



Working Memory Is Complex and Dynamic, Like Your Thoughts

Timothy J. Buschman¹ and Earl K. Miller²

Abstract

■ Working memory is where thoughts are held and manipulated. For many years, the dominant model was that working memory relied on steady-state neural dynamics. A neural representation was activated and then held in that state. However, as often happens, the more we examine working memory (especially with new technology), the more complex

it looks. Recent discoveries show that working memory involves multiple mechanisms, including discontinuous bouts of spiking. Memories are also dynamic, evolving in a task-dependent manner. Cortical rhythms may control those dynamics, thereby endowing top-down “executive” control over our thoughts. ■

INTRODUCTION

Over 30 years ago, working memory was solved. We had found the neural basis for holding an item in working memory. The model was straightforward. A stimulus activates neural spiking in the pFC. That activity is sustained after the stimulus disappears and its memory is held in working memory (Miller, Erickson, & Desimone, 1996; Funahashi, Bruce, & Goldman-Rakic, 1989; Fuster & Alexander, 1971). Decades of research had supported and elaborated this model. We now know working memory representations are seen in a variety of cortical areas (Christophel, Klink, Spitzer, Roelfsema, & Haynes, 2017). We learned about the important role of neuromodulators (Vijayraghavan, Wang, Birnbaum, Williams, & Arnsten, 2007). We gained insight into the biophysical and circuit mechanisms that keep activity elevated (Wang, 1999).

However, as usually happens in science, we came to realize that the neural bases of working memory are more complex than we originally thought. New technology has allowed a more detailed understanding of working memory. These insights confirmed that neural activity seen during a memory delay plays an important role in working memory. However, they also revealed working memory is not simply steady-state maintenance, like a latch circuit in your brain that turns on and off. There are bouts of spiking versus no spiking. There are dynamics and emergent properties that can only be seen at the level of neuron populations and the summed activity of millions of neurons (in local field potentials [LFPs]). Furthermore, new work has shown these dynamics support the most

important thing about working memory: It is under top-down “executive” control. We can choose what to think about and how to think about it.

Mark Stokes was a catalyst in driving this new understanding of working memory. Here, we review our take on the “Stokesian” view of working memory. There were two key insights. First, working memory is not simply the persistent activity of neurons. It is also “activity-silent” with bouts of spiking versus little or no spiking. During the “silent” periods, the memories are held by short-term plasticity mechanisms, like an echo or impression that spiking leaves in the network. Second, working memory activity is not a persistent, veridical, representation of sensory inputs. Instead, it is highly dynamic with representations that change and evolve over time. These are not unrelated insights. The activity-silent dynamics contribute to, and leave room for, emergent properties like oscillatory rhythms at different frequencies. Recent work on those rhythms has captured the neural signatures of top-down control.

WORKING MEMORY IS ACTIVITY-SILENT

The classic view of working memory is that it is represented in the sustained activity of neurons within pFC. For example, when monkeys were trained to remember the location of a reward (Fuster & Alexander, 1971) or remember the location of a stimulus (Funahashi et al., 1989), neurons in pFC were found to be tonically active for as long as the animal held the item in memory. This matched our own experience of working memory as an active process that requires effort. So, naturally, researchers assumed the sustained activity was what maintained representations over memory delays.

¹Princeton University, Princeton, NJ, ²Massachusetts Institute of Technology, Cambridge, MA

However, over the past decade, close inspection of neural activity has found neural responses in pFC are not as sustained as we once believed. As reviewed in Stokes (2015), neural responses often return to “baseline” levels after a few seconds. Furthermore, interrupting working memory maintenance, by having an animal briefly switch to another task, caused working memory representations to disappear. They then re-emerged when the animal re-engaged the original working memory task (Watanabe & Funahashi, 2014). These results led to the idea that other mechanisms may support memory representations. Several possibilities have been raised over the last few years.

First, building on theoretical work (Mongillo, Barak, & Tsodyks, 2008), Stokes and colleagues proposed working memory could be maintained in the short-term synaptic plasticity (STSP). In this model, transient neural representations, such as the ones evoked by a sensory stimulus, can temporarily change synaptic weights in the network (e.g., by altering synaptic vesicle and/or neurotransmitter receptor concentration). These changes are thought to be short term (under 1 sec) but last long enough to maintain the trace of a stimulus in the connectivity within the network over a memory delay. In other words, spiking leaves an “impression” in the network that can maintain the memory between spiking.

