delayed response within the same trial, a design that removed the confound of learning and performing just one working memory task during the several-month duration of an experiment. Their results indicated that the majority of delay-active PFC neurons from which they recorded did not discriminate spatial from object delay periods (Rainer, Asaad, & Miller, 1998; Rao, Rainer, & Miller, 1997). The results of several other electrophysiological studies suggest that, rather than representing a labeled input from posterior sensory areas (as the standard model holds), neurons in the dorsolateral PFC of the monkey adapt flexibly, in a domain-independent manner, to represent whatever information is critical for task performance (e.g., Duncan & Miller, 2002; Fuster, 2002).

Another direct challenge to the standard model has come from Lebedev and colleagues (2004), who directly pitted the “memory buffer” view of PFC against an alternative “attentional” view, with a task that literally pulled apart the foci of stimulus memory and stimulus attention on each trial. They found that the majority of Principal Sulcus neurons from which they recorded represented attended locations during the delay period, with a smaller proportion of their sample demonstrating “hybrid” (i.e., attention and memory) tuning, and a still smaller proportion demonstrating memory tuning. From these results they concluded that attentional functions, such as monitoring and selection, probably account for much more delay-period activity in the PFC than does memory storage. They also applied this reasoning to a reinterpretation of several earlier publications that supported the standard model, suggesting, for example, that electrophysiological results that were initially interpreted as evidence for “purely sensory-mnemonic function for PF[C] neurons” (e.g., Constantinides, Franowicz, & Goldman-Rakic, 2001; Funahashi et al., 1993b) could be reinterpreted as evidence for attentional selection and monitoring, and that evidence for “mnemonic scotomas” produced by punctate lesions in the vicinity of the Principal Sulcus (Funahashi et al., 1993a) “might be
better understood as a localized neglect-like phenomenon or some combination of attention and memory deficits” (pp. 1930–1931, Lebedev et al., 2004).

**Monkey neuropsychology**

In the same year as the Rao et al. (1997) study described in the previous section, Rushworth and colleagues published the results of two lesion experiments that tested the standard model’s prediction that cortex of the Inferior Convexity, ventral to the Principal Sulcus, is important for the short-term retention of visual characteristics of objects (i.e., “what”). After lesions to areas 47 and 12 they found no deficits on delayed visual pattern association or delayed color matching, two results that the standard model cannot easily accommodate (Rushworth, Nixon, Eacott, & Passingham, 1997). A subsequent lesion study by Petrides (2000a) demonstrated a double-dissociation of working memory functions attributable to PFC vs. anterior inferior temporal (IT) cortex: Lesions of PFC did not impair memory for the selection of one among two object stimuli across long (90 and 120 sec) delay periods, but did disrupt memory for one from among a set of three, four, or five items across shorter (10 sec) delays; whereas lesions of anterior IT cortex had the converse effect. These results are logically inconsistent with the idea that PFC is a necessary neural substrate for working memory storage. Rather, they support an alternative view that object working memory storage depends on IT cortex, whereas control functions such as the monitoring of multiple mnemonic representations are supported by PFC.

**Human neuropsychology**

A considerable body of human research is also inconsistent with the standard model. Much of it is summarized in a review of the literature on the cognitive functions of patients with PFC lesions. This review indicated that working memory storage functions of patients with large PFC lesions were unimpaired, as indexed by performance on tests of verbal and nonverbal memory span, and of delayed response and recognition (D’Esposito & Postle, 1999). One of the studies reviewed, for example, found no deficit in forward digit span performance in a group of 24 patients with PFC lesions, although it did find an impairment in a group of 20 patients with lesions in left temporoparietal cortex (Ghent, Mishkin, & Teuber, 1962). Illustrative of the results for delayed-response and -recognition was a study that found intact 60-sec delayed recognition of novel nonsense shapes in a group of PFC-lesioned patients (Prisko, 1963). In another set of delayed-recognition experiments, Chao and Knight (1998; 1995) found deficits in delayed tone recognition in patients with lesions of lateral PFC only when distracting sounds were played during the delay period of the task. On the whole, results from the human neuropsychological literature are difficult to reconcile with a view that PFC is responsible for working memory storage functions.

**Human neuroimaging**

The conclusions arising from review of the patient literature are bolstered by the results of several neuroimaging studies that have failed to find evidence for segregation-by-domain of PFC working memory activity (Arnott, Grady, Hevenor, Graham, & Alain, 2005; D’Esposito, 8

---

5The Petrides (2000a) result has been seen by some as difficult to reconcile with an earlier electrophysiological result from an ABBA task (Miller, Erickson, & Desimone, 1996). The ABBA results had been interpreted by many as evidence that the critical storage site for object working memory was PFC, not IT cortex, because delay-period activity in the latter, but not the former, region was disrupted by intervening distractor stimuli. This seeming inconsistency, however, may be the spurious result of an “aperture effect”, in that it relies on the Miller et al. (1996) results applying to the entirety of the temporal cortex. In fact, however, electrophysiological recordings by Nakamura and Kubota (1995) from a more anterior region of IT cortex, as well as by Suzuki, Miller, and Desimone (1997) from entorhinal cortex, have demonstrated robust distractor-resistant delay-period activity that is compatible with the Petrides (2000a) neuropsychological results. The contributions of anteromedial temporal areas to working memory are considered in more detail in the review by Ranganath (2005) that appears in this issue.

