

*Learning and Memory: A Comprehensive Reference, Third Edition*  
Volume: *Systems Neuroscience*.

# Short-Term and Working Memory

Nathan Tardiff & Clayton E. Curtis

New York University, Departments of Psychology and Neural Science

**Keywords:** working memory; memory; prefrontal cortex; visual cortex; parietal cortex; short term memory; functional magnetic resonance imaging; transcortical magnetic stimulation; lesions; saccades; executive function; top-down control

## Key Points

- Working memory is a system composed of both storage and control processes.
- Unlike long-term memory, working memory has severe capacity limits. As the number of items you are storing in working memory increases, the quality of the memory for those items decreases.
- Classic studies with non-human primates support a model in which working memory is supported by persistent neural activity in neurons in the dorsolateral prefrontal cortex.
- Human studies suggest that working memory depends on persistent activity across a widely distributed network of brain regions.
- Sensory cortices are thought to be recruited for storage, as the visual contents of working memory can be decoded from the patterns of activity in visual cortex.
- Even in visual cortex, working memory representations appear to be abstractions of memoranda related to behavioral goals and demands.
- New evidence suggests that prefrontal cortex controls operations of the content of working memory stored in sensory cortices.

## Abstract

Working memory is one of the most important higher-order cognitive abilities, allowing us to hold information that is no longer present “in mind” in preparation for future action. We summarize three key aspects of working memory: its capacity limits, the neural basis of working memory storage, and the control of working memory contents. We emphasize the cognitive and neural distinctions between storage and control, as well as the evolution of the field of working memory research from a focus on the prefrontal cortex to a broader perspective that identifies working memory as arising from a distributed set of brain regions.

## Introduction

Working memory refers to the ability to maintain and manipulate information over brief periods of time, generally on the order of seconds (Miyake and Shah, 1999; Baddeley, 2012). The information stored in working memory often originates from immediately preceding perceptual experience, or it can be information retrieved from long-term memory. For example, when reading a book, making sense of the current sentence often requires that the subject of immediately preceding sentences be remembered. Similarly, when navigating busy city streets, one must recall that they saw their destination across the street while it is temporarily blocked by a bus waiting at a stop light. As these examples demonstrate, working memory abilities are crucial in any temporally-extended task that requires information that is not available to immediate perception, thus forming a bridge between past experience and future action (Goldman-Rakic, 2011; Diamond, 2013; Fuster, 2015). In keeping with its prominence in cognitive functioning, working memory is predictive of academic achievement and general intelligence, and impairments in working memory are found in both healthy aging and neurological and psychiatric conditions (Salthouse et al., 1991; Miyake et al., 2001; Conway et al., 2003; Alloway, 2006; Alloway et al., 2009; Fukuda et al., 2010b; Johnson et al., 2013; Luck and Vogel, 2013).

Functionally, working memory can be divided into components for the storage of information, which are often thought to be domain specific (e.g., visual, verbal, etc.), and components for the control and manipulation of information, which are often considered to be domain general (Baddeley and Logie, 1999; Miyake et al., 2001; Li et al., 2014; Camos, 2017). Terminologically, some researchers refer to the former as *short-term memory* and reserve *working memory* for processes of control and manipulation of stored information (Cowan, 2017). Here, we take a broad view of the construct of working memory that includes both storage and control components. We will review behavioral and neural signatures of both storage and control. In

doing so, we will argue that jointly considering both components and delineating their interactions is crucial for progress in understanding the mechanisms and neurobiological substrates of working memory. We will give particular focus to visual working memory for three reasons: 1) It affords careful control of stimulus properties that form the basis of the memoranda; 2) It permits linkages with the rich literature on the neurophysiological basis of working memory in nonhuman primates; 3) Sophisticated computational methods and analytic techniques exist for probing the mechanisms of visual working memory and uncovering the nature of the representations underlying visual working memory in human subjects. That said, the findings and principles described below should generalize to working memory content originating from other sources.

Working memory has hallmark signatures at the behavioral and neural levels that differentiate it from other forms of memory. First, unlike long-term memory systems, which seemingly allow for a nearly unlimited (within our lifespans) amount of semantic, episodic, and procedural information to be stored, working memory is severely capacity-limited, typically able to hold only 3-9 items at a time, depending on individual ability and task conditions (Miller, 1956; Cowan, 2001). Second, though long-term memory is fallible and subject to distortions and forgetting, working memory is extremely fragile: it persists on the order of seconds, decays rapidly without active maintenance processes, such as rehearsal, and is susceptible to distraction (Miyake and Shah, 1999; Lorenc et al., 2021). Finally, while long-term memories are created, updated, and maintained via synaptic changes, working memory is primarily reliant on active neural firing (Wang, 2021). In the remainder of the chapter, we will describe three key foundations for understanding working memory: *capacity limits*, *storage*, and *control*.

## Capacity limits in working memory

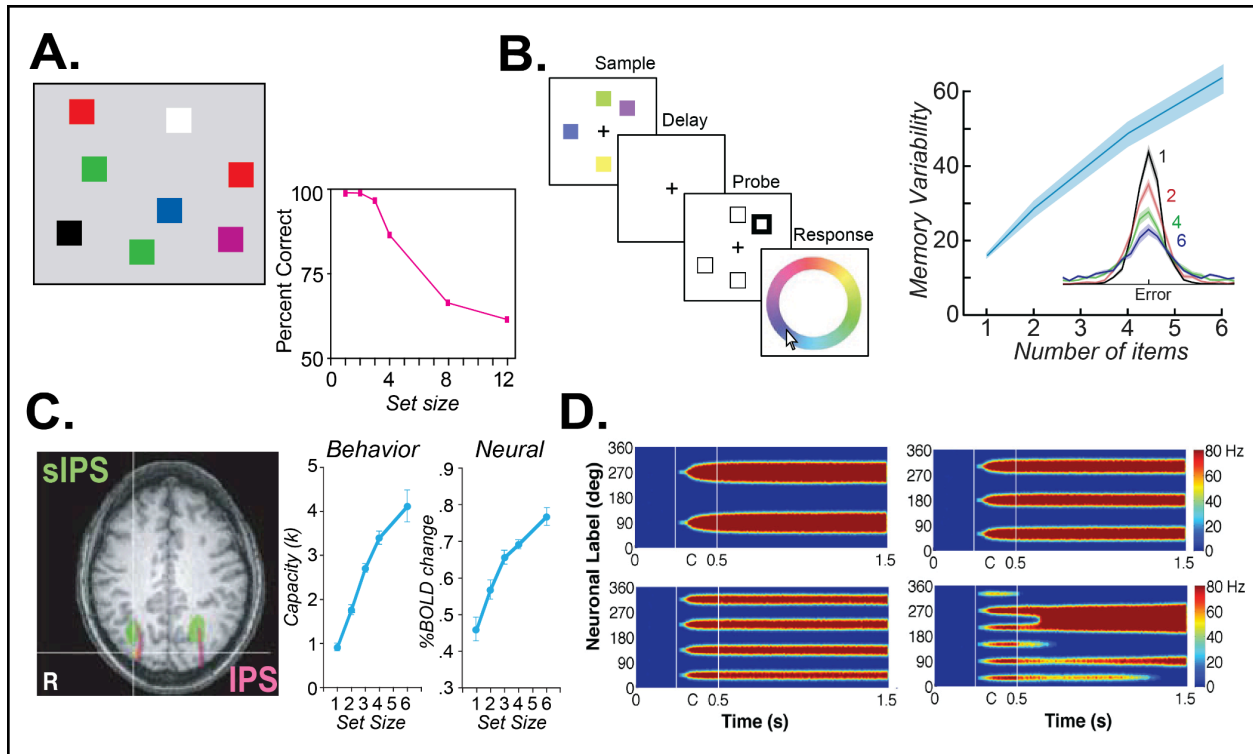
Unlike long-term memory, working memory has severe capacity limits. We can only store so many items in working memory, after which our memory performance declines. While this hallmark of working memory has long been recognized (Miller, 1956), the past few decades have seen considerable research devoted to understanding the nature of those limits (Cowan, 2001; Fukuda et al., 2010a; Luck and Vogel, 2013; Ma et al., 2014). In visual working memory, change-detection tasks have commonly been used to assess working memory capacity limits (Luck and Vogel, 1997). In such tasks, participants are briefly shown an array of items on a screen (e.g., colored squares); after a delay, the display reappears, and participants are tasked with reporting whether an item in the display changed from the first presentation to the second (Figure 1A). By varying the number of items in the display and examining the impact on change-detection performance, a capacity limit in terms of the number of items that can be held in mind can be computed. Across a number of experiments, these paradigms yield a capacity limit of around 3–4 items (Fukuda et al., 2010a). On this basis, one class of highly influential theoretical models proposes that working memory is structured into a limited number of discrete slots, mirroring behavioral capacity limits; when the number of items to be stored exceeds this capacity and an item is not held in working memory, the participant is thought to resort to randomly guessing whether it changed (Luck and Vogel, 1997; Cowan, 2001; Awh et al., 2007; Zhang and Luck, 2008; Rouder et al., 2011). Though exactly how slots would arise in the brain is unclear (Fukuda et al., 2010a), a region of posterior parietal cortex (PPC) increases activity with load up to 4 items, at which point its activity plateaus, in line with the working memory capacity limit estimated from behavior (Todd and Marois, 2004; Xu and Chun, 2006) (Figure 1C).

Objections to the discrete slots model of working memory originated from the observation that, as set size increases, the precision with which individual items are remembered decreases,

even beyond 3-4 items (Palmer, 1990; Wilken and Ma, 2004a; Bays and Husain, 2008). These observations were often made on the basis of continuous-report paradigms, which ask participants to reproduce a stimulus value from the memory array, unlike change-detection paradigms, which require only a binary judgment (Figure 1B). The observation of a relationship between set size and encoding precision led to the development of an alternative theoretical model in which working memory is subserved by a continuous resource, which is divided among the items held in working memory (Wilken and Ma, 2004a; Bays and Husain, 2008; Ma et al., 2014). On this account, decreases in precision with set size arise because less resource is available to each item as it is further divided among the items, and guessing behavior occurs not because an item failed to be held in a particular slot, but because the resource itself is noisy and variable, such that on some trials the noise will be so high the item is essentially not recoverable (van den Berg et al., 2012). While the model is abstract, its proponents identify the resource as likely corresponding to noisy neural populations encoding the memorandum (van den Berg et al., 2012; Ma et al., 2014).

Given that working memory has qualities that appear slot-like alongside qualities that appear resource-like, a number of attempts have been made to bridge across these models, including by allowing the individual slots to behave like noisy resources (Zhang and Luck, 2008). Other work has demonstrated how slot-like capacity-limited behavior can arise from noisy neural populations encoding working memory, either due to dynamics inherent in the neural population (Wei et al., 2012; Standage and Paré, 2018) (Figure 1D), or from the manner in which information is sampled from noisy memory representations (Schneegans et al., 2020). As such, these models support the idea that the underlying resource is continuous. An alternative proposal is that the capacity limits arise not from constraints on the storage of working memory *content*, but from limits on the number of items that can be individuated and tracked by a content-independent pointer system (Thyer et al., 2022); see also (Xu and Chun, 2006). A

similar pointer system has been implemented in a biologically inspired model of prefrontal-basal ganglia circuits (Kriete et al., 2013).



**Figure 1. Evidence for capacity limits in WM.** **A.** Left: In a typical change-detection task, subjects view a sample array of stimuli, followed by a brief delay. After the delay, a test array appears on screen, and subjects report whether the array is identical or whether a feature (e.g., the color of one of the squares) has changed. Alternatively, an item at a particular location may be probed at test, rather than redisplaying the entire array. Right: Accuracy of reporting a change decreases as the set size of the array increases. Adapted from Figure 1 of (Luck and Vogel, 1997). **B.** Left: In a color delayed-estimation task, a typical continuous-report paradigm, rather than responding as to whether an item in the array changed after the delay, subjects are asked to reproduce the color of a probed item. Right: Memory variability, measured as the standard deviation of the distribution of memory errors relative to the true color, increases with increasing set size. Adapted from Figure 1 of (Wilken and Ma, 2004b). **C.** Left: Superior intraparietal sulcus (sIPS; green)—a subregion of the PPC—in an example subject. Middle: Working memory capacity ( $k$ ) measured using a change-detection task and averaged across subjects. Capacity shows a plateau at about 4 items. Right: Neural activity in sIPS demonstrates a similar plateau and correlates with behavioral capacity. Adapted from Figure 2 of (Xu and Chun, 2006). **D.** Simulated neural activity from excitatory units in a neural network model of working memory presented with set sizes of 2, 3, 4, or 6 items. While the network successfully encodes up to 4 items, at set size 6 the activity encoding each item is prone to fading away or merging (bottom right). Neurons are labeled by their preferred stimulus, in degrees (ordinate; e.g., coding for an angle in 2D space), and their activity is plotted across time (abscissa), where C marks the stimulus presentation epoch. Color corresponds to firing rate. Adapted from Figure 2 of (Wei et al., 2012).

## Working memory storage

### Persistent activity in prefrontal cortex as a neural substrate of working memory storage

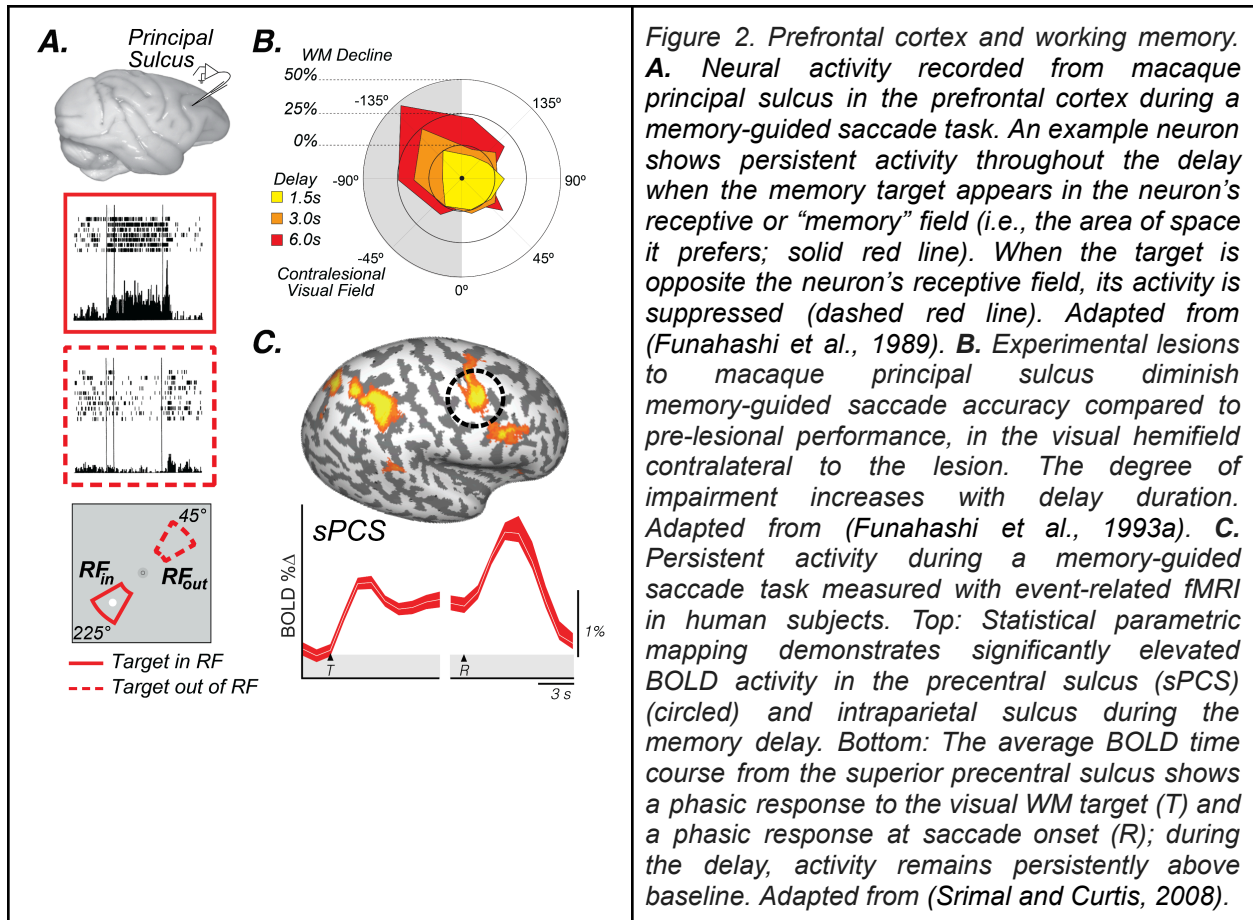
While many researchers now consider working memory to be a cognitive function with distributed substrates throughout the brain (D'Esposito and Postle, 2015; Christophel et al., 2017; Sreenivasan and D'Esposito, 2019; Curtis and Sprague, 2021), the prefrontal cortex has long retained a “preeminent role” in neurobiological accounts of working memory (Goldman-Rakic, 1995). Appreciation for the role of prefrontal cortex in short-term memory processes can be traced back to the studies of Jacobson (Jacobsen, 1935, 1936). When nonhuman primates were confronted with a delayed response task in which they had to search for food hidden in one of two possible locations after a brief delay, animals with bilateral prefrontal lesions were at chance in identifying the baited location, displaying a selective deficit in “immediate memory” (Jacobsen, 1935). Several decades of lesion studies building on this work established that integrity of the lateral prefrontal cortex, and more specifically the principal sulcus (areas 46 and 9/46d of the macaque) was necessary for performance in spatial delayed response tasks (as reviewed in (Curtis and D'Esposito, 2004).