Of course, one inherent difficulty in testing this theory is that we cannot directly observe synaptic weights in the behaving brain—all of our methods detect neural activity. To get around this, the Stokes and Postle laboratories developed a clever approach to measuring the synaptic changes—“ping” the system with a bright visual stimulus (Wolff, Jochim, Akyürek, & Stokes, 2017) or a TMS pulse (Rose et al., 2016). If the memory is stored in short-term synaptic changes, then the neural response induced by the stimulus/pulse should change as a function of what is being held in memory. In other words, the pulse should “re-activate” the memory. Consistent with the activity-silent model, the item in memory could not be decoded (with EEG or fMRI) before the pulse. However, the memory could be decoded in the neural response following the visual stimulus/TMS pulse. Although not entirely excluding alternative explanations, these studies provide the first test of an activity-silent form of memory.

STSP may not be the only activity-silent mechanism at play in working memory tasks. Long-term episodic memory plays an important role in supporting working memory (Beukers, Buschman, Cohen, & Norman, 2021; Sutterer, Foster, Serences, Vogel, & Awh, 2019). However, one limitation of long-term memory is that it suffers from “proactive interference.” This interference occurs when two memories are similar, making it hard to distinguish a current memory from the recent past (e.g., the previous behavioral trial). Theoretical work suggests that episodic memory could mitigate interference by storing the context in which the memory occurred (DuBrow, Rouhani, Niv, & Norman, 2017; Mensink & Raaijmakers, 1988). Such

context information could provide a unique marker to sort and differentiate between different memories, keeping them from interfering with each other (Beukers et al., 2021). In this way, long-term memory could provide another activity-silent mechanism supporting working memory. Consistent with this, recent work has found proactive interference is strongest on trials in which participants must remember a large number of items (likely exceeding the capacity of working memory; Oberauer & Awh, 2022). This suggests participants may engage long-term memory only when it is helpful to supplement working memory.

Altogether, these results suggest the brain uses multiple mechanisms to maintain information in working memory. This makes sense—maintaining short-term memories of sensory inputs is critical to cognition, allowing it to break free from the immediate world. There may have been strong evolutionary pressure to develop multiple mechanisms for maintaining information in working memory. For example, recent modeling work has shown STSP can make working memory more robust. Kozachkov and colleagues (2022) trained artificial recurrent neural networks (RNNs) with and without STSP to perform an object working memory task. Both RNNs with and without STSP were able to maintain memories, even in the face of a distractor. However, RNNs with STSP were more robust to noise and network degradation than RNNs without STSP. Furthermore, RNNs with STSP showed activity that was similar to that seen in the cortex of a non-human-primate performing the same task. RNNs without STSP were more artificial, less brain-like. In short, STSP, and other activity-silent mechanisms, make working memory networks work better. Next, we discuss how working memory is also dynamic.

WORKING MEMORY IS DYNAMIC

The classic view of working memory is that it is a stable representation of recent sensory inputs. Work from Fuster, Goldman-Rakic, and others found neurons in pFC that responded to visual stimuli and then maintained spiking activity over a subsequent memory delay (Miller et al., 1996; Funahashi et al., 1989; Fuster & Alexander, 1971). However, more recent work has shown working memory is more dynamic than once thought. Newer large-scale recordings of populations of neurons have allowed us to decode neural information with far greater sensitivity than the previous single-electrode approach (King & Dehaene, 2014; Meyers, Freedman, Kreiman, Miller, & Poggio, 2008). If representations are stable, then a decoder trained on neural representations at one moment in time should be able to decode the representation at another moment in time. Alternatively, if representations are dynamic, then the decoder should fail to generalize across time. Using this approach, Stokes and colleagues (2013) showed that memory representations are highly dynamic. Decoders trained to decode the

identity of the visual stimulus when it was visible were unable to decode the memory of that same stimulus, even just 250 msec into the memory delay. This suggests that, at the population level, the neural code for sensory inputs and memories are different. Similar results have been seen in rodents (Harvey, Coen, & Tank, 2012). Interestingly, Stokes and colleagues found that, after a few hundred milliseconds, the representation stabilized (Spaak, Watanabe, Funahashi, & Stokes, 2017). Then, toward the end of the delay period, when the animal was preparing to respond, the neural representation again became dynamic.