6These same subjects were impaired on comparable tests for flicker frequency and color, and for auditorily presented click frequency and tones, but these tests were confounded by the fact that they drew stimuli from closed sets, thereby producing high levels of proactive interference (Milner, 1964).
Aguirre, Zarahn, & Ballard, 1998; D'Esposito & Postle, 2002; Mecklinger, Bosch, Gruenewald, Bentin, & von Cramon, 2000; Nystrom et al., 2000; Owen et al., 1998; Postle, Berger, Taich, & D'Esposito, 2000a; Postle & D'Esposito, 1999, 2000; Postle, Stern, Rosen, & Corkin, 2000b; Ranganath et al., 2004b; Stern et al., 2000). For example, Postle and D'Esposito (1999) modified the design of Rao et al. (1997) to evaluate the organization of working memory for the identity and for the location of visually presented stimuli in the human brain, with functional magnetic resonance imaging (fMRI). Our event-related analysis method permitted isolation of delay-related effects, unconfounded by contributions from other epochs of the trial (Postle, Zarahn, & D'Esposito, 2000c; Zarahn, Aguirre, & D'Esposito, 1997). Although the task produced considerable delay-period activity in ventrolateral PFC, dorsolateral PFC, and superior frontal cortex, in no subject, in no PFC ROI, was this activity statistically greater for one stimulus domain than for the other. Clear dissociations of delay-period activity were found, however, in extrastriate regions. Another study (Postle et al., 2000a) evaluated the evidence for a region of superior frontal sulcus immediately anterior to the frontal eye fields (FEF) that is “specialized” for spatial working memory (Courtney et al., 1998). It pitted this standard-model hypothesis against an alternative account, which was that this region may simply be sensitive to the complexity of eye movement tasks – the working memory task of the Courtney et al. (1998) study required saccades to stimuli appearing at unpredictable locations on a 2D display, whereas the eye movement control task from this study required self-paced horizontal saccades. The Postle et al. (2000a) results failed to find evidence that spatial working memory-related activity in the superior frontal sulcus is greater than the activity evoked by a visually guided saccade task of comparable kinematic complexity. (This result has more recently been revisited by Slotnick (in press) and Postle (in press-a).)

In contrast to the studies reviewed in the previous paragraph, there have also been several reports demonstrating PFC working memory-related activity that is largely, but not entirely, overlapping for different stimulus domains (Gruber & von Cramon, 2003; Manoach et al., 2004; Rama et al., 2004; Rama, Sala, Gillen, Pek, & Courtney, 2001; Sala et al., 2003). Courtney and colleagues have interpreted these results as suggesting an organizational scheme of graded degrees of segregation and overlap of working memory storage functions of the PFC (e.g., Courtney, 2004). How does one reconcile the empirical discrepancies between studies that do versus those that do not find evidence for domain segregation of delay-period activity in PFC? It may be that these discrepancies are due, in large part, to methodological factors. First, each of the studies considered in this paragraph employed designs and/or analysis approaches that may have rendered their estimates of delay-period activity vulnerable to contamination by stimulus-related variance. This is because each either used a task with too short a delay period, or modeled delay-period activity with a boxcar covariate that spanned the entire delay period (or both; for detailed discussion of these factors, see Postle, in press-a; Zarahn et al., 1997). Second, each of these studies employed group analyses performed on spatially normalized data sets. This approach typically produces delay-period effect sizes that are on the order of tenths or hundredths of a percentage of signal change. Many of the studies that do not find evidence for domain segregation of PFC delay-period activity, in contrast, use a single-subject approach that finds domain segregation effects in posterior cortex that are one-to-two orders of magnitude larger – on the order of whole percentage points of signal change. The question of how to evaluate the relative functional significance of the results produced by these two types of group analysis will be postponed until the review of the memory load literature, to which these methodological issues are also relevant.

Thus far, this section on neuroimaging has emphasized the empirical question of the domain specificity of working memory-related activity in PFC. Another factor that can shed light on

---

7Consistent with this result is a recent electrophysiological study showing spatial tuning properties in human ventrolateral PFC (Rizzuto, Mamelak, Sutherling, Fineman, & Anderson, 2005)
the neural bases of memory storage processes is that of load – the number of items that must
be held in memory. A region whose activity varies systematically with load is a candidate
substrate for memory storage functions. (Note that such inferences about storage are harder to
draw when load is varied in a complex task like the n-back (e.g., Braver et al., 1997; Cohen et
al., 1997; Jonides et al., 1997), because many processes other than storage might also be
sensitive to this manipulation.) One study has investigated the effects of varying memory load
on delay-period activity in PFC with a task that compared the effects of varying load with those
of varying manipulation demands on delayed recognition of the ordinal position of letter stimuli
(Postle, Berger, & D’Esposito, 1999). Manipulation was operationalized with a contrast
between a task that required subjects to reposition into alphabetical order the 5 randomly
ordered letters of a memory set (Alphabetize 5) and a simple delayed-recognition task in which
no alphabetization was required (Forward 5). Load was operationalized with the contrast
(Delay Forward 5 vs. Delay Forward 2). The results revealed a significant alphabetization effect
(Delay Alphabetize 5 vs. Delay Forward 5) in dorsolateral PFC in 5 of 5 subjects, and load effects
in dorsolateral PFC in only 2 subjects. (Reliable load effects were seen in all subjects in left
posterior perisylvian cortex.) Notably, there was no evidence of sensitivity to memory load in
the voxels that showed the alphabetization effect, indicating that executive control and
mnemonic functions did not colocalize in PFC. Corroborating results came from an fMRI study
by Rypma and D’Esposito (1999) that found load effects in PFC during the encoding period,
but not the delay period, and delay-period load effects in left inferior parietal cortex. Subsequent
research, however, has painted a more complex picture of the effects of load on delay-period
activity in PFC.