Subsequently, extracellular recording techniques enabled the identification of neural activity correlated with working memory. In two pioneering studies, one using a delayed response task, as described above (Fuster and Alexander, 1971), and another using a delayed alternation task, in which subjects were required to alternate between pulling two levers, where a delay between trials induces a memory requirement (Kubota and Niki, 1971), recordings from the area of the macaque principal sulcus discovered neurons that maintained firing rates above baseline during the delay period of the task, a phenomenon now termed *persistent activity*. Funahashi, Bruce, and Goldman-Rakic (Funahashi et al., 1989) developed an oculomotor version of the delayed response task (also known as a memory-guided saccade (MGS) task), which among other



refinements allowed them to probe the relationship between neural activity and specific locations in the visual field. Using this paradigm, they demonstrated that prefrontal cortex neurons with persistent activity were stimulus selective, such that a given neuron typically only responded to one or perhaps two nearby target positions in the contralateral hemifield (Figure 2A). Additionally, lesions of the prefrontal cortex tended to impair memory for targets in the contralesional hemifield (Funahashi et al., 1993a) (Figure 2B).

Together, these studies established fundamental features of prefrontal cortex persistent activity that have been replicated numerous times (as reviewed in (Riley and Constantinidis, 2016)). Persistent activity: 1) Spans the delay, adapting in duration to delays of varying lengths (Fuster and Alexander, 1971; Funahashi et al., 1989); 2) correlates with memory accuracy (Fuster, 1973; Funahashi et al., 1989; Wimmer et al., 2014); see (Rezayat et al., 2022) for review); and 3) Is selective for particular memoranda (Funahashi et al., 1989; Wimmer et al., 2014). These features are consistent with a mechanism that bridges the gap between a past sensory event and subsequent contingent behavior and align closely with cognitive notions of working memory, such that prefrontal persistent activity was identified as the neural basis of working memory (Goldman-Rakic, 1995).



While these fundamental features of persistent activity were established in dorsolateral prefrontal cortex during visuospatial working memory, they have been extended across frontal areas, domains, and species. Visuospatial persistent activity is actually more common, robust, and spatially-selective in neurons of the frontal eye fields (FEF) than dorsolateral prefrontal cortex (Goldberg and Bruce, 1985; Funahashi et al., 1989; Sommer and Wurtz, 2001; Merrikhi et al., 2017; Hart and Huk, 2020). Persistent activity is found in many other frontal areas including the supplementary eye fields (SEFs) (Shichinohe et al., 2009; Fukushima et al., 2011), orbitofrontal cortex (OFC) (Ichihara-Takeda and Funahashi, 2007), and in the dorsal premotor cortex (pMD) (Rossi-Pool et al., 2017; Bastos et al., 2018). Outside the coding of space, prefrontal cortex has neurons that demonstrate persistent activity during working memory tasks for both simple (e.g., color) and complex (e.g., face) objects (Quintana et al., 1988; Wilson et al.,

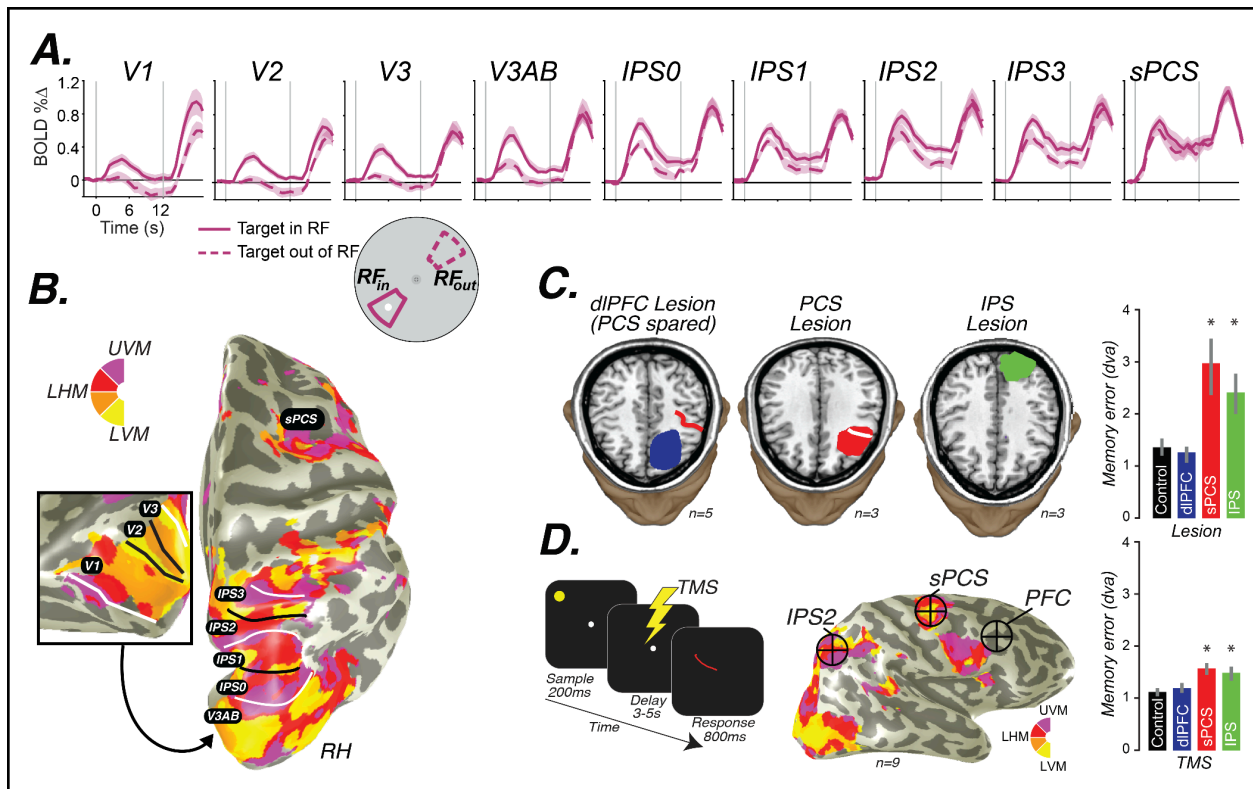
1993; Miller et al., 1996a; Rao et al., 1997; Hoshi et al., 1998; Scialidhe et al., 1999; Fuster et al., 2000; Rainer and Miller, 2000; Constantinidis et al., 2001a; Freedman et al., 2001; Sakagami et al., 2001; Averbeck et al., 2003; Inoue and Mikami, 2006; Genovesio et al., 2009; Roy et al., 2014; Panichello and Buschman, 2021) and the direction of dot motion (Zaksas and Pasternak, 2006; Mendoza-Halliday et al., 2014). Nor are these findings restricted to the visual domain: there are prefrontal cortex neurons that are selective for vibrotactile frequency (Romo et al., 1999), sound location (Fuster et al., 2000; Kikuchi-Yorioka and Sawaguchi, 2000), and audiovisual macaque vocalizations (Hwang and Romanski, 2015). Furthermore, encoding is present not only for retrospective information (i.e., goal-relevant stimuli), but also for prospective motor plans (Funahashi et al., 1993b; Takeda and Funahashi, 2002; Markowitz et al., 2015) and prospective sensory features of delayed paired associates (Rainer et al., 1999; Fuster et al., 2000). In summary, persistent neural activity during delays is thought to carry any information stored in working memory that is useful for memory-guided behavior.

The prominence of persistent activity in prefrontal cortical neurons led Goldman-Rakic to propose that the prefrontal cortex is the principal site of working memory storage (Goldman-Rakic, 1995, 2011), a theory that continues to be highly influential (Riley and Constantinidis, 2016). Nonetheless, persistent activity during working memory has now been identified in every lobe of the brain, as well as in subcortical structures (Leavitt et al., 2017; Sreenivasan and D'Esposito, 2019). Within parietal cortex, neurons in areas LIP and 7a also show spatially selective and robust persistent activity (Gnadt and Andersen, 1988; Barash et al., 1991; Constantinidis and Steinmetz, 1996; Chafee and Goldman-Rakic, 1998; Pesaran et al., 2002; Hart and Huk, 2020). In the temporal lobe, persistent activity is found in monkey inferotemporal (IT) cortex (Fuster and Jervey, 1981; Miyashita and Chang, 1988; Miller et al., 1993a; Chelazzi et al., 1998). Activity also persists in the medial temporal lobe (Miller and Desimone, 1994; Suzuki et al., 1997; Wirth et al., 2003) and in subcortical areas, including the

mediodorsal thalamus (Fuster and Alexander, 1971; Funahashi, 2013), and superior colliculus (SC) (Shen et al., 2011; Dash et al., 2015; Sadeh et al., 2018). Therefore, and to preview the evidence for persistent activity in humans described next, one challenge for future research is to understand the significance of such widespread and distributed working memory signals.

### **Persistent activity measured in humans**

Using functional magnetic resonance imaging (fMRI), persistent activity has been measured in humans. Mirroring studies in the macaque, elevated BOLD activity in prefrontal cortex during the delay period of working memory tasks is present for visuospatial (Courtney et al., 1998; Zarahn et al., 1999; Leung et al., 2002; Sakai et al., 2002; Curtis et al., 2004; Curtis and D'Esposito, 2006; Srimal and Curtis, 2008) (Figure 2C), verbal (Barch et al., 1997; Cohen et al., 1997; Narayanan et al., 2005), auditory (Tark and Curtis, 2009; Kumar et al., 2016; Uluç et al., 2018), and vibrotactile (Preuschhof et al., 2006; Schmidt et al., 2017) memoranda. Beyond prefrontal cortex, persistent activity is found in human posterior parietal cortex (Todd and Marois, 2004; Xu and Chun, 2006; Xu, 2018) and the temporal lobe (Ranganath et al., 2004; Lepsien and Nobre, 2007), including in single neurons of the medial temporal lobe, as measured with invasive recordings in epilepsy patients (Kamiński et al., 2017; Kornblith et al., 2017; Boran et al., 2022). Persistent activity is also found in subcortical areas, including the caudate nucleus (Postle and D'Esposito, 1999), the lateral geniculate nucleus (LGN) (Rahmati et al., 2023), and the SC (Rahmati et al., 2020). Furthermore, the magnitude of persistent activity spans delays of varying lengths (Srimal and Curtis, 2008) and correlates with behavioral performance (Pessoa et al., 2002; Sakai et al., 2002; Curtis et al., 2004).



**Figure 3. Identifying working memory regions in human cortex.** **A.** Persistent activity across brain areas identified in **B.**, measured with fMRI. BOLD activity persists during the delay of an MGS task in many areas. Solid lines are average BOLD activity when the memory target was within the pRF of the voxels. Dashed lines are when the memory target was 180 degrees away from the pRFs. Error bands are standard errors of the mean (SEM). Gray vertical lines mark the onset of the visual target and the end of the delay. Adapted from (Hallenbeck et al., 2021a). **B.** Population receptive field map of retinotopically-organized regions of visual, parietal, and frontal cortex in an example human subject's right hemisphere. Color indicates an area's preferred polar angle in the contralateral visual field (i.e., segment of visual space, see legend). Each labeled area has its own field map of contralateral visual space. Adapted from (Mackey et al., 2017). **C.** Lesions to human sPCS and IPS impair working memory. Dorsolateral prefrontal cortex (dIPFC) lesions that spared sPCS had no effect on memory. Left: Lesion sites in example patients. Right: Average memory error (SEM) on an MGS task for targets in the contralesional visual field, as a function of lesion site. Adapted from (Mackey et al., 2016a, 2016b). **D.** TMS to sPCS and IPS, but not dIPFC, impairs working memory, replicating the lesion results. Left: MGS task. TMS was administered during the delay period. Middle: Stimulation targets from the brain of an example subject. Right: MGS performance, as in **C.**

### Beyond persistent activity in prefrontal cortex: Evidence from early visual cortex

Beyond IT, a high-level visual region, whether persistent activity exists in sensory regions generally—particularly lower-level sensory regions—has proven controversial, with inconsistent findings across studies in visual, auditory, and somatosensory areas (Leavitt et al., 2017).

Focusing on visual cortex, while evidence of persistent activity is found less often than in higher-order regions (Leavitt et al., 2017), some studies have identified persistent activity even in V1, the earliest cortical visual area (Fuster, 1990; Supèr et al., 2001; Van Kerkoerle et al., 2017; Dotson et al., 2018). There are a number of possible reasons for discrepancies in the presence of persistent activity between studies.

First, the presence or absence of persistent activity may depend on the nature of the memorandum. One study found a lack of persistent activity for the memorized direction of random dot motion in motion-selective visual area MT, while persistent activity was found in later visual area MST and in lateral prefrontal cortex, leading the authors to conclude that persistent activity only emerges later in the cortical hierarchy (Mendoza-Halliday et al., 2014). However, whether or not activity persists in an area may depend on the format of the memorandum, which could differ from the format of the perceptual representation. In the case of dot motion, it seems unlikely that memory consists of a replay of hundreds of moving dots; perhaps instead the percept is compressed or recoded into something like a directional vector, which does not drive MT because it does not contain active motion.

Second, delay-period activity in early visual areas, even when stimulus selective, is sometimes found to be below the pretrial baseline (Supèr et al., 2001; Bisley et al., 2004). Some have dismissed these results as a lack of persistent activity because they do not resemble canonical delay-period activity established in dorsolateral prefrontal cortex (Riley and Constantinidis, 2016). However, given the different cortical architectures in dorsolateral prefrontal cortex and visual cortex and likely different roles of these regions in working memory, it should not be surprising that the nature of their delay-period activity can differ.

Finally, delay-active neurons may be relatively less numerous and more topographically isolated in sensory areas. Using large-scale recordings, a recent study was able to find persistently

active neurons in macaque V1 and V2, but the relative proportion of these neurons participating in delay-period activity tended to decrease over the delay, such that there was a much lower incidence of delay-active neurons in V1 and V2 in the late delay period compared to higher areas (Dotson et al., 2018). Similarly, another study found persistent activity in V1, but only in superficial and deep layers of cortex, suggesting the increases in multi-unit activity are due to top-down signals (Van Kerkoerle et al., 2017).

Findings from human fMRI support the premise that in earlier visual areas, delay-period activity may be both below baseline and more topographically specific than in higher areas. Analyses that average over all the voxels in an area fail to find persistent activity in early visual cortex (Ester et al., 2009; Harrison and Tong, 2009; Offen et al., 2009; Serences et al., 2009; Riggall and Postle, 2012; Albers et al., 2013). However, this approach is rather crude, especially compared to the standard approach in nonhuman primate electrophysiology, which involves 1) characterizing each neuron's receptive field (RF), or preferred stimulus feature during perception (e.g., area of the visual field); and 2) contrasting delay-period activity for stimuli placed inside and outside of each neuron's RF (e.g., comparing preferred to non-preferred or anti-preferred stimuli). The development of population receptive field (pRF) mapping (Dumoulin and Wandell, 2008; Wandell and Winawer, 2015; Mackey et al., 2017) has enabled researchers to contrast in-RF and out-RF responses at the voxel level as well. This technique can be used to characterize the part of the visual field preferred by an MRI voxel, representing the location in visual space that—when a stimulus is present—is most dominant in driving the activity of the underlying population of tens of thousands of neurons sampled by the voxel.

Utilizing this technique, Hallenbeck et al. (Hallenbeck et al., 2021a) found that delay-period activity in V1 was higher for in-RF compared to out-RF voxels, though activity tended to be below baseline overall during the late-delay period (Figure 3A&B). These results closely mirror those of Supèr et al. (Supèr et al., 2001) in macaque V1, who also found that activity during the

delay dipped below baseline but was higher for in-RF stimuli. Interestingly, two gradients can be seen in the fMRI data, moving in opposite directions along the visual hierarchy. First, the overall amplitude of persistent activity increases moving up the hierarchy from visual to frontal cortex. Second, the spatial selectivity—captured by the difference between in-RF and out-RF voxels—generally increases going down the hierarchy from frontal to visual cortex. This pattern lends further support to the idea that stimulus-selective delay activity in visual regions may not be absent but rather different in character than in prefrontal and parietal areas.

Furthermore, recently developed techniques based on pRF mapping or inverted encoding models (Brouwer and Heeger, 2009; Serences and Saproo, 2012) of fMRI data have allowed researchers to create visual reconstructions of the contents of working memory in the same feature domain as the stimulus (e.g., orientation or polar angle in visual space). These methods consistently reveal activity corresponding to the memorized feature in visual cortex, as well as in prefrontal and parietal cortices (Kok and de Lange, 2014; Hallenbeck et al., 2021a; Kwak and Curtis, 2022; Li and Curtis, 2023). Because these methods rely on relative differences in activity among voxels sensitive to different stimuli (e.g., different regions of visual space), they can ultimately be thought of as more sensitive measures of stimulus-selective persistent activity (Curtis and Sprague, 2021).