Building on this work, Panichello and Buschman (2021) found dynamics in working memory are under cognitive control. Monkeys performed a task that required them to remember the color of two squares (Figure 1A). After a memory delay, the monkeys were cued to select one of the two squares and then, after a second memory delay, report the color of the selected square by saccading to the matching color on a color wheel. Consistent with previous work, the memory representation in pFC was dynamic. Interestingly, how the memory representation changed depended on whether it was selected for a response (or forgotten).

During the first memory delay (before selection), the color of each item was stably encoded as a ring, forming two color wheels in neural space (schematized in Figure 1B, left). Interestingly, each item's ring existed in its own independent "subspace" of neural activity. However, this changed when a memory was selected. The ring representing the color of the selected item moved from its independent subspace into a new "target" subspace (Figure 1B, right). This target subspace was the same for both items. When Memory 1 was selected, its representation moved into the target subspace, and when Memory 2 was selected, it moved into the same subspace. In other words, the dynamics of the memory depended on which memory was selected: Selecting Memory 1 induced one set of dynamics that moved Memory 1 from its

independent subspace into the target subspace, whereas selecting Memory 2 induced a different set of dynamics that transformed the representation of Memory 2 into the target subspace.

These results show dynamics in working memory are under cognitive control. But, to what purpose? The independent subspaces observed during the first memory delay make sense—the animal's task is to remember the color of each square separately, which is facilitated by the independent subspaces (Libby & Buschman, 2021). However, after selection, the animal's task changes. Now, they must report the color of the selected item, regardless of whether it used to be Memory 1 or 2. This can explain the dynamics observed in working memory. When Memory 1 is selected, the dynamics "move" the Memory 1 representation into the target subspace (and vice versa for Memory 2). Now that the selected item is in the common target subspace, downstream circuits can use this representation to drive the animal's response, regardless of which memory was selected. In this way, cognitive control may induce different dynamics to support different cognitive tasks.

This same model could explain the dynamics observed in other studies. Many of these tasks require the brain to shift from processing a sensory stimulus to preparing a motor response (classically referred to as a shift from retrospective to prospective memory; Rainer, Rao, & Miller, 1999). In other words, working memory does not just maintain a veridical representation of inputs. Rather, it exists to support cognition and behavior. From this perspective, it makes sense that working memory representations would be dynamic—they evolve in a way that facilitates the task at hand.

WORKING MEMORY IS RHYTHMIC (AND RHYTHMS ARE CONTROL)

Working memory is under top-down ("executive") control. We can choose what to encode in working memory,

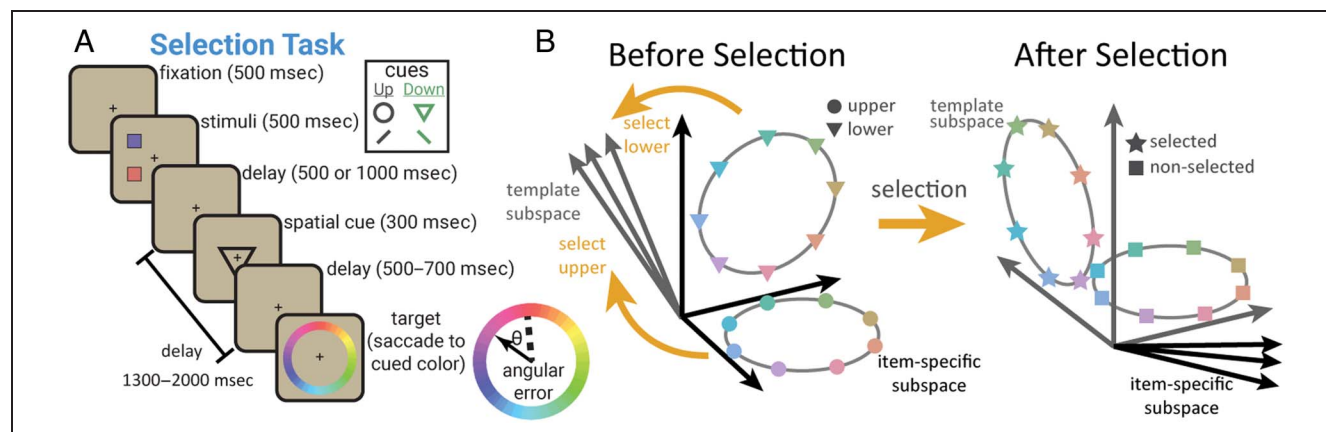


Figure 1. Model of dynamic control of working memory. (A) Behavioral task for selecting an item from working memory. (B) Memory representations transformed in a task-dependent manner. Before selection, the color of each item was represented in an independent subspace within the neural population in LPFC (left). Selection transformed the selected item into a new "target" subspace (right) that was used to guide behavior. Adapted from Panichello and Buschman (2021).