A subsequent study by Rypma and colleagues (2002) that varied load parametrically between
1 and 8 letters revealed no encoding-related load effects, but significant delay- and probe-
related load effects in dorsolateral PFC, and trends in this direction in ventrolateral PFC. An
individual-differences analysis indicated that these patterns of load-dependent effects by trial
epoch varied between high- and low-performing groups, leading the authors to interpret the
delay-period load effects as evidence for strategic reorganization of information, rather than
for storage per se. Another study reported load effects in PFC with an n-back task and an item
recognition task (Veltman, Rombouts, & Dolan, 2003), but because the design of neither task
permitted isolation of delay-period activity, the implications of this study for storage processes
is unclear. Finally, three very recent studies using verbal material have produced inconsistent
results: Narayanan et al. (2005) and Zarahn et al. (2005) find delay-period sensitivity to load
in PFC, whereas Postle et al. (2004c), in a replication and extension of the Postle et al.
(1999) study that uses both fMRI and fMRI-guided repetitive transcranial magnetic stimulation
(rTMS), do not. Additionally, this latter study found that delay-period rTMS of PFC disrupted
performance on Alphabetize 5, but not Forward 5, trials.

As with the “organization-by-stimulus-domain” literature, these discrepant findings in the
“load” literature also break down according to analysis method: Narayanan et al. (2005) and
Zarahn et al. (2005) report spatially normalized group analyses producing PFC delay-period
load effects that, from their figures, appear to be of approximately .2% signal change; whereas
the Postle et al. (2004c) analyses, performed on single-subject data sets, identified delay-period
load effects in PFC in only 3 of 21 subjects, but robust load effects in several nonPFC regions
in 15 different subjects, all contributing to a group mean load effect of 2.4% signal change.
The idea that the two approaches to group analysis of fMRI data can yield such discrepant
results was confirmed when we reanalyzed the Postle et al. (2004c) data set with a spatially
normalized group analysis approach, and this analysis yielded a load effect of .15% at the
boundary of Brodmann’s areas 6 and 9 in the left hemisphere (Feredoes & Postle, 2005). This
raises a question of clear importance for the neuroimaging community: how to interpret the
relative functional significance of the “large” effects that are often observed in single-subject
data sets vs. the “small” effects that are evidently subthreshold at the single-subject level, but
detected by spatially normalized group analyses. Detailed consideration of this question is beyond the scope of the present review. For the purposes of evaluating the relative importance of the domain- and load-related signals identified by the two methods, however, I will suggest that the “large” effects may provide a better index of neural activity that is functionally significant to the organism.

The literature on load manipulation with nonverbal stimuli is also equivocal. Some studies that have varied the number of faces to be remembered have reported delay-period load effects in PFC (Druzgal & D’Esposito, 2003; Ranganath et al., 2004b), although another has not (Jha & McCarthy, 2000). Druzgal et al. (2003) also found delay-period sensitivity to load in the fusiform face area (FFA), whereas Ranganath et al. (2004b) did not. And because the study by Ranganath and colleagues also employed a second stimulus category in addition to faces – scenes – the authors could also assess evidence of category specificity of delay-period activity. Here, they found clear category specificity in regions of ventral temporal cortex (with delay period activity in the FFA greater for face memory than for house memory, and the converse in the Parahippocampal Place Area), but not in PFC. Because load sensitivity and stimulus specificity can both be interpreted as evidence that delay-period activity supports a storage function, the implication of these “face” data for the debate over storage functions of PFC are unclear. In the spatial domain, Leung and colleagues have described a load effect in PFC for 5 vs. 3 items (Leung et al., 2002), and a more complicated set of results in a more recent study that varied load parametrically between 1 and 4 items (Leung et al., 2004). In the latter study, the authors partitioned data from their delayed-recognition task into target, early delay, late delay, and probe epochs. Effects were qualitatively similar in the two components of the delay, with those from the late delay the most clearly free of contamination from the target epoch. Here, the strongest effects of load were seen in parietal cortex (Precuneus, Superior Parietal Lobule, Inferior Parietal Lobule), and were roughly monotonic, but with signal flattening out or decreasing slightly between load 3 to load 4. The pattern in FEF was qualitatively similar, but was compressed over a narrower range. In contrast to these regions, the Middle Frontal Gyrus of the PFC showed a qualitatively different pattern, increasing from load 1 to load 2, then decreasing from 2 to 3 and from 3 to 4. The authors interpreted their fMRI results in terms of a physiological signal that tracks behavioral performance (i.e., increasing over “easy” loads, then decreasing as a capacity limit was reached), although one cannot know from the behavioral data whether the parietal regions or the PFC may have done this with higher fidelity. (Indeed, Zarahn and colleagues (2005) interpreted a qualitatively similar pattern of load effects -- linear for several cortical regions, inverted-U for hippocampus – as “positive evidence against human hippocampal involvement in working memory maintenance” (p. 303.).)