In sum, the most widely agreed upon mechanism underlying working memory—memoranda-specific persistent activity—is not exclusive to prefrontal cortex. Rather it appears in many parts of the brain. The ubiquity of persistent activity raises important questions that continue to vex researchers about the relative roles of each region in supporting working memory, as well as the nature of interactions between regions.

### **Questions regarding the role of persistent activity in prefrontal cortex**



One of these crucial questions is: If persistent activity is found all over the brain, does that imply that all of these regions play a role in working memory storage, or can persistent activity serve different purposes depending on where it is found? Work in cognitive psychology has long suggested that working memory requires both storage functions and processes that act over stored representations. Notably, Baddeley's seminal psychological model of working memory (Baddeley and Hitch, 1974; Baddeley and Logie, 1999) established a distinction between domain-specific storage buffers that hold the contents of working memory (the phonological loop for verbal information and the visuospatial sketchpad for visual information), and a central executive responsible for controlling and coordinating the operations of working memory, including allocating limited-capacity attentional resources, filtering distracting or irrelevant stimuli that could interfere with existing memoranda, and updating the contents of working memory based on task goals. Well prior to many of the findings of widespread persistent activity outlined above, the identification of lateral prefrontal cortex with storage, as opposed to the processing requirements of working memory, was called into question.

First, lateral prefrontal cortex lesions do not always impair working memory (Curtis and D'Esposito, 2004). Notably, Malmö (Malmö, 1942) found that macaques with bilateral resection of the frontal lobes were able to perform a delayed response task if conducted in darkness, which implied that the frontal lobes are not involved in the storage of working memory but rather in protecting the contents of working memory from interference by distracting stimuli during the delay period. Second, when researchers began trying to translate findings regarding the role of dorsolateral prefrontal cortex in spatial working memory from nonhuman primates to human neuroimaging, they encountered a serious discrepancy. Several imaging studies failed to find spatial-working-memory-related delay-period activity in the presumed homologous part of human dorsolateral prefrontal cortex (Jonides et al., 1993; Smith et al., 1996; Courtney et al., 1998; Zarahn et al., 1999; Rowe et al., 2000). This failure does not appear to be caused by a

lack of sensitivity to detect persistent activity associated with working memory for a single item using fMRI, as subsequent work demonstrated persistent activity in the superior precentral sulcus (sPCS) in the frontal cortex and the posterior part of the intraparietal sulcus (IPS) during an MGS task (Curtis et al., 2004; Curtis and D'Esposito, 2006; Schluppeck et al., 2006; Srimal and Curtis, 2008; Tark and Curtis, 2009; Jerde et al., 2012; Sprague et al., 2014; Saber et al., 2015; Rahmati et al., 2020; Hallenbeck et al., 2021a). Further ruling out a limitation of fMRI, memory-guided saccade accuracy was spared by dorsolateral prefrontal cortex lesions, whereas sPCS and IPS lesions increased memory errors to targets in the contralesional hemifield (Mackey et al., 2016a, 2016b; see also Ploner et al., 1999) (Figure 3C). Ruling out factors such as cortical reorganization or compensation, transcranial magnetic stimulation (TMS) applied to sPCS, IPS, and the dorsolateral prefrontal cortex of healthy subjects during the delay period recapitulated the lesion findings: TMS to the sPCS, but not dorsolateral prefrontal cortex, caused an increase in memory-guided saccade errors (MacKey and Curtis, 2017) (Figure 3D).

Additional findings from fMRI suggest that dorsolateral prefrontal cortex is activated by task demands outside a pure requirement for storage (e.g., D'Esposito et al., 1999; Rowe et al., 2000; Sakai et al., 2002). One might think that an area involved in storage would be sensitive to load, or the number of items required to be held in mind. However, dorsolateral prefrontal cortex appears insensitive to load unless storage demands increase beyond the capacity limit of working memory, a pattern of results that is more consistent with a role in the strategic organization and compression of memoranda (e.g., chunking), rather than storage (Rypma et al., 2002; Bor et al., 2003; Curtis and D'Esposito, 2003). Other evidence indicates that dorsolateral prefrontal cortex activity is sensitive not to storage requirements but to the need to select among competing responses or representations (Passingham, 1985; Rowe et al., 2000; Rowe and Passingham, 2001). Indeed, dorsolateral prefrontal cortex is sensitive to selection

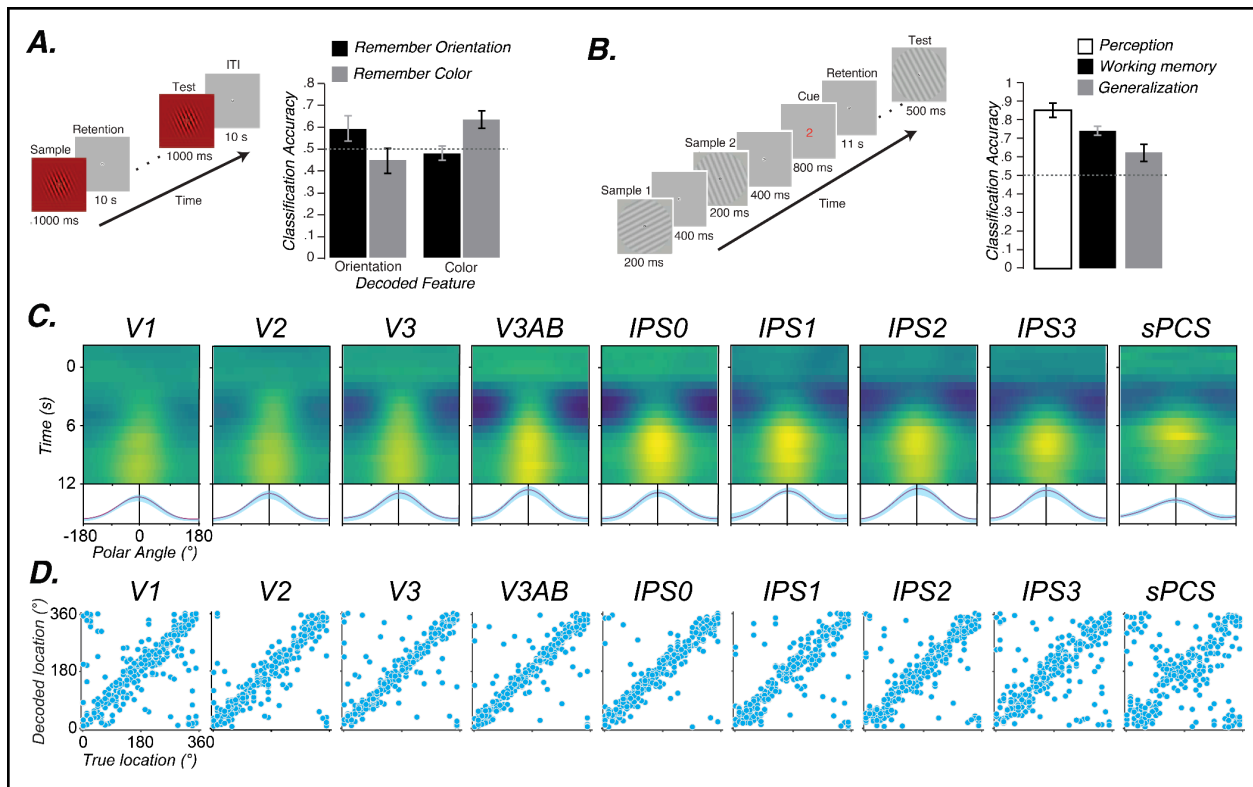
demands even in the absence of working memory requirements (Schumacher and D'Esposito, 2002; Schumacher et al., 2003), suggesting that it subserves rule-based response selection, not working memory storage per se. These and related findings have led to a variety of proposals that the primary role of lateral prefrontal cortex is not working memory storage but executive or cognitive control operations, which allow for the selection and filtering of mental representations and the planning and execution of behavior in service of task goals, particularly in the face of conflict or distraction (Petrides, 2000; Shimamura, 2000; Miller and Cohen, 2001; Chrysikou et al., 2014; Fuster, 2015). Furthermore, these findings argue against the basic premise that delay-period persistent activity should be associated primarily with storage, not only in prefrontal cortex but anywhere in the brain (Curtis and D'Esposito, 2003). This raises two questions: 1) Where might the primary site(s) of working memory storage reside, if not prefrontal cortex? 2) Besides persistent activity, what criteria can we use to identify these sites and differentiate storage from other processes?

### **The sensory recruitment hypothesis for working memory storage**

While the preceding evidence is suggestive of regions that may be involved in working memory storage, including sPCS and PPC, one of the most prominent current hypotheses is that working memory storage is subserved by sensory cortex. This *sensory-recruitment hypothesis* is not a specific, mechanistic proposal, but rather a general framework positing that the neural mechanisms that subserve sensory and perceptual processing are used for working memory. Moreover, it also assumes that the role of prefrontal cortex is to exert top-down control over the representations held in sensory areas (Curtis and D'Esposito, 2003; Pasternak and Greenlee, 2005; Postle, 2006; D'Esposito, 2007; D'Esposito and Postle, 2015; Serences, 2016).

The genesis of this hypothesis stemmed in part from many of the observations described above, including a lack of delay-period activity in dorsolateral prefrontal cortex during simple

memory-maintenance tasks measured with fMRI, combined with the discovery that information about what one has stored in working memory exists in sensory cortex. A methodological breakthrough occurred around the turn of the millennium, when neuroimaging researchers began to analyze multivariate patterns of activation across voxels in a region as a means of discriminating between different stimuli or conditions. Using simple machine learning classifiers, they found that the contents of perception could be decoded based on the different patterns that were evoked by different stimuli, so-called *multi-voxel pattern analysis* (MVPA; (Haxby et al., 2001; Norman et al., 2006). This advance represented an important increase in sensitivity over traditional *mass univariate* approaches to analyzing fMRI data, which can only distinguish elevated or suppressed average activation between conditions. Such an increase in sensitivity to fine-grained patterns of activation may be particularly important for examining working memory phenomena in early sensory areas, where it might be expected that only certain small subpopulations of neurons tuned to the memorandum (e.g., a specific visual orientation) would be active during the delay (Serences et al., 2009; Curtis and Sprague, 2021). Indeed, in two seminal papers, multivariate classifiers were trained on fMRI activation patterns measured from visual cortex during the delay period of working memory tasks (Harrison and Tong, 2009; Serences et al., 2009). These studies demonstrated that the memorandum could be successfully decoded from delay-period activation patterns. As an important control, they showed that visual information that was presented in the same trial but was not cued to be remembered [either another stimulus (Harrison and Tong, 2009) or another feature of the same stimulus (Serences et al., 2009)] was not maintained, ruling out any confound from lingering sensory-evoked activation that might persist in the slow hemodynamic signals measured with fMRI (Figure 4A&B).



**Figure 4.** Decoding working memory in early visual cortex and beyond. **A.** Left: Participants were asked to remember the orientation or the color of a sample stimulus over a delay in separate blocks. Right: The task relevant, but not irrelevant, feature was decodable from the pattern of BOLD activity in V1, suggesting that behavioral goals impact the stimulus features maintained in V1 during the delay. Adapted from (Serences et al., 2009). **B.** Left: Using a retro-cue design, participants maintained the cued grating in working memory over a delay. Right: Using pattern classification of voxel BOLD activity in V1, the orientation in working memory could be decoded (black bar), but not as accurately as during a perceptual task (white bar). When training on one task and testing on the other, classification dropped (gray bar) suggesting differences in the format of perceptual and working memory representations. Adapted from (Harrison and Tong, 2009). **C.** ‘Inverted encoding models’ (IEM; (Brouwer and Heeger, 2009) model the activation of each voxel as a weighted combination of neural information channels mapping a feature space, here angular visual space. Shown are average reconstructed channel response profiles for each timepoint across memory-guided saccade trials (i.e., the activation of polar angle representations inferred from voxel activations), for several visual field maps. Targets are aligned at 0°, and the bottom inset is the average channel response profile over the final 1.5 s of the delay (blue line). Data adapted from (Hallenbeck et al., 2021b). **D.** Using a generative model to estimate working memory representations (van Bergen & Jehee, 2021), one can decode precise memory content, here from an example participant. For each visual field map, the decoded location as a function of the memorized target location is plotted, where each dot is a trial. Notice how closely the points lie along the diagonal, as well as the decrease in decoded accuracy as one moves up the visual hierarchy. Adapted from (Li et al., 2021a).

Subsequently, dozens of studies using a variety of decoding and encoding methods have found that the patterns of activity in visual, parietal, and frontal cortex contain information about visual features held in working memory (Jerde et al., 2012; Riggall and Postle, 2012; Albers et al.,

2013; Emrich et al., 2013; Ester et al., 2013, 2015; Lee et al., 2013; Xing et al., 2013; Sprague et al., 2014, 2016; van Bergen et al., 2015; Bettencourt and Xu, 2016; Yu and Shim, 2017; Christophel et al., 2018a, 2018b; Lorenc et al., 2018; Rahmati et al., 2018, 2020; Rademaker et al., 2019; Brissenden et al., 2021; Hallenbeck et al., 2021a; Li et al., 2021b; van Bergen and Jehee, 2021; Miller et al., 2022; Master et al., 2023); see (Christophel et al., 2017) for review) (Figure 4C&D). Stimulus-selective patterns during working memory have also been found in subcortical structures, including the SC (Rahmati et al., 2020), LGN (Rahmati et al., 2023), and cerebellum (Brissenden et al., 2021). These methods have also successfully identified working memory contents in other modalities, including auditory (Linke et al., 2011; Kumar et al., 2016; Uluç et al., 2018; Czoschke et al., 2021; Deutsch et al., 2023) and vibrotactile (Schmidt et al., 2017).

In support of the sensory recruitment hypothesis, the quality of working memory representations in early visual areas predicts behavioral performance (Emrich et al., 2013; Ester et al., 2013; Bettencourt and Xu, 2016; Hallenbeck et al., 2021a; Li et al., 2021b). While quality also predicts performance in parietal and frontal areas (Bettencourt and Xu, 2016; Li et al., 2021b), an interesting pattern emerges when the cortical topography of sustained delay-period activation is contrasted with stimulus-selective activation patterns: regions with elevated delay-period activity in frontal and parietal cortices tend to have robust persistent activity but weak or absent decoding, while visual regions tend to show weak or absent persistent activity but robust decoding (Riggall and Postle, 2012; Emrich et al., 2013; Hallenbeck et al., 2021a; Li et al., 2021b) (Compare Figures 3A and 4D). These counter-gradients of delay-period activity and decodability have been taken as further evidence that frontal and parietal regions serve to control or coordinate working memory representations, whereas sensory regions encode the representations themselves (D'Esposito and Postle, 2015; Curtis and Sprague, 2021) .

### **Evaluating arguments for regional specialization for working memory storage**

We now turn to the second question raised above: Besides the mere existence of persistent activity, what criteria can we use to identify a region as having a role in working memory storage (or other working memory processes)? Here we will survey and evaluate the primary arguments related to this question that are present in the working memory literature. Arguments can be divided into two main classes: *arguments from architecture*, which infer regional involvement in working memory from neural architecture, and *arguments from function*, which infer involvement in working memory from known functional specializations of each region (e.g., their coding properties) or from appeals to functional organizations that would lead to putatively more efficient or robust neural processing.

### ***Neural architecture: the canonical microcircuit model of working memory***

Given that stimulus-selective persistent activity is the primary neural signature of working memory, it is important to identify what features of cortical microarchitecture support this activity. Efforts to do so started in prefrontal cortex with a number of complementary observations. In layer III of the prefrontal cortex, pyramidal (excitatory) neurons synapse onto clusters of other layer III pyramidal neurons at regular intervals (Levitt et al., 1993; Lund et al., 1993; Kritzer and Goldman-Rakic, 1995). This architecture is reminiscent of V1, where a motif of patchy horizontal connectivity (Gilbert and Wiesel, 1983) links neurons that tend to have similar orientation tuning (Gilbert and Wiesel, 1989). By analogy with V1, Goldman-Rakic proposed that pyramidal neurons in layer III of the prefrontal cortex have excitatory recurrent connectivity with other similarly-tuned neurons, which enables stimulus-selective persistent activity via recurrent excitation, while inhibitory interneurons further sculpt this activity by suppressing excitatory units with dissimilar tuning to those coding for the memorandum (Goldman-Rakic, 1995). Goldman-Rakic's theory was successfully translated into a computational model, demonstrating that memory-selective persistent activity can in fact be generated via excitatory recurrent dynamics balanced and sculpted by inhibition (Compte et al., 2000; Wang, 2001). The model

further suggested that the slow kinetics of excitatory NMDA receptors, compared to the faster kinetics of AMPA receptors (Wang, 1999), were necessary for activity to persist.

Empirical evidence largely supports this prefrontal cortex microcircuit model. Persistent activity of similarly-tuned prefrontal cortex neurons are correlated (Constantinidis et al., 2001b), their activity is sculpted by lateral inhibition (Rao et al., 2000; Constantinidis and Goldman-Rakic, 2002), and excitatory persistent activity depends on NMDA receptors (Wang et al., 2013). Furthermore, the model predicts that random drift in the “bump” of activity encoding the memorandum is the cause of memory errors (Compte et al., 2000), which is supported by the fact that directional biases (e.g., clockwise or counterclockwise) in the identity of the memorandum decoded from macaque prefrontal cortex correlate with memory errors in behavior (Wimmer et al., 2014).