we can manipulate those thoughts, and we can ignore distractions and choose to stop thinking those thoughts. pFC plays a key role in controlling working memory (Panichello & Buschman, 2021; Gazzaley & Nobre, 2012; Miller & Cohen, 2001). As discussed above, top-down control of neural dynamics may change how working memory representations are used. Mounting evidence suggests control may arise from oscillatory dynamics that emerge at a higher level of integration—in the LFPs. LFPs are the summation of coordinated, oscillatory, activity of millions of neurons. The electrical fields that arise from this activity can act as “guard rails” that control higher-dimensional, neuron-level activity by funneling it along stable low-dimensional routes (Pinotsis & Miller, 2022).

The central idea is that sensory information (the contents of working memory) and control signals use different frequency bands that interact. As reviewed by Miller, Lundqvist, and Bastos (2018), recent work suggests sensory information is carried by spiking associated with bursts of gamma (>30 Hz) power. The top-down control signals are carried by alpha/beta rhythms (8–30 Hz). Alpha/beta inhibits gamma wherever they “collide” in cortex. Thus, top-down (alpha/beta) controls bottom-up (gamma/spiking). Figure 2 shows how this works. The top-down alpha/beta are carried in the deep layers of cortex. The deep cortical layers carry feedback signals down the cortical hierarchy. Bottom-up sensory information in gamma/spiking is carried in the superficial cortical layers that send signals in a feedforward manner, up the cortical hierarchy. Alpha/beta originating in the deep layers inhibits gamma/spiking in the superficial layers (Bastos, Loonis, Kornblith, Lundqvist, & Miller, 2018).

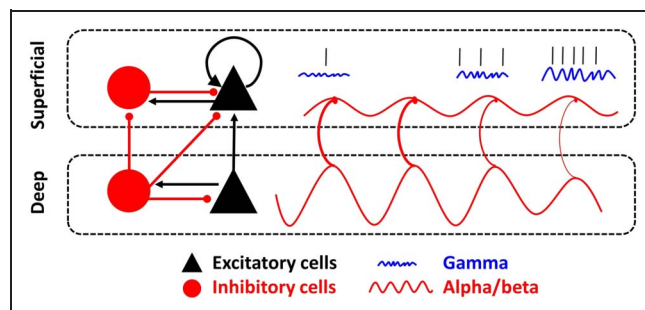


Figure 2. Top-down control model of working memory by brain rhythms. Inhibitory connections are line segments with a red, rounded end, and excitatory connections are line segments with a black, arrow end. The sinusoidal red line in deep layers reflects beta oscillations and their driving influence on superficial beta oscillations. Beta oscillations are phase-amplitude coupled with gamma oscillations (blue squiggly lines), and these gamma oscillations organize delay-period spiking representing working memory content (straight black marks). Over time, moving from left to right in the figure, the deep beta reduces in power and releases inhibition onto the superficial layers. This results in enhanced superficial gamma and spiking. The reversed process (enhancement of deep layer beta, enhanced suppression of superficial layer gamma/spiking) would “clear out” the contents of working memory. From the work of Bastos and colleagues (2018).

To gain access to working memory, deep-layer alpha/beta power and/or its coupling to superficial layer beta weakens. This disinhibits recurrent excitation of superficial layer neurons, generating bursts of gamma and spiking to sensory inputs. During memory maintenance, the balance between alpha/beta and gamma can regulate the level of spiking to occasionally refresh the synaptic weight changes that help maintain memories (Miller et al., 2018).

To read out information from working memory, alpha/beta power/coherence drops. This allows increased gamma bursting and the ramp-up of spiking often seen near the end of memory delays (Hussar & Pasternak, 2010; Roesch & Olson, 2005). The disinhibition of gamma increases spiking so that the memories can acquire control of behavior. Balance between alpha/beta and gamma during the memory delay can keep gamma to a moderate level. That way, spiking does not prematurely gain control over behavior. To clear out working memory, beta power/coupling increases. This suppresses gamma and the spiking that was maintaining the memory. Examples of these dynamics can be found in Lundqvist and colleagues (2016) and Lundqvist, Herman, Warden, Brincat, and Miller (2018).