The studies of Leung and colleagues (2004) and of Rypma and colleagues (2002) illustrate the importance of constraining the interpretation of neuroimaging data with careful analysis of behavioral performance. This approach has been realized most effectively in studies of VSTM. In these studies, also known as “change detection” studies, subjects’ memory for arrays of differently colored (but identically shaped) stimuli – typically squares or circles, varying in number from 1 to 8 or more – is assessed with Yes/No recognition. VSTM capacity is estimated with a procedure that estimates the underlying capacity (presumably a stable value) that accounts for performance at different loads. The advantage of this approach is that it removes the subjectivity from interpreting behavioral performance, thereby providing a powerful tool for analyses that take into account individual differences, estimated independently of the imaging study itself. The power of this approach was illustrated by Vogel and Machizawa (2004), who described an ERP correlate of VSTM that scales with load before saturating at a level commensurate with an individual’s VSTM capacity. This signal was strongest over posterior parietal and lateral occipital electrode sites, suggesting that its generators may have corresponded to the (bilateral) region of intraparietal sulcus/introccipital sulcus whose fMRI signal scaled with a group estimate of VSTM capacity in a study by Todd.
and Marois (2004). Subsequently, Todd and Marois (in press) have confirmed that fMRI signal in this region varies systematically with individual differences in VSTM capacity. These studies provide compelling evidence that the short-term retention of visual information is supported by extrastriate visual areas.

As a final note in this section of the human neuroimaging literature, one group that has challenged the standard model suggests that many neuroimaging studies of delay task performance fail to find any reliable PFC delay-period activity (Passingham & Sakai, 2004; Passingham & Rowe, 2002). For example, Rowe and colleagues (2000) only found delay-period activity in the PFC when their task placed high demands on attentional selection. It is certainly true that PFC activity has more reliably been shown to be sensitive to the experimental manipulation of nonstorage factors such as attentional selection, manipulation, or proactive interference, than to the storage-related factors reviewed in this section. It is also true, however, that many neuroimaging studies of tasks that only require the simple, short-term retention of information across a brief delay period find delay-period activity in the PFC (e.g., Courtney, Ungerleider, Keil, & Haxby, 1996; Courtney, Ungerleider, Keil, & Haxby, 1997; D’Esposito, Postle, Ballard, & Lease, 1999; Narayanan et al., 2005; Pessoa et al., 2002; Postle, Awh, Jonides, Smith, & D’Esposito, 2004a; Postle et al., 1999; Postle & D’Esposito, 1999; Postle et al., 2003). The question of how to interpret the function of this delay-period activity, if not as being directly storage-related, is taken up in the final section of this review.

In summary, there is a large body of results from neuropsychological, neurophysiological, and neuroimaging studies that is difficult to reconcile with the standard model’s depiction of the PFC as a critical substrate for working memory storage. Many of these data, however, are consistent with the alternative account that will be advanced in the next section.

**An Alternative View: Working Memory as an Emergent Property**

Although there may not exist specialized, PFC-based working memory storage systems, it is clear that many different kinds of information can nonetheless be retained “in” working memory. Indeed, it seems reasonable to assume that if the brain can represent it, the brain can also demonstrate working memory for it (see also, for example, Engle & Kane, 2004; Schneider & Detweiler, 1987). The view that will be advanced here is that working memory functions are produced when attention is directed to systems that have evolved to accomplish sensory-, representation-, or action-related functions. From this perspective, working memory may simply be a property that emerges from a nervous system that is capable of representing many different kinds of information, and that is endowed with flexibly deployable attention.

Predictions about the nature of representations contributing to the short-term retention of any particular kind of information are made by considering the nature of the information that is to be remembered, and the mental processes that are afforded by the task that is being performed. In this section I will illustrate this idea with examples from working memory for locations in space, for visually presented objects, and for linguistic stimuli.

**Spatial working memory**

Because a target location can be remembered both in terms of its retinal position and in terms of the metrics of the saccade and/or grasp that would be required to acquire it, one can hypothesize that short-term retention of locations can be supported by both a (retrospective) perceptual code and a (prospective) motor code (Postle & D’Esposito, 2003). At the level of mechanism, there is empirical evidence that the former can be implemented with attention-based rehearsal, the rehearsal of spatial information via covert shifts of spatial selective attention to memorized locations (Awh & Jonides, 2001; Awh, Vogel, & Oh, 2005). The latter can be implemented when the target location can be represented in relation to the viewer’s body (Postle & D’Esposito, 2003). Thus, from the emergent property perspective, the short-
term retention of spatial information does not depend on a specialized memory system – not, at the theoretical level, by an inner scribe of a visuospatial scratch pad; nor, at the neurobiological level, by a region of PFC that is specialized for spatial working memory. Rather, it bootstraps off capabilities afforded by spatial selective attention and motor control. Mechanistically, attention-based rehearsal is accomplished by allocating attention (via activity in the FEF (e.g., Corbetta et al., 1998; Moore & Fallah, 2004) and parietal attention centers (e.g., Yantis et al., 2002)) to regions of extrastriate and parietal cortex responsible for the perception of location (Awh, Anllo-Vento, & Hillyard, 2000; Awh et al., 1999; Postle et al., 2004a). Prospective motor coding, in turn, is accomplished by transforming vision-based coordinates into a motor plan, retaining this motor plan throughout the delay period (both processes that engage the caudate nucleus (Postle & D’Esposito, 2003) and possibly the superior colliculus (Theeuwes, Olivers, & Chizk, 2005)), and using it either to execute the response (in the case of delayed response) or to evaluate the validity of the memory probe (in the case of delayed recognition). Precentral and subcortical brain systems that represent the motor effector(s) engaged by the task can also participate in retention and probe evaluation/response-related activity (Balan & Ferrera, 2003; Curtis, 2005; Curtis, Rao, & D’Esposito, 2004; Postle & D’Esposito, 2003; Theeuwes et al., 2005), as can the PFC (Takeda & Funahashi, 2002, 2004). Consistent with the prospective motor coding idea is the fact that working memory for locations is disrupted by concurrent motor activity (e.g., concurrent finger tapping (Farmer, Berman, & Fletcher, 1986; Salway & Logie, 1995; Smyth, Pearson, & Pendleton, 1988), pointing (Hale, Myerson, Rhee, Weiss, & Abrams, 1996), eye movements (Baddeley, 1986; Hale et al., 1996; Lawrence, Myerson, Oonk, & Abrams, 2001; Pearson & Sahraie, 2003; Postle, Idzikowski, Della Salla, Logie, & Baddeley, in press-b), and arm movements (Baddeley & Lieberman, 1980; Lawrence et al., 2001; Logie & Marchetti, 1991; Quinn & Ralston, 1986)).