Differences in cortical architecture between prefrontal cortex and visual cortex have led some to argue that V1 is not well-suited to support persistent activity (Wang, 2001, 2020; Riley and Constantinidis, 2016; Leavitt et al., 2017). While V1 shares some architectural similarities with prefrontal cortex that inspired the canonical model, there are several properties that likely constrain its ability to support persistent activity. NMDA receptors are less densely expressed in V1 compared to prefrontal cortex (Wang et al., 2008). Pyramidal neurons in prefrontal cortex have more extensive horizontal connectivity in Layers II and III (Kritzer and Goldman-Rakic, 1995) and are twice as likely to form reciprocal connections (Wang et al., 2006) compared to those in visual cortex. Changes in the relative proportions of different types of inhibitory interneurons from visual to frontal cortex may also lead to greater ability of prefrontal cortex to support persistent activity (Torres-Gomez et al., 2020). Functionally, the timescale of intrinsic neural activity slows from visual to frontal cortex (Murray et al., 2014). Computational models suggest that the faster timescales in visual cortex would limit persistent activity (Wang, 2001; Mejías and Wang, 2022), while the architecture of prefrontal cortex particularly lends itself to the



slow recurrent dynamics necessary to support working memory. While these gradients in the putative ability to support persistent activity match the gradient of persistent activity we find in fMRI, the counter-gradient of stronger decoding of working memory content in visual cortex indicates that information likely persists there in some form.

***Neural architecture: top-down feedback and differentiating storage from maintenance***

Working memory activity observed in sensory cortex likely reflects top-down feedback. Both laminar recording in macaque V1 (Van Kerkoerle et al., 2017) and fMRI measurements of human V1 (Lawrence et al., 2018) demonstrate that persistent activity is largely confined to superficial and deep layers of cortex, where top-down inputs terminate. Local-field potentials in sensory areas, which are thought to largely reflect synaptic inputs to a region, can be elevated and selective for the contents of working memory even in the absence of evidence for persistent activity (Mendoza-Halliday et al., 2014; Bahmani et al., 2018). Furthermore, TMS and lesions to human lateral prefrontal cortex disrupt response selectivity in visual cortex in both perception (Miller et al., 2011) and working memory (Lee and D'Esposito, 2012), and activity levels in prefrontal and parietal cortices predict the fidelity of decoded information in visual cortex during both working (Master et al., 2023) and long-term memory (St-Laurent et al., 2015).

While the role of top-down feedback in sensory cortices during working memory is widely acknowledged, there is substantial disagreement over how to interpret these empirical observations. Some view feedback in sensory cortices as evidence that prefrontal or parietal areas must be the site of working memory storage, whereas working memory signals in sensory regions may or may not have a functional role separate from storage (Riley and Constantinidis, 2016; Leavitt et al., 2017; Xu, 2017, 2020). The logic here is that sensory cortex inherits its working memory information from association cortex; therefore, sensory cortex is unnecessary for working memory storage. However, this view fails to appreciate the distinction between

*storage* and *maintenance* that is made by proponents of sensory recruitment (Curtis and D'Esposito, 2003). As noted above, feedback from prefrontal cortex may be particularly important for the *maintenance* of working memory by protecting it from noise or distraction, but this does not mean the prefrontal cortex is the primary site in which working memory representations are stored. Indeed, computational modeling suggests that nonspecific modulation of neural gain, such as that caused by neuromodulatory input or top-down attention, (Niyogi and Wong-Lin, 2013) or reciprocal connectivity between regions (Mejías and Wang, 2022) can enable persistent activity in neural circuits that cannot independently support persistent activity. This suggests that, in principle, prefrontal cortex need not store working memory representations to enable working memory maintenance.

Taken together, arguments from neural architecture are most convincing when implemented in computational models (Teng and Postle, 2021), as demonstrated by the canonical model of working memory. Such models allow bridging between levels of analysis, from neurobiology to circuits to behavior. They force researchers to make their theories concrete, and ideally they generate specific, testable hypotheses. They also reduce the need for intuition, which can go astray when attempting to reason about complex dynamical systems such as the brain.

### ***Integrating working memory and sensory processing for efficiency***

Some argue that utilizing sensory cortex for working memory is sensible from the standpoint of efficient use of neural resources. For example, there could be costs (e.g., metabolic) to having separate sites of perception and working memory with similar coding properties (Postle, 2006; Sreenivasan et al., 2014). Situating working memory in sensory cortex also may allow for rapid comparisons between the contents of working memory and perception, potentially facilitating behavioral performance (Pasternak and Greenlee, 2005; Sugase-Miyamoto et al., 2008; Woloszyn and Sheinberg, 2009; Rademaker et al., 2019). Indeed, cells in IT cortex show

selective response patterns to stimuli that match the contents of working memory (Miller et al., 1996a; Sugase-Miyamoto et al., 2008; Woloszyn and Sheinberg, 2009). Note, however, that this view does not require a commitment to sensory cortex being the site of working memory storage (Xu, 2020), and such comparison processes could be accomplished through alternative mnemonic mechanisms to persistent activity, such as those involving short-term synaptic plasticity or modulation of subthreshold synaptic currents (Sugase-Miyamoto et al., 2008; Rademaker et al., 2019; Comeaux et al., 2023).

### ***Neural coding and the integration or segregation of function***

Storage of working memory in sensory cortex may also improve the fidelity of working memory representations (Sreenivasan et al., 2014; Christophel et al., 2017). For example, it is not clear that the neural tuning in prefrontal cortex is sufficient to support fine-grained memoranda (Zaksas and Pasternak, 2006). However, the relationship between the tuning precision of individual neurons and the fidelity of the representation is likely to be complex. For example, theoretical analysis suggests that for a commonly used linear readout of population activity, the amount of information that can be extracted is a nonlinear function of the tuning width and signal-to-noise characteristics of the individual units (Seung and Sompolinsky, 1993). Therefore, differences in tuning precision across regions may only serve as a rough guide, with detailed computational analyses necessary to determine the relative coding fidelity of each region. Furthermore, population tuning in visual cortex is wider in working memory than in perception, suggesting that working memory does not take full advantage of the coding capacity of visual cortex (Li and Curtis, 2023).

A separate set of arguments make inferences about a region's role in working memory storage based on its other functions. Some argue that because sensory cortex serves perception, it could not simultaneously maintain working memory without interference (Xu, 2017). On the

other hand, because prefrontal cortex neurons generally show selectivity for various task features, including abstract rules, rather than strictly memoranda derived from perception, one could argue that its role is unlikely to be primarily for storage (Sreenivasan and D'Esposito, 2019). These arguments presuppose that a neural area can only process one thing at a time. However, there is increasing evidence that neural areas, including sensory areas, can multiplex information, either in different subspaces of neural activity (Semedo et al., 2019, 2022; Hajnal et al., 2023) or across different lamina or frequency bands (Bastos et al., 2018, 2020). Therefore, arguments that merely invoke the fact that a region primarily does one function do not bear on whether it could also (simultaneously) perform another function. The converse is also true: frontal and parietal cortices show sensory activity during perception, but we would not discount the role of sensory cortex in perception; neither should the reverse position be taken for the role of sensory cortex in working memory (Scimeca et al., 2018).

### ***Separating working memory and sensory processing for robustness: the case study of distraction***

More compelling arguments for the potential need to segregate perception and working memory come from studies of the effect of distracting stimuli on signatures of working memory in different regions of the brain. The logic here is that if a region is involved in working memory storage, it needs to have some level of robustness in the face of intervening inputs, or else the memorandum would be corrupted or lost. Indeed, studies of distractor resistance have been some of the most challenging to the sensory recruitment hypothesis. Intervening distractors disrupt the selectivity of delay-period activity in macaque IT (Miller et al., 1993b, 1996b) and PPC (di Pellegrino and Wise, 1993; Constantinidis and Steinmetz, 1996; Qi et al., 2010; Suzuki and Gottlieb, 2013), while prefrontal cortex delay activity appears resistant to distraction. These are some of the most widely cited results in support of the idea that prefrontal cortex, rather than sensory cortex, is critical for working memory storage. Evidence from human fMRI has been

mixed, with some reporting that memoranda in visual cortex are disrupted by intervening irrelevant distractors and others reporting decoding of both the memory target and the distractor (Bettencourt and Xu, 2016; Lorenc et al., 2018; Rademaker et al., 2019; Hallenbeck et al., 2021a; Iamshchinina et al., 2021). A number of factors could contribute to these mixed results, including differences in statistical power and decoding methods used (Iamshchinina et al., 2021), and differences in the effects of distractors on behavior between studies. Nevertheless, because memoranda cannot always be decoded from visual cortex during distraction, its role in working memory storage has been questioned (Xu, 2017, 2020).

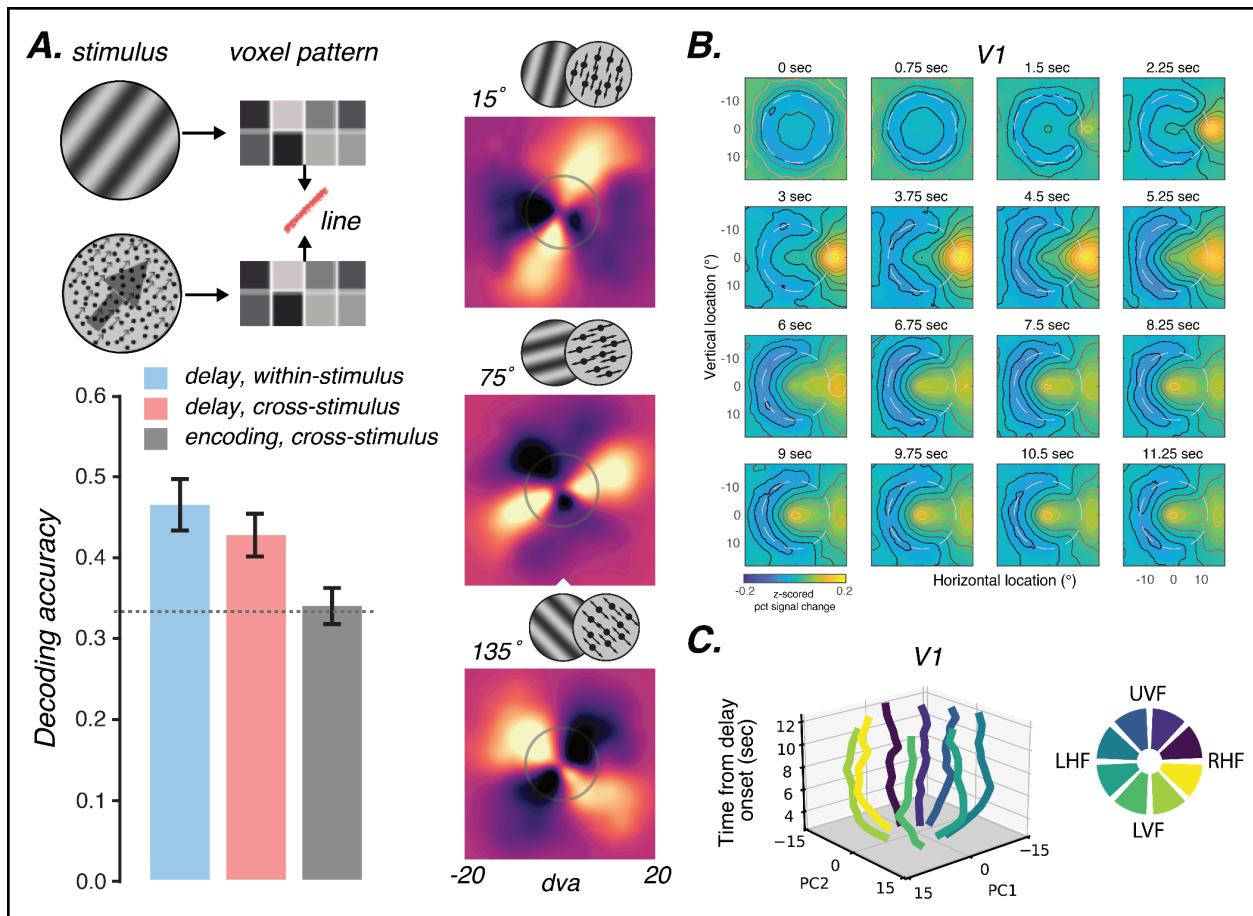
However, the requirement for robustness cannot ignore the fact that distracting stimuli bias working memory behavior (Magnussen et al., 1991; Magnussen and Greenlee, 1992; Smyth and Scholey, 1994; Smyth, 1996; Rademaker et al., 2015), and conversely that working memory contents can impact both perception and visuomotor selection (Gayet et al., 2013; Hollingworth et al., 2013), suggesting tight links between working memory and perceptual processing. Importantly, in fMRI studies in which distractors bias behavior, reconstructions from visual cortex show similar bias (Lorenc et al., 2018; Hallenbeck et al., 2021a), with a recent study demonstrating that biases in memory errors induced by distractors are predicted by trial-by-trial errors in the working memory representations in early visual cortex, but not association cortex (Hallenbeck et al., 2021a). Together, these findings provide strong evidence for a role of visual cortex in working memory storage.

As the case of distraction demonstrates, strong evidence for the role of a brain region in working memory storage is provided if representations reconstructed from this region predict behavioral performance. Such evidence is made even stronger if neural activity predicts behavior on a trial-by-trial basis within-subject, and stronger still if a region *differentially* predicts behavior relative to other areas. While such brain-behavior correlations could always in principle be

driven by other areas, finding differential relationships supplies important evidence in favor of region-specific hypotheses (Xu, 2020).

### **The format and dynamics of working memory**

We have already argued against the idea that perceptual and working memory representations necessarily interfere with each other, given the ability of brain regions to multiplex information. However, ideas of this nature have an underlying assumption that bears examination: namely, that working memory representations have a similar format as perceptual representations (Xu, 2017). While some early efforts to find evidence for sensory recruitment made this explicit assumption (Serences et al., 2009), the idea that the contents of working memory are different from perception has a long history. For example, when tasked to remember letters from a known alphabet, we do not store the visual form of the letters, but rather transform the letters into phonological or sound-based codes (Conrad and Hull, 1964). Such codes effectively compress the information that needs to be recalled, allowing us to recall more letters. Thus, transformations of working memory may serve two purposes: protecting it from perceptual interference, and compressing it into a format that maximizes working memory capacity and supports behavioral goals.



**Figure 5.** Format of working memory representations in visual cortex. **A.** We predicted that distinct visual stimuli—the orientation of a grating or the direction of dot motion—would be re-coded into a line-like format during working memory maintenance (left). Indeed, classifiers trained to decode one stimulus (e.g., orientation) could be used to predict the other stimulus (e.g., motion direction) based on the patterns of activity during the memory delay (red bar), but not during perception (gray bar). For comparison, the average within-stimulus (e.g., test and train on gratings) decoding results are shown (blue bar). We visualized our hypothesis that a line-like representation was being stored by projecting voxel activity in V1 from cortical space to visual space using the receptive field parameters of the voxels, estimated in a separate pRF mapping experiment (right). Adapted from (Kwak and Curtis, 2022). **B.** During an MGS task, participants maintained the location of a target in working memory over a 12 second delay period, then generated a saccade to the stored location. Each TR of BOLD data from the delay is projected onto the screen coordinates of visual space using the estimated receptive field parameters from pRF mapping. Data from all trials are rotated such that the targets are all to the right. Note how over time activity builds at the target location, then forms a line between fixation and the target location, perhaps indicating a transformation into the trajectory of the memory-guided saccade. **C.** BOLD activity in V1 during each time point during the delay period of **B** was visualized in a subspace using principal components analysis (PCA; left). PCA is a technique that can reduce complex, high-dimensional data (e.g., hundreds of voxels) into a small number of dimensions that best explain the variability in the data. PC1 and PC2, the first two principal components (x-axis and y-axis) captured horizontal and vertical visual space, respectively. The z-axis represents time from delay onset, and the color wheel (right) indicates the location, in the visual field, represented by each of the eight colored lines. Despite the dynamics visualized in **B**, this subspace was stable, as indicated by the relatively constant position of each location over time. **B** and **C** adapted from (Li and Curtis, 2023)

In visual working memory, decoders trained on activity patterns from perception of a stimulus are often poor predictors of patterns during working memory, especially when compared with training on working memory patterns themselves (Harrison and Tong, 2009; Serences et al., 2009; Albers et al., 2013; Spaak et al., 2017; Rademaker et al., 2019; Hallenbeck et al., 2021a; lamshchinina et al., 2021). These results suggest that perceptual representations have been reformatted during working memory but have the disadvantage of drawing conclusions based on null results, which could arise for many reasons apart from reformatting. Recently developed techniques allow projecting voxel activity from cortex into spatial visualizations in the same screen coordinates as the presented stimuli (Kok and de Lange, 2014; Favila et al., 2022; Kwak and Curtis, 2022; Yoo et al., 2022; Zhou et al., 2022; Li and Curtis, 2023; Woodry et al., 2024). In one study using this technique (Kwak and Curtis, 2022), participants were required to make orientation judgments for oriented grating stimuli and direction judgments for moving dots stimuli. In both cases, reconstructions from neural activity in V1 and other areas demonstrated that the task relevant features of the stimuli were extracted and re-coded into a shared mnemonic format that took on the form of an abstract, line-like pattern aligned with the grating's orientation or the direction of dot motion. Importantly, this reformatting did not occur for perception (Figure 5A). Other studies have now found similar results using slightly different stimuli and techniques (Li and Curtis, 2023; Duan and Curtis, 2024) (Figure 5B). A related behavioral and modeling study that carefully controlled for the sensory stimuli, actions, rules, and predictability of the response needed during a working memory task found that recurrent neural networks trained on the task did not merely store the sensory and/or motor action required to complete the behavior (Ehrlich and Murray, 2022). Rather, they maintained a "contingency representation," providing a mapping from possible future task events to the required actions, and patterns of response time from human subjects were consistent with the use of contingency representations. Together, these studies provide strong evidence that



working memory representations are efficient, compressed abstractions optimized to support behavior, putting more of the “working” into working memory.

More generally, neural activity during working memory seems to be far more dynamic than expected based on the canonical model of working memory. The neurons encoding the memorandum and their population dynamics appear to change over the course of the working memory delay, though the population appears to maintain a “stable subspace” allowing a consistent readout of the memorandum (Murray et al., 2017; Spaak et al., 2017; Li and Curtis, 2023) (Figure 5C). Modeling suggests that such dynamics help to optimally store information in working memory, where optimality is defined as maximizing the ability to decode the contents of working memory at the end of the delay period (Stroud et al., 2023). Interestingly, in the model such “optimal loading” further minimized the neural activity needed to maintain the memorandum, and the activity subspace at perceptual input was orthogonal to that during the late delay. Signatures of optimal loading were confirmed in data from macaque prefrontal cortex (Stroud et al., 2023) and orthogonality between perceptual and working memory representations has also been observed in mouse auditory cortex (Libby and Buschman, 2021). These results indicate that even in the absence of reformatting into a more abstract representation, neural dynamics may transform the working memory representation to protect it from perceptual interference.

### **Integrative accounts of working memory storage**

While we have focused on contrasting prefrontal cortex-focused with sensory-focused accounts of working memory storage, there is a growing consensus that working memory is not a property of any one brain region but rather is supported by interactions among multiple regions (Postle, 2006; D’Esposito, 2007; Fuster, 2015; Serences, 2016; Christophel et al., 2017; Iamshchinina et al., 2021; Lorenc and Sreenivasan, 2021; Mejías and Wang, 2022; Comeaux et al., 2023).

Distributed working memory may have multiple benefits over working memory localized to a particular region. If the same or similar working memory contents are spread out over multiple regions, it could afford greater robustness to noise and distraction (Hallenbeck et al., 2021a; lamshchinina et al., 2021), a proposal supported by computational modeling (Mejías and Wang, 2022). Indeed, disruptions to regions with persistent activity can bidirectionally affect the content of other regions (e.g., (Fuster et al., 1985; Chafee and Goldman-Rakic, 2000), and there is suggestive evidence that interactions between regions may be a stronger determinant of working memory performance than activity within individual regions (Rezayat et al., 2022).

Besides redundancy, which implies that working memory representations in different regions are merely copies, distributed working memory can take advantage of the different strengths afforded by each region via its coding properties, depending on the demands of the task (Serences, 2016; Christophel et al., 2017; Lorenc and Sreenivasan, 2021). Factors affecting the loci of working memory storage might include distractor strength and relevance (Bettencourt and Xu, 2016; Rademaker et al., 2019); whether the memorandum is abstract or more sensory in nature, as well as its familiarity (Christophel et al., 2017; Rademaker et al., 2019); whether the memory item is currently in the focus of attention or not (Christophel et al., 2018b), whether the representation reflects the contents of working memory or its context (Teng and Postle, 2021, 2024), and whether the action to be taken at the end of the delay is known or not (Curtis et al., 2004; Curtis and D'Esposito, 2006; Henderson et al., 2022). Supporting such flexibility, working memory decoding from visual cortex is weaker when the correct motor response is known (Henderson et al., 2022). Similarly, working memory is stronger in visual cortex when fine-grained maintenance of the stimulus is required, but stronger in prefrontal cortex during rule-based categorization of the stimulus (Shao et al., 2023). Furthermore, while most studies find better distractor resistance in prefrontal cortex than other areas for spatial memoranda, PPC may be more robust to distractors for numerical stimuli (Jacob and Nieder, 2014). In sum,

the consensus emerging from these integrative perspectives shifts the emphasis of working memory research from trying to identify *the* brain region(s) responsible for working memory storage to asking how different regions and their interactions contribute to working memory, and how those contributions may shift based on task context (Lorenc and Sreenivasan, 2021).

## **Working memory control processes**

Up to this point, we have primarily focused on the storage, or memory, aspects of working memory. But what about the “working” part, those processes that manipulate and update the contents of working memory? While we know quite a bit about how working memory is stored in the brain, understanding the control of working memory is far more challenging. Part of the challenge stems from the fact that it can be very difficult to differentiate control processes from storage on the basis of commonly used measures of persistent activity, as discussed above (Curtis and D’Esposito, 2003). Moreover, multivariate encoding/decoding methods are best suited to measure information about stimulus features (e.g., orientation, position, color, etc.) stored in working memory. It remains challenging to decode aspects of control processes, as researchers have less experimental control over their implementation. While there are many control processes associated with working memory, here we focus on two of the more well-studied classes of processes: those that control access to working memory, and those that prioritize the contents of working memory.

### **Access to working memory**

Access to working memory must be controlled if it is to function properly. A working memory system that allows too many items to enter would be more susceptible to distraction and interference, leading to poor performance. Conversely, a working memory system that does not admit goal-relevant information when necessary would fail to form the necessary bridge between previous perception and later action. Indeed, damage to prefrontal cortex can lead to

distractibility, such that behavior that is more controlled by external stimuli, and perseveration, indicating a failure to update behavior on the basis of new information (Goldman-Rakic, 2011), both deficits suggesting the importance of prefrontal cortex for the control of working memory.

One mechanism for access to working memory appeals to the concept of a filter, a function often attributed to the prefrontal cortex (Shimamura, 2000; Chrysikou et al., 2014). However, there are debates about how this filtering operation might be implemented. Some accounts suggest that filtering is an inhibitory function, suppressing irrelevant information from entering working memory. For example, one proposal for why working memory declines with age is that the aging brain has weaker inhibitory control, leading to a greater amount of information entering working memory in older than in younger adults (Hasher and Zacks, 1988; Lustig et al., 2007). An alternative view is that filtering is accomplished not by suppressing irrelevant information or representations, but by top-down biasing signals from the prefrontal cortex that enhance relevant representations, which could then suppress irrelevant representations via competitive local inhibition (Munakata et al., 2011; Liesefeld et al., 2020).

Another mechanism for access to working memory appeals to the concept of a gate. When the gate is open, working memory can be updated with new information; when it is closed, the contents of working memory are protected from distraction. Both computational modeling and neural evidence suggest that gating may be implemented by connectivity between prefrontal cortex and the striatum, a subcortical region in the basal ganglia, where dopamine signals in the striatum control or gate access to prefrontal working memory (O'Reilly and Frank, 2006; Badre and Frank, 2012; D'Ardenne et al., 2012; Chatham et al., 2014; Furman et al., 2020).

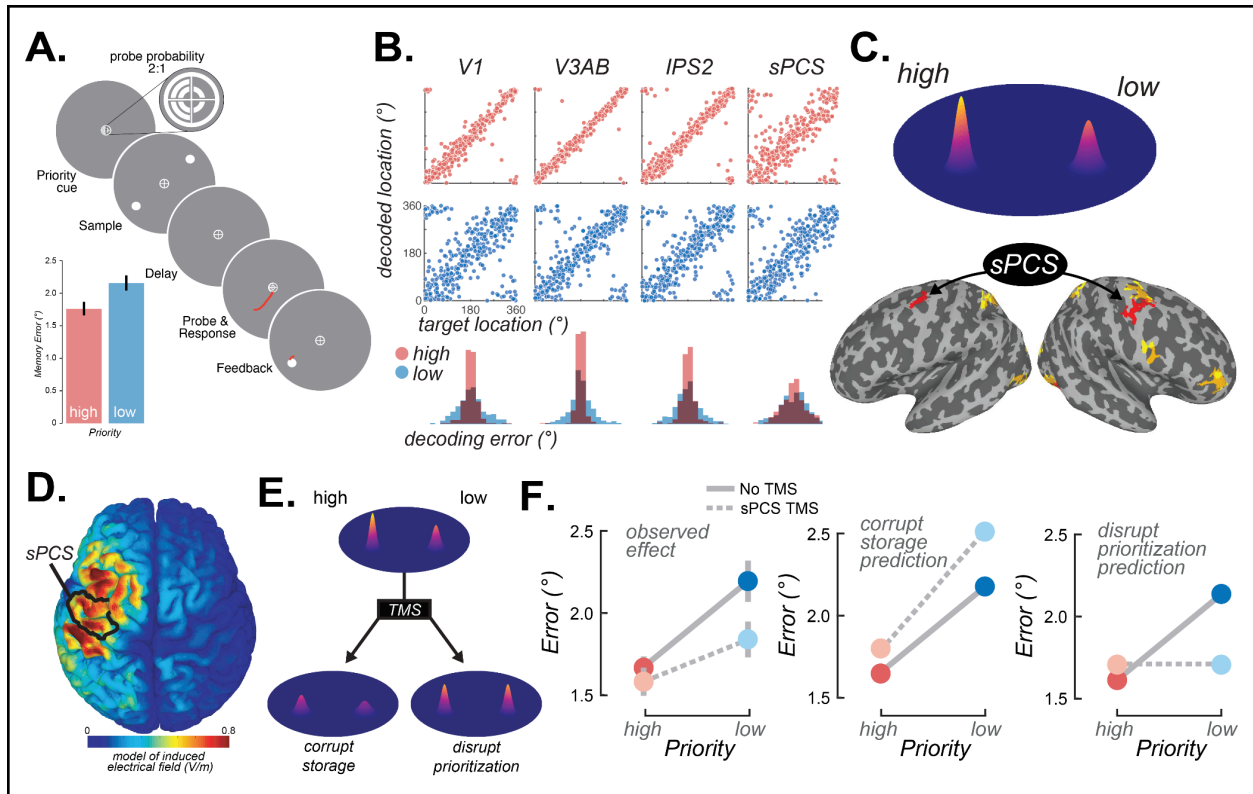
### **Prioritization in working memory**

The fact that working memory is capacity limited presents an optimization problem: if multiple items need to be remembered, how should resources be allocated in order to maximize our

memory? In particular, when some items are more important for recall than others, it would make sense for our working memory system to have the capability to store more relevant items with higher precision than less relevant items. This prioritization function thus provides more subtle control over the current contents of working memory, complementing the access functions described above.

Prioritization, as a control process, is amenable to study using the methods for understanding storage described above and is an important example of how prefrontal cortex and visual cortex interact in support of working memory. Prioritization has been tested using paradigms that cue participants as to which of multiple items is most likely to be the target of recall. In these conditions, people prioritize items in memory according to their behavioral relevance, resulting in better memory for prioritized items but at a cost of worse memory for other items (Zhang and Luck, 2008; Klyszejko et al., 2014; Emrich et al., 2017; Yoo et al., 2018). fMRI findings suggest that the neural basis of prioritization in spatial working memory relies on interactions between frontal and visual cortex. First, neural populations encoding high-priority items have higher delay-period activity than those encoding low-priority items, suggesting that prioritization is implemented at the neural level by changing the relative gain of populations responsible for storing the memoranda (Yoo et al., 2022). Importantly, memorandum-specific gain changes were not found in parietal or frontal cortex. Second, trial-by-trial amplitudes of delay-period activity in sPCS of the frontal cortex were found to predict prioritization in visual cortex (Li et al., 2024) (Figure 6A–C). This relationship was unique to sPCS, suggesting that it controls prioritization. Providing causal support for this hypothesis, disrupting activity in sPCS using TMS disrupted the behavioral prioritization of working memory (Hallenbeck et al., 2024) (Figure 6D–F). Together, these results suggest a clear division of labor, in which prefrontal cortex is involved in the control of working memory, while visual cortex plays a role in working memory storage. This line of work serves as a critical example that understanding the control and

storage subcomponents of working memory will require new analytic techniques that allow researchers to measure trial-wise interactions between brain areas.



**Figure 6.** Prioritizing working memory resources. **A.** Both an fMRI and TMS study used a two-item memory-guided saccade task. A pre-cue prior to the target stimuli indicated which of the two items was more likely to be tested, where the high-priority target was tested twice as often (2:1 probe probability). Behaviorally, high-priority items had smaller memory errors, indicating participants prioritized the items in memory. Adapted from (Hallenbeck et al., 2024; Li et al., 2024). **B.** Decoded representations of high- and low-priority items for an example participant in striate, extrastriate, parietal, and frontal visual field maps. Note the better precision of the high-priority representations. **C.** Top: Cartoon model of populations encoding a high- and low-priority item. The high-priority item has a higher gain (activity level) than the low, leading to higher precision. Bottom: Differences in memory precision were estimated from the neural data in **B**. Trial-by-trial differences in precision covaried with the amplitude of persistent delay-period activity in bilateral sPCS (red maps). The sum of the precisions of both items, perhaps related to the total amount of allocated resource, correlated with the delay-period amplitude of a broader network in frontal and parietal cortex (yellow). **D.** Modeled electrical field evoked by TMS to the sPCS (left). **E.** Hypothesized effects of TMS on low- and high-priority items in working memory. Without TMS (top), the neural populations representing the high-priority item could have a higher gain than those for the low-priority item, as in **C**. TMS could either corrupt the storage of the memoranda (bottom left) or it could disrupt the prioritization, such that the items are maintained with closer to equal gain (bottom right). **F.** Using a modified variable-precision model of working memory, the authors tested which of the two hypothesized effects of TMS best accounted for memory errors. Left: Without TMS, errors were smaller for high compared to low-priority items in the observed data. TMS reduced this difference, mainly by reducing the errors for low-priority items. Middle: Modeled effects on memory error if storage were corrupted by TMS. Right: Modeled effects on memory error if prioritization were disrupted by TMS. Clearly, the observed data most resembled the modeled disruption of prioritization. A, B, and C adapted from (Li et al., 2024) and D, E and F adapted from (Hallenbeck et al., 2024).

## Conclusion

In this review, we have outlined key behavioral signatures of working memory and what is known about their neural substrates across three fundamental aspects: capacity, storage, and control. As we have highlighted, rather than being subserved by any one brain area, working memory is made possible by complex interactions across many parts of the brain, befitting its crucial role in higher-order cognition. As is often the case, the development of our understanding of working memory in cognitive neuroscience from more local to more distributed perspectives has paralleled the development of neuroimaging techniques over the decades. The first generation of PET and fMRI studies used gross amplitude measures to identify the network of brain areas activated during working memory. The second generation of fMRI studies used analyses of multivoxel patterns to decode and identify where working memory information is stored in the brain. The third generation of fMRI studies, which has just begun, combines decoding and amplitude measures and their interactions to test theories about how the control processes of working memory are implemented in the brain.

Utilizing this new generation of techniques will facilitate solving important unanswered challenges for understanding working memory:

- 1) We must better delineate the role of specific regions and their interactions in working memory, which requires careful consideration of their potential contributions to both storage and control.
- 2) With respect to storage, increasing our ability to reconstruct the contents of working memory from neural data can further our understanding of why these contents are stored where they are, how they guide behavior, and the transformations they undergo.

3) Control processes in working memory have been comparatively harder to study. Given that operations on neural representations occur via interactions between brain areas (Miller and Cohen, 2001; DiCarlo et al., 2012), an increasing focus on interactions within and between regions could provide new avenues for identifying the neural substrates and computations by which control processes maintain and manipulate the contents of working memory.



## References

- Albers AM, Kok P, Toni I, Dijkerman HC, de Lange FP (2013) Shared representations for working memory and mental imagery in early visual cortex. *Curr Biol* 23:1427–1431. PMID: 23871239.
- Alloway TP (2006) How does working memory work in the classroom? *Educational Research and Reviews* 1:134–139. .
- Alloway TP, Gathercole SE, Kirkwood H, Elliott J (2009) The Cognitive and Behavioral Characteristics of Children With Low Working Memory. *Child Dev* 80:606–621. PMID: 19467014.
- Averbeck BB, Chafee MV, Crowe DA, Georgopoulos AP (2003) Neural activity in prefrontal cortex during copying geometrical shapes. I. Single cells encode shape, sequence, and metric parameters. *Exp Brain Res* 150:127–141. PMID: 12669170.
- Awh E, Barton B, Vogel EK (2007) Visual Working Memory Represents a Fixed Number of Items Regardless of Complexity. *Psychol Sci* 18:622–628. .
- Baddeley A (2012) Working memory: theories, models, and controversies. *Annu Rev Psychol* 63:1–29. PMID: 21961947.
- Baddeley AD, Hitch G (1974) Working Memory. In: *Psychology of Learning and Motivation* (Bower GH, ed), pp 47–89. Academic Press.
- Baddeley AD, Logie RH (1999) Working memory: The multiple-component model. In: *Models of working memory: Mechanisms of active maintenance and executive control* (pp (Miyake A, ed), pp 28–61. New York, NY, US: Cambridge University Press, xx.
- Badre D, Frank MJ (2012) Mechanisms of hierarchical reinforcement learning in cortico-striatal circuits 2: Evidence from fMRI. *Cereb Cortex* 22:527–536. PMID: 21693491.
- Bahmani Z, Daliri MR, Merrikhi Y, Clark K, Noudoost B (2018) Working Memory Enhances Cortical Representations via Spatially Specific Coordination of Spike Times. *Neuron* 97:967–979.e6. PMID: 29398360.
- Barash S, Bracewell RM, Fogassi L, Gnadt JW, Andersen RA (1991) Saccade-related activity in the lateral intraparietal area. II. Spatial properties. *J Neurophysiol* 66:1109–1124. PMID: 1753277.
- Barch DM, Braver TS, Nystrom LE, Forman SD, Noll DC, Cohen JD (1997) Dissociating working memory from task difficulty in human prefrontal cortex. *Neuropsychologia* 35:1373–1380. PMID: 9347483.
- Bastos AM, Loonis R, Kornblith S, Lundqvist M, Miller EK (2018) Laminar recordings in frontal cortex suggest distinct layers for maintenance and control of working memory. *Proc Natl Acad Sci U S A* 115:1117–1122. PMID: 29339471.
- Bastos AM, Lundqvist M, Waite AS, Kopell N, Miller EK (2020) Layer and rhythm specificity for predictive routing. *Proc Natl Acad Sci U S A* 117:31459–31469. PMID: 33229572.