Thus far, what has been described are two separate mechanisms for the retention of location information in working memory – a retrospective attention-based rehearsal mechanism and a prospective motor coding mechanism – but it may be that these two are better thought of as two points along a single sensorimotor continuum. This would follow from evidence for close linkages between spatial attention and oculomotor control (e.g., Andersen et al., 2004; Goldberg, Bisley, Powell, Gottlieb, & Kusunoki, 2002; Goodale & Westwood, 2004; Hoffman & Subramaniam, 1995), and may be reflected in the trend of decreasing contralateral bias in delay-period activity (an index of attention-based rehearsal) that is seen as one proceeds rostrally from peristriate, through parietal, premotor, and finally prefrontal cortex (Figure 2.C., Postle et al., 2004a). This notion is demonstrated directly when Moore and Fallah (2004) apply subthreshold stimulation to a region of the FEF with a known motor field and produce a covert shift of spatial attention to the very same region of space represented by the motor field. (I.e., stimulation in a putatively motoric area results in improved visual perceptual performance, in this case, detection of a luminance change.) Such results suggest that, at least to a certain extent, motor intention and sensory attention arise from the same underlying network dynamics (Hamker, 2005). From this perspective, the apparent distinction between putatively sensory-based vs. motor-based mechanisms for spatial working memory can seem to be a false one, one that may not be honored by the neural systems that give rise to spatial working memory. Consistent with this reasoning, Theeuwes and colleagues (2005) have demonstrated that remembering a location has the same deviating effect on eye movements as does visual presentation of a distracting stimulus at that same location. From this result they see “evidence for a strong overlap between visual working memory, spatial attention, and the oculomotor system” and raise “the possibility that working memory is “nothing more” than the preparation to perform an action, whether it be oculomotor, manual, verbal, or otherwise” (pp. 198–199).
Object working memory

fMRI studies of n-back (Postle et al., 2000b) and delayed-recognition (Postle & D’Esposito, 1999) working memory for location versus identity of abstract geometric shapes have each found object-specific memory-related activity in ventral temporal and occipital cortex, but not in PFC. A subsequent fMRI study employed a multistep ABBA-like design intended to winnow out delay-period activity that may be correlated with, but not necessary for, working memory for faces. The task featured three 7 sec delay periods that were interposed between the presentation of the first and second, second and third, and third and fourth stimuli. We reasoned that the multiple distracting events in this task might serve to “weed out” activity from the first delay period that wasn’t involved directly in storage, because only voxels whose activity was necessary for retaining the memory trace to the end of the trial would be expected to maintain their activity across distracting stimuli. Our hierarchical analysis procedure proceeded in three steps: First, we identified Delay 1-sensitive voxels (presumed to represent the superset of the neural correlate of mnemonic representation of the target face); second, we determined how many of these Delay 1 voxels remained active during Delay 2; and third, we identified the voxels from Step 2 that retained their activity during Delay 3. As expected, the results from each subject revealed Delay 1-specific activity in many brain areas, including PFC, posterior Fusiform Gyrus, and posterior parietal cortex. In each subject, only a subset of these voxels retained the Delay 1 signal during Delay 2. And posterior Fusiform Gyrus was the only region in which voxels retained the signal during Delay 3 in each subject (Postle et al., 2003). Other studies of working memory for faces are also consistent with an important role in retention for posterior Fusiform Gyrus (Druzel & D’Esposito, 2003; Ranganath et al., 2004a; Ranganath et al., 2004b). Thus, working memory for the identity of objects is associated with sustained activity in the very brain systems that are responsible for the visual perception of these stimuli. An important goal for future research on object working memory will be to understand, by analogy to spatial working memory, the relation between object-based attention and object working memory (for further consideration of this question, see Awh et al., 2005).

In addition to the studies reviewed thus far, behavioral studies indicate that working memory for objects relies on more than the sustained activation of ventral stream representations. Studies by Simons (1996) and Postle and colleagues (2005; in press-b) indicate that the short-term retention of the identity of visually presented stimuli, no matter how abstract or “nonnameable”, is selectively sensitive (in comparison to location memory) to verbal distraction. In one study, for example, delayed recognition of Attneave (1956) shapes was selectively sensitive to the RSVP presentation of concrete nouns during the delay period. Delayed recognition of locations, in contrast, was selectively sensitive to the endogenously generation of saccades during the delay period (Postle, Idzikowski, Della Salla, Logie, & Baddeley, in press). Based on these and subsequent studies (Postle, Messner, & Cappa, unpublished data), we have proposed that humans automatically, obligatorily recode information about nonspatial features of visually perceived stimuli into a verbal code, and that this verbal code is retained as a part of the short-term memory representation of the stimulus.