- Bays PM, Husain M (2008) Dynamic shifts of limited working memory resources in human vision. *Science* 321:851–854. PMC2532743.
- Bettencourt KC, Xu Y (2016) Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nat Neurosci* 19:150–157. .
- Bisley JW, Zaksas D, Droll JA, Pasternak T (2004) Activity of Neurons in Cortical Area MT during A Memory for Motion Task. *J Neurophysiol* 91:286–300. PMID: 14523065.
- Boran E, Hilfiker P, Stieglitz L, Sarnthein J, Klaver P (2022) Persistent neuronal firing in the medial temporal lobe supports performance and workload of visual working memory in humans. *Neuroimage* 254:119123. PMID: 35321857.
- Bor D, Duncan J, Wiseman RJ, Owen AM (2003) Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron* 37:361–367. PMID: 12546829.
- Brissenden JA, Tobyne SM, Halko MA, Somers DC (2021) Stimulus-Specific Visual Working Memory Representations in Human Cerebellar Lobule VIIb/VIIIa. *J Neurosci* 41:1033–1045. PMC7880273.
- Brouwer GJ, Heeger DJ (2009) Decoding and reconstructing color from responses in human visual cortex. *J Neurosci* 29:13992–14003. PMC2799419.
- Camos V (2017) Chapter Five - Domain-Specific Versus Domain-General Maintenance in Working Memory: Reconciliation Within the Time-Based Resource Sharing Model. In: *Psychology of Learning and Motivation* (Ross BH, ed), pp 135–171. Academic Press.
- Chafee MV, Goldman-Rakic PS (1998) Matching patterns of activity in primate prefrontal area 8a and parietal area 7ip neurons during a spatial working memory task. *J Neurophysiol* 79:2919–2940. PMID: 9636098.
- Chafee MV, Goldman-Rakic PS (2000) Inactivation of parietal and prefrontal cortex reveals interdependence of neural activity during memory-guided saccades. *J Neurophysiol* 83:1550–1566. PMID: 10712479.
- Chatham CH, Frank MJ, Badre D (2014) Corticostriatal output gating during selection from working memory. *Neuron* 81:930–942. PMC3955887.
- Chelazzi L, Duncan J, Miller EK, Desimone R (1998) Responses of neurons in inferior temporal cortex during memory-guided visual search. *J Neurophysiol* 80:2918–2940. PMID: 9862896.
- Christophel TB, Allefeld C, Endisch C, Haynes J-D (2018a) View-Independent Working Memory Representations of Artificial Shapes in Prefrontal and Posterior Regions of the Human Brain. *Cereb Cortex* 28:2146–2161. PMID: 28505235.
- Christophel TB, Iamshchinina P, Yan C, Allefeld C, Haynes JD (2018b) Cortical specialization for attended versus unattended working memory. *Nature Neuroscience* 2018 21:4 21:494–496. PMID: 29507410.
- Christophel TB, Klink PC, Spitzer B, Roelfsema PR, Haynes J-D (2017) The Distributed Nature of Working Memory. *Trends Cogn Sci* 21:111–124. .

- Chrysiou EG, Weber MJ, Thompson-Schill SL (2014) A matched filter hypothesis for cognitive control. *Neuropsychologia* 62:341–355. PMID: 24200920.
- Cohen JD, Perlstein WM, Braver TS, Nystrom LE, Noll DC, Jonides J, Smith EE (1997) Temporal dynamics of brain activation during a working memory task. *Nature* 386:604–608. PMID: 9121583.
- Comeaux P, Clark K, Noudoost B (2023) A recruitment through coherence theory of working memory. *Prog Neurobiol* 228:102491. .
- Compte A, Brunel N, Goldman-Rakic PS, Wang XJ (2000) Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb Cortex* 10:910–923. PMID: 10982751.
- Conrad R, Hull AJ (1964) INFORMATION, ACOUSTIC CONFUSION AND MEMORY SPAN. *Br J Psychol* 55:429–432. PMID: 14237884.
- Constantinidis C, Franowicz MN, Goldman-Rakic PS (2001a) The sensory nature of mnemonic representation in the primate prefrontal cortex. *Nature Neuroscience* 2001 4:3 4:311–316. PMID: 11224549.
- Constantinidis C, Franowicz MN, Goldman-Rakic PS (2001b) Coding Specificity in Cortical Microcircuits: A Multiple-Electrode Analysis of Primate Prefrontal Cortex. *Journal of Neuroscience* 21:3646–3655. PMID: 11331394.
- Constantinidis C, Goldman-Rakic PS (2002) Correlated discharges among putative pyramidal neurons and interneurons in the primate prefrontal cortex. *J Neurophysiol* 88:3487–3497. PMID: 12466463.
- Constantinidis C, Steinmetz MA (1996) Neuronal activity in posterior parietal area 7a during the delay periods of a spatial memory task. *J Neurophysiol* 76:1352–1355. PMID: 8871242.
- Conway ARA, Kane MJ, Engle RW (2003) Working memory capacity and its relation to general intelligence. *Trends Cogn Sci* 7:547–552. PMID: 14643371.
- Courtney SM, Petit L, Maisog JM, Ungerleider LG, Haxby JV (1998) An area specialized for spatial working memory in human frontal cortex. *Science* 279:1347–1351. PMID: 9478894.
- Cowan N (2001) The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behav Brain Sci* 24:87–114. PMID: 11515286.
- Cowan N (2017) The many faces of working memory and short-term storage. *Psychon Bull Rev* 24:1158–1170. PMID: 27896630.
- Curtis CE, D’Esposito M (2003) Persistent activity in the prefrontal cortex during working memory. *Trends Cogn Sci* 7:415–423. .
- Curtis CE, D’Esposito M (2004) The effects of prefrontal lesions on working memory performance and theory. *Cogn Affect Behav Neurosci* 4:528–539. PMID: 15849895.
- Curtis CE, D’Esposito M (2006) Selection and maintenance of saccade goals in the human frontal eye fields. *J Neurophysiol* 95:3923–3927. PMID: 16467423.

- Curtis CE, Rao VY, D'Esposito M (2004) Maintenance of spatial and motor codes during oculomotor delayed response tasks. *J Neurosci* 24:3944–3952. PMC6729424.
- Curtis CE, Sprague TC (2021) Persistent Activity During Working Memory From Front to Back. *Front Neural Circuits* 15:696060. .
- Czoschke S, Fischer C, Bahador T, Bledowski C, Kaiser J (2021) Decoding Concurrent Representations of Pitch and Location in Auditory Working Memory. *J Neurosci* 41:4658–4666. PMC8260242.
- D'Ardenne K, Eshel N, Luka J, Lenartowicz A, Nystrom LE, Cohen JD (2012) Role of prefrontal cortex and the midbrain dopamine system in working memory updating. *Proc Natl Acad Sci U S A* 109:19900–19909. PMC3523834.
- Dash S, Yan X, Wang H, Crawford JD (2015) Continuous updating of visuospatial memory in superior colliculus during slow eye movements. *Curr Biol* 25:267–274. PMID: 25601549.
- D'Esposito M (2007) From cognitive to neural models of working memory. *Philos Trans R Soc Lond B Biol Sci* 362:761–772. PMID: 17400538.
- D'Esposito M, Postle BR (2015) The Cognitive Neuroscience of Working Memory. <https://doi.org/10.1146/annurev-psych-010814-015031> 66:115–142. PMID: 25251486.
- D'Esposito M, Postle BR, Ballard D, Lease J (1999) Maintenance versus manipulation of information held in working memory: an event-related fMRI study. *Brain Cogn* 41:66–86. PMID: 10536086.
- Deutsch P, Czoschke S, Fischer C, Kaiser J, Bledowski C (2023) Decoding of Working Memory Contents in Auditory Cortex Is Not Distractor-Resistant. *Journal of Neuroscience* 43:3284–3293. PMID: 36944488.
- Diamond A (2013) Executive functions. *Annu Rev Psychol* 64:135–168. PMC4084861.
- DiCarlo JJ, Zoccolan D, Rust NC (2012) How does the brain solve visual object recognition? *Neuron* 73:415–434. PMC3306444.
- di Pellegrino G, Wise SP (1993) Visuospatial versus visuomotor activity in the premotor and prefrontal cortex of a primate. *J Neurosci* 13:1227–1243. PMC6576597.
- Dotson NM, Hoffman SJ, Goodell B, Gray CM (2018) Feature-Based Visual Short-Term Memory Is Widely Distributed and Hierarchically Organized. *Neuron* 99:215–226.e4. PMID: 29909999.
- Duan Z, Curtis CE (2024) Visual working memories are abstractions of percepts. *Elife* 13.
- Dumoulin SO, Wandell BA (2008) Population receptive field estimates in human visual cortex. *Neuroimage* 39:647–660. PMID: 17977024.
- Ehrlich DB, Murray JD (2022) Geometry of neural computation unifies working memory and planning. *Proceedings of the National Academy of Sciences* 119:e2115610119. .
- Emrich SM, Lockhart HA, Al-Aidroos N (2017) Attention mediates the flexible allocation of visual working memory resources. *J Exp Psychol Hum Percept Perform* 43:1454–1465. PMID:

28368161.

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* 33:6516–6523. PMC3664518.

Ester EF, Anderson DE, Serences JT, Awh E (2013) A neural measure of precision in visual working memory. *J Cogn Neurosci* 25:754–761. PMC4041615.

Ester EF, Serences JT, Awh E (2009) Spatially global representations in human primary visual cortex during working memory maintenance. *J Neurosci* 29:15258–15265. PMC2830793.

Ester EF, Sprague TC, Serences JT (2015) Parietal and Frontal Cortex Encode Stimulus-Specific Mnemonic Representations during Visual Working Memory. *Neuron* 87:893–905. PMC4545683.

Favila SE, Kuhl BA, Winawer J (2022) Perception and memory have distinct spatial tuning properties in human visual cortex. *Nat Commun* 13:5864. PMC9579130.

Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2001) Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* 291:312–316. PMID: 11209083.

Fukuda K, Awh E, Vogel EK (2010a) Discrete capacity limits in visual working memory. *Curr Opin Neurobiol* 20:177–182. PMID: 20362427.

Fukuda K, Vogel E, Mayr U, Awh E (2010b) Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychon Bull Rev* 17:673–679. PMID: 21037165.

Fukushima J, Akao T, Shichinohe N, Kurkin S, Kaneko CRS, Fukushima K (2011) Neuronal activity in the caudal frontal eye fields of monkeys during memory-based smooth pursuit eye movements: comparison with the supplementary eye fields. *Cereb Cortex* 21:1910–1924. PMC3138517.

Funahashi S (2013) Thalamic mediodorsal nucleus and its participation in spatial working memory processes: comparison with the prefrontal cortex. *Front Syst Neurosci* 7:36. PMC3728470.

Funahashi S, Bruce CJ, Goldman-Rakic PS (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J Neurophysiol* 61:331–349. PMID: 2918358.

Funahashi S, Bruce CJ, Goldman-Rakic PS (1993a) Dorsolateral prefrontal lesions and oculomotor delayed-response performance: evidence for mnemonic “scotomas.” *J Neurosci* 13:1479–1497. PMC6576716.

Funahashi S, Chafee MV, Goldman-Rakic PS (1993b) Prefrontal neuronal activity in rhesus monkeys performing a delayed anti-saccade task. *Nature* 365:753–756. PMID: 8413653.

Furman DJ, Zhang Z, Chatham CH, Good M, Badre D, Hsu M, Kayser AS (2020) Augmenting Frontal Dopamine Tone Enhances Maintenance over Gating Processes in Working Memory. *J Cogn Neurosci*:1–13. .

Fuster J (2015) *The Prefrontal Cortex*. Academic Press.

- Fuster JM (1973) Unit activity in prefrontal cortex during delayed-response performance: neuronal correlates of transient memory. *J Neurophysiol* 36:61–78. PMID: 4196203.
- Fuster JM (1990) Inferotemporal units in selective visual attention and short-term memory. *J Neurophysiol* 64:681–697. PMID: 2230917.
- Fuster JM, Alexander GE (1971) Neuron Activity Related to Short-Term Memory. *Science* 173:652–654. PMID: 4998337.
- Fuster JM, Bauer RH, Jervey JP (1985) Functional interactions between inferotemporal and prefrontal cortex in a cognitive task. *Brain Res* 330:299–307. PMID: 3986545.
- Fuster JM, Bodner M, Kroger JK (2000) Cross-modal and cross-temporal association in neurons of frontal cortex. *Nature* 405:347–351. PMID: 10830963.
- Fuster JM, Jervey JP (1981) Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science* 212:952–955. PMID: 7233192.
- Gayet S, Paffen CLE, Van der Stigchel S (2013) Information matching the content of visual working memory is prioritized for conscious access. *Psychol Sci* 24:2472–2480. PMID: 24121415.
- Genovesio A, Tsujimoto S, Wise SP (2009) Feature- and order-based timing representations in the frontal cortex. *Neuron* 63:254–266. PMC2803114.
- Gilbert CD, Wiesel TN (1983) Clustered intrinsic connections in cat visual cortex. *J Neurosci* 3:1116–1133. PMC6564507.
- Gilbert CD, Wiesel TN (1989) Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *J Neurosci* 9:2432–2442. PMC6569760.
- Gnadt JW, Andersen RA (1988) Memory related motor planning activity in posterior parietal cortex of macaque. *Exp Brain Res* 70:216–220. PMID: 3402565.
- Goldberg ME, Bruce CJ (1985) Cerebral cortical activity associated with the orientation of visual attention in the rhesus monkey. *Vision Res* 25:471–481. PMID: 4024466.
- Goldman-Rakic PS (1995) Cellular basis of working memory. *Neuron* 14:477–485. PMID: 7695894.
- Goldman-Rakic PS (2011) Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. *Comprehensive Physiology*.
- Hajnal MA, Tran D, Einstein M, Martelo MV, Safaryan K, Polack P-O, Golshani P, Orbán G (2023) Continuous multiplexed population representations of task context in the mouse primary visual cortex. *Nat Commun* 14:6687. PMC10590415.
- Hallenbeck GE, Sprague TC, Rahmati M, Sreenivasan KK, Curtis CE (2021a) Working memory representations in visual cortex mediate distraction effects. *Nature Communications* 2021 12:1 12:1–18. PMID: 34354071.
- Hallenbeck GE, Sprague TC, Rahmati M, Sreenivasan KK, Curtis CE (2021b) Working Memory Representations in Visual Cortex Mediate the Effects of Distraction. *Cold Spring Harbor*

Laboratory:2021.02.01.429259. .

- Hallenbeck G, Tardiff N, Sprague TC, Curtis CE (2024) Prioritizing the allocation of working memory resources depends on prefrontal cortex. *bioRxiv*.
- Harrison SA, Tong F (2009) Decoding reveals the contents of visual working memory in early visual areas. *Nature* 458:632–635. PMID: 19225460.
- Hart E, Huk AC (2020) Recurrent circuit dynamics underlie persistent activity in the macaque frontoparietal network. *Elife* 9PMC7205463.
- Hasher L, Zacks RT (1988) Working memory, comprehension, and aging: A review and a new view. *Psychol Learn Motiv* 22:193–225. .
- 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* 293:2425–2430. PMID: 11577229.
- Henderson MM, Rademaker RL, Serences JT (2022) Flexible utilization of spatial-and motor-based codes for the storage of visuo-spatial information. *Elife* 11PMID: 35522567.
- Hollingworth A, Matsukura M, Luck SJ (2013) Visual working memory modulates rapid eye movements to simple onset targets. *Psychol Sci* 24:790–796. PMC3901956.
- Hoshi E, Shima K, Tanji J (1998) Task-dependent selectivity of movement-related neuronal activity in the primate prefrontal cortex. *J Neurophysiol* 80:3392–3397. PMID: 9862940.
- Hwang J, Romanski LM (2015) Prefrontal neuronal responses during audiovisual mnemonic processing. *J Neurosci* 35:960–971. PMC4300334.
- Iamshchinina P, Christophel TB, Gayet S, Rademaker RL (2021) Essential considerations for exploring visual working memory storage in the human brain. <https://doi.org/10.1080/1350628520211915902> 29:425–436. .
- Ichihara-Takeda S, Funahashi S (2007) Activity of primate orbitofrontal and dorsolateral prefrontal neurons: task-related activity during an oculomotor delayed-response task. *Exp Brain Res* 181:409–425. PMID: 17443317.
- Inoue M, Mikami A (2006) Prefrontal activity during serial probe reproduction task: encoding, mnemonic, and retrieval processes. *J Neurophysiol* 95:1008–1041. PMID: 16207786.
- Jacobsen CF (1935) FUNCTIONS OF FRONTAL ASSOCIATION AREA IN PRIMATES. *Arch NeurPsych* 33:558–569. .
- Jacobsen CF (1936) Studies of cerebral function in primates. I. The functions of the frontal association areas in monkeys. *Comparative Psychology Monographs* 13:1–60. .
- Jacob SN, Nieder A (2014) Complementary Roles for Primate Frontal and Parietal Cortex in Guarding Working Memory from Distractor Stimuli. *Neuron* 83:226–237. PMID: 24991963.
- Jerde TA, Merriam EP, Riggall AC, Hedges JH, Curtis CE (2012) Prioritized Maps of Space in Human Frontoparietal Cortex. *Journal of Neuroscience* 32:17382–17390. PMID: 23197729.

- Johnson MK, McMahon RP, Robinson BM, Harvey AN, Hahn B, Leonard CJ, Luck SJ, Gold JM (2013) The relationship between working memory capacity and broad measures of cognitive ability in healthy adults and people with schizophrenia. *Neuropsychology* 27:220–229. .
- Jonides J, Smith EE, Koeppel RA, Awh E, Minoshima S, Mintun MA (1993) Spatial working memory in humans as revealed by PET. *Nature* 363:623–625. PMID: 8510752.
- Kamiński J, Sullivan S, Chung JM, Ross IB, Mamelak AN, Rutishauser U (2017) Persistently active neurons in human medial frontal and medial temporal lobe support working memory. *Nature Neuroscience* 20:4 20:590–601. PMID: 28218914.
- Kikuchi-Yorioka Y, Sawaguchi T (2000) Parallel visuospatial and audiospatial working memory processes in the monkey dorsolateral prefrontal cortex. *Nat Neurosci* 3:1075–1076. PMID: 11036261.
- Klyszejko Z, Rahmati M, Curtis CE (2014) Attentional priority determines working memory precision. *Vision Res* 105:70–76. PMID: 25240420.
- Kok P, de Lange FP (2014) Shape perception simultaneously up- and downregulates neural activity in the primary visual cortex. *Curr Biol* 24:1531–1535. PMID: 24980501.
- Kornblith S, Quiñan Quiroga R, Koch C, Fried I, Mormann F (2017) Persistent Single-Neuron Activity during Working Memory in the Human Medial Temporal Lobe. *Curr Biol* 27:1026–1032. PMID: 28318972.
- Kriete T, Noelle DC, Cohen JD, O'Reilly RC (2013) Indirection and symbol-like processing in the prefrontal cortex and basal ganglia. *Proc Natl Acad Sci U S A* 110:16390–16395. PMID: 24062434.
- Kritzer MF, Goldman-Rakic PS (1995) Intrinsic circuit organization of the major layers and sublayers of the dorsolateral prefrontal cortex in the rhesus monkey. *J Comp Neurol* 359:131–143. PMID: 8557842.
- Kubota K, Niki H (1971) Prefrontal cortical unit activity and delayed alternation performance in monkeys. *J Neurophysiol* 34:337–347. PMID: 4997822.
- Kumar S, Joseph S, Gander PE, Barascud N, Halpern AR, Griffiths TD (2016) A Brain System for Auditory Working Memory. *J Neurosci* 36:4492–4505. PMC4837683.
- Kwak Y, Curtis CE (2022) Unveiling the abstract format of mnemonic representations. *Neuron* 110:1822–1828.e5. .
- Lawrence SJD, van Mourik T, Kok P, Koopmans PJ, Norris DG, de Lange FP (2018) Laminar Organization of Working Memory Signals in Human Visual Cortex. *Curr Biol* 28:3435–3440.e4. PMID: 30344121.
- Leavitt ML, Mendoza-Halliday D, Martinez-Trujillo JC (2017) Sustained Activity Encoding Working Memories: Not Fully Distributed. *Trends Neurosci* 40:328–346. .
- Lee S-H, Kravitz DJ, Baker CI (2013) Goal-dependent dissociation of visual and prefrontal cortices during working memory. *Nat Neurosci* 16:997–999. PMC3781947.



- Lee TG, D'Esposito M (2012) The dynamic nature of top-down signals originating from prefrontal cortex: a combined fMRI-TMS study. *J Neurosci* 32:15458–15466. PMC3511853.
- Lepsien J, Nobre AC (2007) Attentional modulation of object representations in working memory. *Cereb Cortex* 17:2072–2083. PMID: 17099066.
- Leung H-C, Gore JC, Goldman-Rakic PS (2002) Sustained mnemonic response in the human middle frontal gyrus during on-line storage of spatial memoranda. *J Cogn Neurosci* 14:659–671. PMID: 12126506.
- Levitt JB, Lewis DA, Yoshioka T, Lund JS (1993) Topography of pyramidal neuron intrinsic connections in macaque monkey prefrontal cortex (areas 9 and 46). *J Comp Neurol* 338:360–376. PMID: 8113445.
- Libby A, Buschman TJ (2021) Rotational dynamics reduce interference between sensory and memory representations. *Nat Neurosci* 24:715–726. .
- Li D, Christ SE, Cowan N (2014) Domain-general and domain-specific functional networks in working memory. *Neuroimage* 102 Pt 2:646–656. PMC4252243.
- Liesefeld HR, Liesefeld AM, Sauseng P, Jacob SN, Müller HJ (2020) How visual working memory handles distraction: cognitive mechanisms and electrophysiological correlates. *Vis cogn* 28:372–387. .
- Li H-H, Curtis CE (2023) Neural population dynamics of human working memory. *Curr Biol* 33:3775–3784.e4. .
- Li H-H, Sprague TC, Yoo AH, Ma WJ, Curtis CE (2021a) Joint representation of working memory and uncertainty in human cortex.
- Li H-H, Sprague TC, Yoo AH, Ma WJ, Curtis CE (2021b) Joint representation of working memory and uncertainty in human cortex. *Neuron* 109:3699–3712.e6. .
- Li H-H, Sprague TC, Yoo AH, Ma WJ, Curtis CE (2024) Controlling the allocation of working memory resource. *bioRxiv*.
- Linke AC, Vicente-Grabovetsky A, Cusack R (2011) Stimulus-specific suppression preserves information in auditory short-term memory. *Proc Natl Acad Sci U S A* 108:12961–12966. PMC3150893.
- Lorenc ES, Mallett R, Lewis-Peacock JA (2021) Distraction in Visual Working Memory: Resistance is Not Futile. *Trends Cogn Sci* 25:228–239. PMC7878345.
- Lorenc ES, Sreenivasan KK (2021) Reframing the debate: The distributed systems view of working memory. *Vis cogn* 29:416–424. .
- Lorenc ES, Sreenivasan KK, Nee DE, Vandenbroucke ARE, D'Esposito M (2018) Flexible Coding of Visual Working Memory Representations during Distraction. *Journal of Neuroscience* 38:5267–5276. PMID: 29739867.
- Luck SJ, Vogel EK (1997) The capacity of visual working memory for features and conjunctions. *Nature* 390:279–281. PMID: 9384378.

- Luck SJ, Vogel EK (2013) Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends Cogn Sci* 17:391–400. PMID: 23850263.
- Lund JS, Yoshioka T, Levitt JB (1993) Comparison of intrinsic connectivity in different areas of macaque monkey cerebral cortex. *Cereb Cortex* 3:148–162. PMID: 8490320.
- Lustig C, Hasher L, Zacks RT (2007) Inhibitory deficit theory: Recent developments in a “new view.” In: *Inhibition in cognition* (pp (Gorfein DS, ed), pp 145–162. Washington, DC, US: American Psychological Association, xvii.
- MackKey WE, Curtis CE (2017) Distinct contributions by frontal and parietal cortices support working memory. *Scientific Reports* 2017 7:1 7:1–7. PMID: 28733684.
- Mackey WE, Devinsky O, Doyle WK, Golfinos JG, Curtis CE (2016a) Human parietal cortex lesions impact the precision of spatial working memory. *J Neurophysiol* 116:1049–1054. PMID: 27306678.
- Mackey WE, Devinsky O, Doyle WK, Meager MR, Curtis CE (2016b) Human Dorsolateral Prefrontal Cortex Is Not Necessary for Spatial Working Memory. *Journal of Neuroscience* 36:2847–2856. PMID: 26961941.
- Mackey WE, Winawer J, Curtis CE (2017) Visual field map clusters in human frontoparietal cortex. *Elife* 6.
- Magnussen S, Greenlee MW (1992) Retention and disruption of motion information in visual short-term memory. *J Exp Psychol Learn Mem Cogn* 18:151–156. PMID: 1532017.
- Magnussen S, Greenlee MW, Asplund R, Dyrnes S (1991) Stimulus-specific mechanisms of visual short-term memory. *Vision Res* 31:1213–1219. PMID: 1891813.
- Malmö RB (1942) INTERFERENCE FACTORS IN DELAYED RESPONSE IN MONKEYS AFTER REMOVAL OF FRONTAL LOBES. *J Neurophysiol* 5:295–308. .
- Markowitz DA, Curtis CE, Pesaran B (2015) Multiple component networks support working memory in prefrontal cortex. *Proc Natl Acad Sci U S A* 112:11084–11089. PMC4568266.
- Master SL, Li S, Curtis CE (2023) Trying harder: how cognitive effort sculpts neural representations during working memory. *bioRxiv:2023.12.07.570686*. .
- Ma WJ, Husain M, Bays PM (2014) Changing concepts of working memory. *Nature Neuroscience* 2014 17:3 17:347–356. PMID: 24569831.
- Mejías JF, Wang X-J (2022) Mechanisms of distributed working memory in a large-scale network of macaque neocortex. *Elife* 11.
- Mendoza-Halliday D, Torres S, Martínez-Trujillo JC (2014) Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. *Nature Neuroscience* 2014 17:9 17:1255–1262. PMID: 25108910.
- Merrikhi Y, Clark K, Albarran E, Parsa M, Zirnsak M, Moore T, Noudoost B (2017) Spatial working memory alters the efficacy of input to visual cortex. *Nat Commun* 8:15041. PMC5414175.

- Miller BT, Vytlačil 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. PMID: 20586702.
- Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24:167–202. PMID: 11283309.
- Miller EK, Desimone R (1994) Parallel neuronal mechanisms for short-term memory. *Science* 263:520–522. PMID: 8290960.
- Miller EK, Erickson CA, Desimone R (1996a) Neural Mechanisms of Visual Working Memory in Prefrontal Cortex of the Macaque. *Journal of Neuroscience* 16:5154–5167. PMID: 8756444.
- Miller EK, Erickson CA, Desimone R (1996b) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J Neurosci* 16:5154–5167. PMC6579322.
- Miller EK, Li L, Desimone R (1993a) Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *J Neurosci* 13:1460–1478. PMC6576733.
- Miller EK, Li L, Desimone R (1993b) Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *J Neurosci* 13:1460–1478. PMID: 8463829.
- Miller GA (1956) The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychol Rev* 63:81–97. PMID: 13310704.
- Miller JA, Tambini A, Kiyonaga A, D'Esposito M (2022) Long-term learning transforms prefrontal cortex representations during working memory. *Neuron* 0.
- Miyake A, Friedman NP, Rettinger DA, Shah P, Hegarty M (2001) How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *J Exp Psychol Gen* 130:621–640. PMID: 11757872.
- Miyake A, Shah P (1999) *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*. Cambridge University Press.
- Miyashita Y, Chang HS (1988) Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature* 331:68–70. PMID: 3340148.
- Munakata Y, Herd SA, Chatham CH, Depue BE, Banich MT, O'Reilly RC (2011) A unified framework for inhibitory control. *Trends Cogn Sci* 15:453–459. PMID: 21889391.
- Murray JD, Bernacchia A, Freedman DJ, Romo R, Wallis JD, Cai X, Padoa-Schioppa C, Pasternak T, Seo H, Lee D, Wang X-J (2014) A hierarchy of intrinsic timescales across primate cortex. *Nat Neurosci* 17:1661–1663. PMC4241138.
- Murray JD, Bernacchia A, Roy NA, Constantinidis C, Romo R, Wang XJ (2017) Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. *Proc Natl Acad Sci U S A* 114:394–399. PMID: 28028221.
- Narayanan NS, Prabhakaran V, Bunge SA, Christoff K, Fine EM, Gabrieli JDE (2005) The role of the prefrontal cortex in the maintenance of verbal working memory: an event-related fMRI analysis. *Neuropsychology* 19:223–232. PMID: 15769206.

- Niyogi RK, Wong-Lin KF (2013) Dynamic Excitatory and Inhibitory Gain Modulation Can Produce Flexible, Robust and Optimal Decision-making. *PLoS Comput Biol* 9:e1003099. PMID: 23825935.
- Norman KA, Polyn SM, Detre GJ, Haxby JV (2006) Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn Sci* 10:424–430. PMID: 16899397.
- Offen S, Schluppeck D, Heeger DJ (2009) The role of early visual cortex in visual short-term memory and visual attention. *Vision Res* 49:1352–1362. PMID: 18329065.
- O'Reilly RC, Frank MJ (2006) Making Working Memory Work: A Computational Model of Learning in the Prefrontal Cortex and Basal Ganglia. *Neural Comput* 18:283–328. .
- Palmer J (1990) Attentional Limits on the Perception and Memory of Visual Information. *J Exp Psychol Hum Percept Perform* 16:332–350. PMID: 2142203.
- Panichello MF, Buschman TJ (2021) Shared mechanisms underlie the control of working memory and attention. *Nature*:1–5. .
- Passingham RE (1985) Memory of monkeys (*Macaca mulatta*) with lesions in prefrontal cortex. *Behav Neurosci* 99:3–21. PMID: 4041231.
- Pasternak T, Greenlee MW (2005) Working memory in primate sensory systems. *Nat Rev Neurosci* 6:97–107. PMID: 15654324.
- 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* 5:805–811. PMID: 12134152.
- Pessoa L, Gutierrez E, Bandettini P, Ungerleider L (2002) Neural correlates of visual working memory: fMRI amplitude predicts task performance. *Neuron* 35:975–987. PMID: 12372290.
- Petrides M (2000) The role of the mid-dorsolateral prefrontal cortex in working memory. In: *Executive Control and the Frontal Lobe: Current Issues* (Schneider WX, Owen AM, Duncan J, eds), pp 44–54. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Ploner CJ, Rivaud-Péchoux S, Gaymard BM, Agid Y, Pierrot-Deseilligny C (1999) Errors of memory-guided saccades in humans with lesions of the frontal eye field and the dorsolateral prefrontal cortex. *J Neurophysiol* 82:1086–1090. PMID: 10444703.
- Postle BR (2006) Working memory as an emergent property of the mind and brain. *Neuroscience* 139:23–38. PMID: 16324795.
- Postle BR, D'Esposito M (1999) Dissociation of human caudate nucleus activity in spatial and nonspatial working memory: an event-related fMRI study. *Brain Res Cogn Brain Res* 8:107–115. PMID: 10407200.
- Preuschhof C, Heekeren HR, Taskin B, Schubert T, Villringer A (2006) Neural correlates of vibrotactile working memory in the human brain. *J Neurosci* 26:13231–13239. PMC6675015.
- Qi X-L, Katsuki F, Meyer T, Rawley JB, Zhou X, Douglas KL, Constantinidis C (2010) Comparison of neural activity related to working memory in primate dorsolateral prefrontal

and posterior parietal cortex. *Front Syst Neurosci* 4:12. PMC2876875.

Quintana J, Yajeya J, Fuster JM (1988) Prefrontal representation of stimulus attributes during delay tasks. I. Unit activity in cross-temporal integration of sensory and sensory-motor information. *Brain Res* 474:211–221. PMID: 3208130.

Rademaker RL, Bloem IM, De Weerd P, Sack AT (2015) The impact of interference on short-term memory for visual orientation. *J Exp Psychol Hum Percept Perform* 41:1650–1665. PMID: 26371383.

Rademaker RL, Chunharas C, Serences JT (2019) Coexisting representations of sensory and mnemonic information in human visual cortex. *Nature Neuroscience* 2019 22:8 22:1336–1344. PMID: 31263205.

Rahmati M, Curtis CE, Sreenivasan KK (2023) Mnemonic representations in human lateral geniculate nucleus. *Front Behav Neurosci* 17:1094226. .

Rahmati M, DeSimone K, Curtis CE, Sreenivasan KK (2020) Spatially Specific Working Memory Activity in the Human Superior Colliculus. *J Neurosci* 40:9487–9495. PMC7724141.

Rahmati M, Saber GT, Curtis CE (2018) Population Dynamics of Early Visual Cortex during Working Memory. *J Cogn Neurosci* 30:219–233. .

Rainer G, Miller EK (2000) Effects of visual experience on the representation of objects in the prefrontal cortex. *Neuron* 27:179–189. PMID: 10939341.

Rainer G, Rao SC, Miller EK (1999) Prospective coding for objects in primate prefrontal cortex. *J Neurosci* 19:5493–5505. PMC6782318.