Verbal working memory

Finally, we turn to overtly language-based stimuli, perhaps the most studied by working memory researchers. Within this domain, too, neuroimaging studies designed to identify the neural loci of working memory storage have, for the most part, pinpointed left posterior perisylvian areas (e.g., Awh et al., 1996; Hickock, Buchsbaum, Humphreys, & Muftuler, 2003; Paulesu, Frith, & Frackowiak, 1993; Postle et al., 1999; Rypma & D’Esposito, 1999) – areas associated with speech perception and the phonological lexicon – and not PFC8. (Two

8Note that many studies have implicated left inferior PFC and premotor cortex in the rehearsal (as opposed to the storage) of information represented in an articulatory code, and that these brain areas are also implicated in speech production.

Neuroscience. Author manuscript; available in PMC 2006 April 4.
recent exceptions, reviewed earlier, are from Narayanan et al (2005) and Zarahn et al.
(2005). Additionally, as with visuo-object material, we have seen evidence that letters are
represented with multiple codes in working memory. This is seen, for example, when running
span (or “updating”) performance reveals sensitivity to manipulations of context that is
unrelated to the task (Postle, 2003). In these studies, performance suffers when the grouping
of stimuli at presentation is violated by the requirement to “drop” some of a group’s items from
the memory set but to keep others, and this effect persists despite explicit instructions to
subjects to ignore grouping information, despite intensive training, and across different rates
of stimulus presentation. Other work has shown that the short-term retention of letters is also
sensitive to study-test manipulations of the color in which items are presented (despite
instructions to ignore color), and to study-test changes in environmental context (Postle
2003, unpublished observations).

**Principles of the emergent property view**

Two principles emerge from the data reviewed in this section. First, the retention of information
in working memory is associated with sustained activity in the same brain regions that are
responsible for the representation of that information in non-working memory situations, such
as perception, semantic memory, oculo- and skeletomotor control, and speech comprehension
and production. Second, humans opportunistically, automatically, recruit as many mental
codes as are afforded by a stimulus when representing that stimulus in working memory. Thus,
for example, if the perception of an object triggers an association with one’s previous
experience, the name and the semantic content of that previous experience will contribute to
the representation in working memory of that object. So, too, will information about the
context in which the object was presented. Ideas that are consonant with this first principle of
the emergent processes framework have also been derived from reviews of the monkey
neuropsychological and human psychophysical literature (Pasternak & Greenlee,
2005; Theeuwes et al., 2005), and from the human neuroimaging literature (Jonides, Lacey, &
Nee, 2005; Slotnick, 2005).

This first principle of the emergent-property view also overlaps with the idea that the
representation of information in working memory is accomplished via the “temporary
activation of representations in long-term memory” (as proposed, for example, by Anderson,
1983; Cowan, 1995; Oberauer, 2002; Ruchkin, Grafman, Cameron, & Berndt, 2003). This idea
has met with considerable resistance from many quarters (see, for example, the commentary
following Ruchkin et al., 2003). Some of the concerns are in the following vein: If working
memory is accomplished via the temporary activation of long-term memory representations,
how is the ordinal position of items in working memory retained? How can one account for
the flexible transformation of information in working memory? and Working memory often
represents the “here and now” of a situation, and thus contains detail specific to the present
that cannot be derived from the mere activation of representations in semantic or lexical
memory. One important and novel contribution of the emergent property view may be the way
in which it addresses these important questions. It does so by appealing to its second principle,
that of multiple encoding. Thus, in the emergent-processes account, maintenance of ordinal
position is a process to which the speech production system is well-suited. So if the information
in working memory is being represented, in part, in an articulatory code, (covertly) cycling

---

9 This principle can be seen as a broadening of the multiple encoding theory of Wickens (1973), which holds that “[humans] tend to
encode … word[s] … on a multiple number of attributes” (p. 490) and that “such cognitive processing is highly automatic and
compulsive” (p. 485).

10 An exception seems to be when the task is to remember the location of a stimulus (Postle et al., 2005; Postle et al., in press-b; Postle
et al., unpublished data). This may be because egocentric locations don’t have any meaning in and of themselves, absent objects occupying
that space. Thus, when we attend to an empty location in space this perceptual act engages visual and motor systems, but not richly
elaborated semantic or episodic memories of the kind that we have for objects.

*Neuroscience. Author manuscript; available in PMC 2006 April 4.*
this information through the speech production apparatus (e.g., Page & Norris, 1998) would be a way to accomplish memory for order without resorting to a special-purpose memory system. (In such a case, one could invoke the operation of an “articulatory loop”, but to do so would seem to needlessly relabel a system that already exists.) The principle of multiple encoding permits this explanation to extend to working memory for all types of information, with the exception of egocentric location (see footnote #10). Consistent with this reasoning is the fact that memory for order is known to be superior when items can be represented with a verbal code (e.g., Glenberg & Fernandez, 1988; O’Connor & Hermelin, 1972; O’Connor & Hermelin, 1973). An analogous explanation can be invoked for ordinal memory for egocentrically encoded locations (e.g., as is required by the Corsi blocks task) if one allows for contributions from the oculo-and/or skeletomotor system, because ordering and sequencing are also fundamental to the control of these systems. What about the flexible transformation of information held in working memory? We know from the problem solving literature that the ability to represent information in a different format, or to consider it from a different perspective, can be important for solving problems. Similarly, the ability to represent an item (or a piece of information) in multiple codes, despite the unimodal channel by which it may have been perceived, should facilitate one’s ability to manipulate or transform the representation of this information. And what of the representation of the subjective present? The multiple encoding principle holds that, for example, when one is asked to remember the seven digits of a telephone number, there is more to this process than the retrieval into conscious awareness of the seven lexical representations that were a part of one’s knowledge base prior to being given the number to remember. Also represented in working memory can be, for example, information about who spoke the telephone number, about the timber, volume, and tone of the talker’s voice, about one’s affective classification of the talker, about whose telephone number it is, about other telephone numbers that are similar, about the ambient illumination in the room in which the number was spoken, and so on.