Ranganath C, Cohen MX, Dam C, D'Esposito M (2004) Inferior temporal, prefrontal, and hippocampal contributions to visual working memory maintenance and associative memory retrieval. *J Neurosci* 24:3917–3925. PMC6729418.

Rao SC, Rainer G, Miller EK (1997) Integration of what and where in the primate prefrontal cortex. *Science* 276:821–824. PMID: 9115211.

Rao SG, Williams GV, Goldman-Rakic PS (2000) Destruction and creation of spatial tuning by disinhibition: GABA(A) blockade of prefrontal cortical neurons engaged by working memory. *J Neurosci* 20:485–494. PMC6774140.

Rezayat E, Clark K, Dehaqani M-RA, Noudoost B (2022) Dependence of Working Memory on Coordinated Activity Across Brain Areas. *Front Syst Neurosci* 15:787316. .

Riggall AC, Postle BR (2012) The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. *J Neurosci* 32:12990–12998. PMC3470886.

Riley MR, Constantinidis C (2016) Role of Prefrontal Persistent Activity in Working Memory. *Front Syst Neurosci* 9:168684. .

Romo R, Brody CD, Hernández A, Lemus L (1999) Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature* 399:470–473. PMID: 10365959.

- Rossi-Pool R, Zainos A, Alvarez M, Zizumbo J, Vergara J, Romo R (2017) Decoding a Decision Process in the Neuronal Population of Dorsal Premotor Cortex. *Neuron* 96:1432–1446.e7. PMID: 29224726.
- Rouder JN, Morey RD, Morey CC, Cowan N (2011) How to measure working memory capacity in the change detection paradigm. *Psychon Bull Rev* 18:324–330. PMID: 21331668.
- Rowe JB, Passingham RE (2001) Working memory for location and time: activity in prefrontal area 46 relates to selection rather than maintenance in memory. *Neuroimage* 14:77–86. PMID: 11525340.
- Rowe JB, Toni I, Josephs O, Frackowiak RS, Passingham RE (2000) The prefrontal cortex: response selection or maintenance within working memory? *Science* 288:1656–1660. PMID: 10834847.
- Roy JE, Buschman TJ, Miller EK (2014) PFC neurons reflect categorical decisions about ambiguous stimuli. *J Cogn Neurosci* 26:1283–1291. PMC4034266.
- Rypma B, Berger JS, D'Esposito M (2002) The influence of working-memory demand and subject performance on prefrontal cortical activity. *J Cogn Neurosci* 14:721–731. PMID: 12167257.
- Saber GT, Pestilli F, Curtis CE (2015) Saccade planning evokes topographically specific activity in the dorsal and ventral streams. *J Neurosci* 35:245–252. PMC4287145.
- Sadeh M, Sajad A, Wang H, Yan X, Crawford JD (2018) The Influence of a Memory Delay on Spatial Coding in the Superior Colliculus: Is Visual Always Visual and Motor Always Motor? *Front Neural Circuits* 12:74. PMC6204359.
- Sakagami M, Tsutsui Ki, Lauwereyns J, Koizumi M, Kobayashi S, Hikosaka O (2001) A code for behavioral inhibition on the basis of color, but not motion, in ventrolateral prefrontal cortex of macaque monkey. *J Neurosci* 21:4801–4808. PMC6762341.
- Sakai K, Rowe JB, Passingham RE (2002) Active maintenance in prefrontal area 46 creates distractor-resistant memory. *Nat Neurosci* 5:479–484. PMID: 11953754.
- Salthouse TA, Babcock RL, Shaw RJ (1991) Effects of adult age on structural and operational capacities in working memory. *Psychol Aging* 6:118–127. PMID: 2029360.
- Scalaidhe SP, Wilson FA, Goldman-Rakic PS (1999) Face-selective neurons during passive viewing and working memory performance of rhesus monkeys: evidence for intrinsic specialization of neuronal coding. *Cereb Cortex* 9:459–475. PMID: 10450891.
- Schluppeck D, Curtis CE, Glimcher PW, Heeger DJ (2006) Sustained activity in topographic areas of human posterior parietal cortex during memory-guided saccades. *J Neurosci* 26:5098–5108. PMC1538982.
- Schmidt TT, Wu YH, Blankenburg F (2017) Content-Specific Codes of Parametric Vibrotactile Working Memory in Humans. *Journal of Neuroscience* 37:9771–9777. PMID: 28893928.
- Schneegans S, Taylor R, Bays PM (2020) Stochastic sampling provides a unifying account of visual working memory limits. *Proc Natl Acad Sci U S A* 117:20959–20968. PMID: 32788373.

- Schumacher EH, D'Esposito M (2002) Neural implementation of response selection in humans as revealed by localized effects of stimulus–response compatibility on brain activation. *Hum Brain Mapp.*
- Schumacher EH, Elston PA, D'Esposito M (2003) Neural evidence for representation-specific response selection. *J Cogn Neurosci* 15:1111–1121. PMID: 14709230.
- Scimeca JM, Kiyonaga A, D'Esposito M (2018) Reaffirming the Sensory Recruitment Account of Working Memory. *Trends Cogn Sci* 22:190–192. PMID: 29475635.
- Semedo JD, Jasper AI, Zandvakili A, Krishna A, Aschner A, Machens CK, Kohn A, Yu BM (2022) Feedforward and feedback interactions between visual cortical areas use different population activity patterns. *Nature Communications* 2022 13:1 13:1–14. .
- Semedo JD, Zandvakili A, Machens CK, Yu BM, Kohn A (2019) Cortical Areas Interact through a Communication Subspace. *Neuron* 102:249–259.e4. PMID: 30770252.
- Serences JT (2016) Neural mechanisms of information storage in visual short-term memory. *Vision Res* 128:53–67. PMID: 27668990.
- Serences JT, Ester EF, Vogel EK, Awh E (2009) Stimulus-specific delay activity in human primary visual cortex. *Psychol Sci* 20:207–214. PMID: 19170936.
- Serences JT, Saproo S (2012) Computational advances towards linking BOLD and behavior. *Neuropsychologia* 50:435–446. .
- Seung HS, Sompolinsky H (1993) Simple models for reading neuronal population codes. *Proceedings of the National Academy of Sciences* 90:10749–10753. PMID: 8248166.
- Shao Z, Zhang M, Yu Q (2023) Preserved stimulus representations in frontal cortex as a potential neural mechanism for flexible control in working memory. *bioRxiv:2023.07.28.551058.* .
- Shen K, Valero J, Day GS, Paré M (2011) Investigating the role of the superior colliculus in active vision with the visual search paradigm. *Eur J Neurosci* 33:2003–2016. PMID: 21645096.
- Shichinohe N, Akao T, Kurkin S, Fukushima J, Kaneko CRS, Fukushima K (2009) Memory and decision making in the frontal cortex during visual motion processing for smooth pursuit eye movements. *Neuron* 62:717–732. PMC2734332.
- Shimamura AP (2000) The role of the prefrontal cortex in dynamic filtering. *Psychobiology* 28:207–218. PMID: 19809315.
- Smith EE, Jonides J, Koeppel RA (1996) Dissociating verbal and spatial working memory using PET. *Cereb Cortex* 6:11–20. PMID: 8670634.
- Smyth MM (1996) Interference with rehearsal in spatial working memory in the absence of eye movements. *Q J Exp Psychol A* 49:940–949. PMID: 8962542.
- Smyth MM, Scholey KA (1994) Interference in immediate spatial memory. *Mem Cognit* 22:1–13. PMID: 8035679.

- Sommer MA, Wurtz RH (2001) Frontal eye field sends delay activity related to movement, memory, and vision to the superior colliculus. *J Neurophysiol* 85:1673–1685. PMID: 11287490.
- Spaak E, Watanabe K, Funahashi S, Stokes MG (2017) Stable and Dynamic Coding for Working Memory in Primate Prefrontal Cortex. *Journal of Neuroscience* 37:6503–6516. PMID: 28559375.
- Sprague TC, Ester EF, Serences JT (2014) Reconstructions of information in visual spatial working memory degrade with memory load. *Curr Biol* 24:2174–2180. PMC4181677.
- Sprague TC, Ester EF, Serences JT (2016) Restoring Latent Visual Working Memory Representations in Human Cortex. *Neuron* 91:694–707. PMC4978188.
- Sreenivasan KK, Curtis CE, D'Esposito M (2014) Revisiting the role of persistent neural activity during working memory. *Trends Cogn Sci* 18:82–89. PMID: 24439529.
- Sreenivasan KK, D'Esposito M (2019) The what, where and how of delay activity. *Nature Reviews Neuroscience* 2019 20:8 20:466–481. PMID: 31086326.
- Srimal R, Curtis CE (2008) Persistent neural activity during the maintenance of spatial position in working memory. *Neuroimage* 39:455–468. PMID: 17920934.
- Standage D, Paré M (2018) Slot-like capacity and resource-like coding in a neural model of multiple-item working memory. *J Neurophysiol* 120:1945–1961. PMID: 29947585.
- St-Laurent M, Abdi H, Buchsbaum BR (2015) Distributed Patterns of Reactivation Predict Vividness of Recollection. *J Cogn Neurosci* 27:2000–2018. PMID: 26102224.
- Stroud JP, Watanabe K, Suzuki T, Stokes MG, Lengyel M (2023) Optimal information loading into working memory explains dynamic coding in the prefrontal cortex. *Proceedings of the National Academy of Sciences* 120:e2307991120. .
- 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. PMC2366068.
- Supèr H, Spekreijse H, Lamme VAF (2001) A neural correlate of working memory in the monkey primary visual cortex. *Science* 293:120–124. PMID: 11441187.
- Suzuki M, Gottlieb J (2013) Distinct neural mechanisms of distractor suppression in the frontal and parietal lobe. *Nat Neurosci* 16:98–104. PMC4207121.
- Suzuki WA, Miller EK, Desimone R (1997) Object and place memory in the macaque entorhinal cortex. *J Neurophysiol* 78:1062–1081. PMID: 9307135.
- Takeda K, Funahashi S (2002) Prefrontal task-related activity representing visual cue location or saccade direction in spatial working memory tasks. *J Neurophysiol* 87:567–588. PMID: 11784772.
- Tark K-J, Curtis CE (2009) Persistent neural activity in the human frontal cortex when maintaining space that is off the map. *Nat Neurosci* 12:1463–1468. PMC3171293.



- Teng C, Postle BR (2021) Understanding occipital and parietal contributions to visual working memory: Commentary on Xu (2020). *Vis cogn* 29:401–408. .
- Teng C, Postle BR (2024) Investigating the Roles of the Visual and Parietal Cortex in Representing Content versus Context in Visual Working Memory. *eNeuro* 11:ENEURO.0270–20.2024. .
- Thyer W, Adam KCS, Diaz GK, Velázquez Sánchez IN, Vogel EK, Awh E (2022) Storage in Visual Working Memory Recruits a Content-Independent Pointer System. *Psychol Sci* 33:1680–1694. PMID: 36006809.
- Todd JJ, Marois R (2004) Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* 428:751–754. PMID: 15085133.
- Torres-Gomez S, Blonde JD, Mendoza-Halliday D, Kuebler E, Everest M, Wang XJ, Inoue W, Poulter MO, Martinez-Trujillo J (2020) Changes in the Proportion of Inhibitory Interneuron Types from Sensory to Executive Areas of the Primate Neocortex: Implications for the Origins of Working Memory Representations. *Cereb Cortex* 30:4544–4562. PMID: 32227119.
- Uluç I, Schmidt TT, Wu Y-H, Blankenburg F (2018) Content-specific codes of parametric auditory working memory in humans. *Neuroimage* 183:254–262. PMID: 30107259.
- van Bergen RS, Jehee JFM (2021) TAFKAP: An improved method for probabilistic decoding of cortical activity. *bioRxiv:2021.03.04.433946*. .
- van Bergen RS, Ji Ma W, Pratte MS, Jehee JFM (2015) Sensory uncertainty decoded from visual cortex predicts behavior. *Nat Neurosci* 18:1728–1730. .
- van den Berg R, Shin H, Chou W-C, George R, Ma WJ (2012) Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences* 109:8780–8785. .
- Van Kerkoerle T, Self MW, Roelfsema PR (2017) Layer-specificity in the effects of attention and working memory on activity in primary visual cortex. *Nature Communications* 2017 8:1 8:1–14. PMID: 28054544.
- Wandell BA, Winawer J (2015) Computational neuroimaging and population receptive fields. *Trends Cogn Sci* 19:349–357. PMC4484758.
- Wang H, Stradtman GG 3rd, Wang X-J, Gao W-J (2008) A specialized NMDA receptor function in layer 5 recurrent microcircuitry of the adult rat prefrontal cortex. *Proc Natl Acad Sci U S A* 105:16791–16796. PMC2575498.
- Wang M, Yang Y, Wang C-J, Gamo NJ, Jin LE, Mazer JA, Morrison JH, Wang X-J, Arnsten AFT (2013) NMDA receptors subserve persistent neuronal firing during working memory in dorsolateral prefrontal cortex. *Neuron* 77:736–749. PMC3584418.
- Wang XJ (1999) Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. *J Neurosci* 19:9587–9603. PMC6782911.
- Wang XJ (2001) Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci* 24:455–463. PMID: 11476885.

- Wang X-J (2020) Macroscopic gradients of synaptic excitation and inhibition in the neocortex. *Nat Rev Neurosci* 21:169–178. PMID: 334830.
- Wang XJ (2021) 50 years of mnemonic persistent activity: quo vadis? *Trends Neurosci* 44:888–902. PMID: 34654556.
- Wang Y, Markram H, Goodman PH, Berger TK, Ma J, Goldman-Rakic PS (2006) Heterogeneity in the pyramidal network of the medial prefrontal cortex. *Nat Neurosci* 9:534–542. PMID: 16547512.
- Wei Z, Wang XJ, Wang DH (2012) From Distributed Resources to Limited Slots in Multiple-Item Working Memory: A Spiking Network Model with Normalization. *Journal of Neuroscience* 32:11228–11240. PMID: 22895707.
- Wilken P, Ma WJ (2004a) A detection theory account of change detection. *J Vis* 4:11–11. PMID: 15669916.
- Wilken P, Ma WJ (2004b) A detection theory account of visual short-term memory for color. *J Vis* 4:150–150. .
- Wilson FA, Scialoja SP, Goldman-Rakic PS (1993) Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science* 260:1955–1958. PMID: 8316836.
- Wimmer K, Nykamp DQ, Constantinidis C, Compte A (2014) Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nature Neuroscience* 2014 17:3 17:431–439. PMID: 24487232.
- Wirth S, Yanike M, Frank LM, Smith AC, Brown EN, Suzuki WA (2003) Single neurons in the monkey hippocampus and learning of new associations. *Science* 300:1578–1581. PMID: 12791995.
- Woloszyn L, Sheinberg DL (2009) Neural Dynamics in Inferior Temporal Cortex during a Visual Working Memory Task. *Journal of Neuroscience* 29:5494–5507. PMID: 19403817.
- Woodry R, Curtis CE, Winawer J (2024) Feedback scales the spatial tuning of cortical responses during visual memory. *bioRxiv*.
- Xing Y, Ledgeway T, McGraw PV, Schluppeck D (2013) Decoding working memory of stimulus contrast in early visual cortex. *J Neurosci* 33:10301–10311. PMID: 23722490.
- Xu Y (2017) Reevaluating the Sensory Account of Visual Working Memory Storage. *Trends Cogn Sci* 21:794–815. PMID: 28774684.
- Xu Y (2018) A Tale of Two Visual Systems: Invariant and Adaptive Visual Information Representations in the Primate Brain. <https://doi.org/10.1146/annurev-vision-091517-033954> 4:311–336. PMID: 29949722.
- Xu Y (2020) Revisit once more the sensory storage account of visual working memory. *Vis cogn* 28:433–446. .
- Xu Y, Chun MM (2006) Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature* 440:91–95. PMID: 16382240.

- Yoo AH, Bolaños A, Hallenbeck GE, Rahmati M, Sprague TC, Curtis CE (2022) Behavioral Prioritization Enhances Working Memory Precision and Neural Population Gain. *J Cogn Neurosci* 34:365–379. PMC9017245.
- Yoo AH, Klyszejko Z, Curtis CE, Ma WJ (2018) Strategic allocation of working memory resource. *Sci Rep* 8:16162. .
- Yu Q, Shim WM (2017) Occipital, parietal, and frontal cortices selectively maintain task-relevant features of multi-feature objects in visual working memory. *Neuroimage* 157:97–107. PMID: 28559190.
- Zaksas D, Pasternak T (2006) Directional signals in the prefrontal cortex and in area MT during a working memory for visual motion task. *J Neurosci* 26:11726–11742. PMC6674769.
- Zarahn E, Aguirre GK, D'Esposito M (1999) Temporal isolation of the neural correlates of spatial mnemonic processing with fMRI. *Brain Res Cogn Brain Res* 7:255–268. PMID: 9838152.
- Zhang W, Luck SJ (2008) Discrete fixed-resolution representations in visual working memory. *Nature* 453:233–235. PMID: 18385672.
- Zhou Y, Curtis CE, Sreenivasan KK, Fougne D (2022) Common Neural Mechanisms Control Attention and Working Memory. *J Neurosci* 42:7110–7120. .