**PFC contributions to working memory**

If the two principles outlined in the preceding section are tenets of an emergent property view of working memory, a corollary of the first tenet is that PFC is not a substrate for the storage of information in working memory. Previous sections of this review have summarized neuropsychological, electrophysiological, and neuroimaging evidence that is consistent with this view. There is also ample evidence for contributions of PFC to encoding-, manipulation-, and probe evaluation- and response-related processes (e.g., D’Esposito & Postle, 2002). But an important question remains. Virtually every electrophysiological and neuroimaging study of primate working memory, regardless of the procedure used or of the species tested, finds delay-period activity in the PFC. This includes tasks with unfilled delay periods that would seem to make no overt demands on updating or manipulation operations. If this activity does not correspond to the storage of information, what are alternative explanations of its function? This is one of the most important questions facing contemporary cognitive neuroscience of working memory (e.g., Curtis & D’Esposito, 2003; Lebedev et al., 2004; Passingham & Sakai, 2004), and some alternatives are reviewed below.

**Mediation of interference and distraction**

Interference can come from two sources, external and internal. With regard to the former, working memory deficits are observed in monkeys (Malmo, 1942) and in humans (Chao & Knight, 1995) with PFC lesions when distracting stimuli are present in the environment. One hypothesized contribution of the PFC to working memory function, therefore, is to control the gain of activity in sensory processing areas of posterior cortex (Knight, Staines, Swick, & Chao, 1999), in a manner that would minimize the disruption of working memory storage processes by suppressing the sensory processing of potentially distracting information in the environment.
(Jensen, Gelfand, Kounios, & Lisman, 2002; Klimesch, Doppelmayr, Schwaiger, Auinger, & Winker, 1999; Worden, Foxe, Wang, & Simpson, 2000). Therefore, a portion of PFC delay-period activity that is often observed on unfilled trials may correspond to a basal level of operation of such a distraction detection mechanism. This idea has recently been tested and confirmed (Postle, in press-b). A proposal for a process that accomplishes a similar function, but via a different mechanism, comes from Sakai, Rowe, and Passingham (2002). Their study found that PFC activity during an unfilled delay period predicted task accuracy on trials when the unfilled delay was followed by a distractor task. They attributed this to a PFC-controlled “active maintenance” process that strengthened mnemonic representations via the strengthening of the coupling of activity between superior frontal cortex (BA 8) and Intraparietal Sulcus.

Internally derived, or proactive interference (PI), arises when the products of earlier mental activity disrupt current memory performance. There are two types of PI: *item-specific* PI results when an invalid memory probe matches a memorandum from the preceding trial (Monsell, 1978); whereas *item non-specific* PI is produced by the accumulation of no-longer-relevant items from previous trials (Wickens, Born, & Allen, 1963). There is a great deal of evidence that high levels of item-specific PI are associated with a phasic signal restricted anatomically to Brodmann’s area 45 in the left inferior PFC, and temporally to the memory probe/response epoch of the task (for a review, see the contribution by Jonides and Nee (2005) in this issue). We have recently confirmed that item-nonspecific PI is also associated with this anatomical and temporal profile (Postle & Brush, 2004; Postle, Brush, & Nick, 2004b), making it seem unlikely that the mediation of PI is associated with delay-period activity in PFC. Theoretical interpretations of this PI-related activity include inhibition (Jonides, Smith, Marshuetz, Koepepe, & Reuter-Lorenz, 1998; May, Hasher, & Kane, 1999) (although not response inhibition, Nelson, Reuter-Lorenz, Sylvester, Jonides, & Smith, 2003), selection (Thompson-Schill et al., 2002), probe evaluation processes (Postle et al., 2004b; Zhang, Leung, & Johnson, 2003), “executive attention” (Mecklinger, Weber, Gunter, & Engle, 2003), and “reactive control” (Braver, Gray, & Burgess, in press). These last two perspectives, to be summarized in more detail further along in this section, also offer caveats to the conclusion that PI-related processes do not contribute to delay-period activity in PFC.

**Attention and selection**

From the perspective of cognitive psychology, Engle and Kane and their colleagues propose that “executive attention” is the mental construct that underlies gF and working memory performance (Engle et al., 1999), and they suggest that executive attention underlies much of the working memory-related activity of the PFC, particularly in situations in which interference must be overcome (Kane & Engle, 2003). From neuroscience, Passingham and colleagues argue that PFC activity during working memory tasks reflects attentional selection (Passingham & Rowe, 2002; Rowe, Stephan, Friston, Frackowiak, & Passingham, 2005; Rowe et al., 2000). Lebedev et al. (2004) suggest that the majority of delay-period activity in the PFC relates to attentional monitoring and selection, a view consistent with that of Petrides (1994a; 2000a; 2000b).

**Flexible control**

The variously named “guided activation” (Miller & Cohen, 2001) or “adaptive coding” (Duncan & Miller, 2002) theories emphasize the role of PFC in biasing stimulus-response circuits such that overlearned, prepotent associations can be overcome in favor of novel, or otherwise less salient behaviors, thereby enabling flexible behavioral response to unfamiliar or atypical situations. From this perspective, delay-period activity in PFC may correspond to the representation/maintenance of behavioral goals or task rules, the expectation of reward, or associating two events that are separated in time. Miller and Cohen (2001) have...
also advanced the idea of an “active maintenance” process that supports sustained activity in the face of interference. This view, similar to the executive attention of Kane and Engle (2003), may offer a framework within which to fit many recent empirical observations: that the strength of PFC delay-period activity predicts trial accuracy (as does the strength of delay-period activity in several other regions, Jha, Fabian, & Aguirre, 2004; Pessoa et al., 2002); that the functional connectivity between PFC and posterior regions predicts trial accuracy (Sakai et al., 2002); and that PFC delay-period activity can show load-dependence in a domain-independent manner (Leung et al., 2002; Leung et al., 2004; Narayan et al., 2005; Ranganath et al., 2002a; Rypma et al., 2002). Another example from this class of models is the Dual Mechanisms of Control model (Braver et al., in press), which posits both “proactive” and a “reactive” modes of control, the latter of which was referred to in the earlier discussion of PI. According to this model, the anticipatory application of proactive control (a mode whose use correlates with general fluid intelligence (gF)) might be associated with PFC activity during the encoding and delay epochs of working memory trials, whereas probe-related activity, such as that associated with PI, can correspond to reactive control. Empirical evidence for this idea comes from the fact that activity in PFC regions other than BA 45 – which is strongly implicated in probe epoch-related PI effects – was correlated with performance on high PI trials of an n-back working memory task in high gF individuals (Gray, Chabris, & Braver, 2003). The extent to which the concept of proactive control may relate to effective encoding strategies (e.g., Bor, Duncan, Wiseman, & Owen, 2003; Rypma & D’Esposito, 1999; Rypma & D’Esposito, 2003), or to the possible empirical examples of “active maintenance” described above, remains to be explored.

**Transformation/response preparation**

Several studies have demonstrated PFC involvement in response selection (e.g., Jiang & Kanwisher, 2003; Schumacher & D’Esposito, 2002; Schumacher, Elston, & D’Esposito, 2003). And particularly relevant to the present question are studies demonstrating PFC activity when a response can be prepared in advance of the response cue (i.e., during the delay period, e.g., Curtis et al., 2004; D’Esposito, Ballard, Zarahn, & Aguirre, 2000; Fukushima, Hasegawa, & Miyashita, 2004; Fuster, 1995; Pochon et al., 2001). Relatedly, the PFC may also contribute to transformations of mnemonic representations when such a transformation would permit anticipatory response preparation, such as with the sensorimotor coordinate transformation of a visually presented target stimulus (Fuster, 1995; Takeda & Funahashi, 2002, 2004). The PFC also mediates other types of transformations that can be employed strategically, such as when a chunking strategy would facilitate subsequent retention in memory (Bor et al., 2003). Note that, in this latter case, the mental operations undertaken on a volitional, strategic basis, may be the same as, or similar to, those that are required on tasks that explicitly require “manipulation” of stimuli (e.g., D’Esposito et al., 1999; Postle et al., 2000a).

**Motivation and reward expectancy**

Watanabe (2002) has characterized an important role for the PFC in integrating cognitive and motivational information. For example, the gain of delay-period activity of spatially tuned units in dorsolateral PFC is modulated by the type of reward anticipated on a particular trial (Watanabe, 1996). As working memory research increasingly incorporates the principles and methods of affective neuroscience and neuroeconomics, our appreciation for the contributions of affective, motivational, and trait-related factors to delay-period activity will certainly increase.

In summary, there are many candidate functions other than storage that can account for delay-period activity in PFC. Further, although PFC contributions to working memory performance can be manifold, none of the control processes reviewed here are specific to or specialized for working memory. For example, the mediation of interference is important in many nonworking
memory situations, and distractibility is characteristic of the “lateral prefrontal syndrome” across a wide spectrum of nonworking memory behaviors (e.g., Knight & D’Esposito, 2003). The same can be said for attention, for flexible control, and so on. Indeed, many important theories of PFC and executive control neither limit the range of behaviors to be controlled to working memory, nor posit specialized subcomponents to effect the control of working memory (e.g., Duncan & Miller, 2002; Miller & Cohen, 2001; Norman & Shallice, 1986). From the emergent-processes perspective, then, the control of working memory does not differ qualitatively from the control of any other behavioral or mental function.

Conclusion

When considering the development of the concept of the Central Executive, Baddeley (2002) characterized its early homuncular instantiation as serving a useful “holding function” until such a time that sufficient progress had been made that “the homunculus was no longer necessary, and hence could be pensioned off” (pp. 246–247). By analogy, the standard model has been invaluable in advancing our understanding of working memory. It may be time, however, to retire it in favor of a framework that can more effectively integrate the ever-growing, ever-more-multidisciplinary stream of data that enriches our understanding of this essential cognitive phenomenon.

Acknowledgements

The author receives support from NIH grant MH064498.

References


Curtis, C. E. (2005). Prefrontal and parietal contributions to spatial working memory. *Neuroscience, this special issue*


Jacobsen CF. The functions of the frontal association areas in monkeys. Comparative Psychology Monographs 1936;13:1–60.