These dynamics may have a role in many cognitive functions, not just working memory. The superficial layer gamma and deep-layer alpha-beta is a ubiquitous motif seen across all of cortex (Mendoza-Halliday et al., 2022; Lundqvist, Bastos, & Miller, 2020). Recent work suggests that the same dynamics play a role in predictive coding (Bastos, Lundqvist, Waite, Kopell, & Miller, 2020). It is possible that much of cognitive control stems from the balance and control of these rhythms.

TOP-DOWN CONTROL BY SPATIAL COMPUTING

Thus far, we have discussed alpha/beta as if it were a coarse-gating signal. It turns working memory on and off like turning a faucet or a light switch on and off. However, its control can be more specific, operating on the level of individual contents of working memory. In Predictive Coding, for example, alpha/beta targets representations of specific stimuli in visual cortex to inhibit processing of predicted sensory inputs (Bastos et al., 2020).

Recent work has shown top-down information, such as the task at hand, is carried by unique patterns of alpha/beta synchrony across cortex (Antzoulatos & Miller, 2014, 2016; Buschman, Denovellis, Diogo, Bullock, & Miller, 2012). In other words, patterns of alpha/beta form neural “ensembles” that reflect top-down information. Importantly, the spatial resolution of these patches is on the macro-scale. It is seen at the level of LFPs that reflect the summed activity of millions of neurons. These LFPs synchronize across millimeters or more of cortex (sometimes across large expanses of cortex). It is these patterns that may provide the control.

The idea is called “spatial computing” (Lundqvist et al., 2022). It suggests that patterns of alpha/beta power and coherence create a macro-scale patchwork of higher power alpha/beta versus higher power gamma across cortical networks. Wherever alpha-beta power is low, gamma and spiking are high and vice versa. Different patterns of alpha/beta result in different patterns of gamma/spiking. By contrast, stimulus information (e.g., the contents of working memory) is represented at a much finer scale. It is carried by patterns of activity of (and connectivity between) individual neurons (rather than millions of neurons). Stimulus information is widely distributed and repeated across the networks, like sand across a larger scale “checkboard” pattern of alpha/beta and gamma. In other words, the contents (stimuli) and the control of working memory operate on very different spatial scales. Stimulus representation is high-dimensional, reflected by spiking patterns of populations of individual neurons. By contrast, control is low-dimensional, operating at the level of groups of millions of neurons via patterns of alpha–beta versus gamma coherence.

Control comes from *where* in network space a stimulus representation is currently expressed. The patterns of alpha/beta versus gamma and *changes in those patterns* are computations. Applying a set of operations (e.g., executing a task’s rules) corresponds to imposing different macro scale patterns of alpha/beta and gamma. Items can be accessed and operated on just by knowing their place in network space.

To understand how this works, consider a task requiring an animal to remember two objects (A and B) in the order in which they appeared (first or second; Lundqvist et al., 2018; Warden & Miller, 2007). Just before the first object is shown, the alpha/beta patterns create a mirror-image pattern of gamma. That specific pattern corresponds to “1st item.” When the object appears (say, object A), neurons in the gamma patches selective for that object are activated, priming them (via STSP). Next, before the second object is shown, a different pattern of alpha/beta sets up a different pattern of gamma that corresponds to “2nd item.” When the second object (say object B) appears, neurons in those gamma patches are activated and primed. To maintain and read out which object was first or second, the pattern corresponding to the first or second item is re-established. The primed neurons in the corresponding gamma patches will “ring back” and spike more strongly. When, for example, the patchwork corresponding to the first object is re-established, the neurons ring back with “object A” because they were primed by that object when it appeared first. In short, spatial computing posits that working memory control stems from the spatio-temporal activity patterns across network space that reflect and change with top–down task demands.

The separation of content versus control to high versus low-dimensional scales solves a critical issue in many neural network models. In typical models, the rules of the task (the control) and the content (e.g., the items

held in working memory) are both encoded in the high-dimensional details of connectivity between spiking neurons. Because of this lack of separation, if one wants to introduce novel items into working memory, the network has to be retrained. They do not show the flexibility of working memory seen in humans and animals. Typical network models cannot do “zero-shot” learning (instant generalization) that real brains can (see Bouchacourt & Buschman, 2019, for a different model of flexibility; O’Reilly & Frank, 2006). Spatial Computing solves this by separating control versus content into different scales of representation.

From a normative perspective, the brain may use a small (low-dimensional) set of control states to flexibly adapt to new situations (MacDowell, Tafazoli, & Buschman, 2022). Adapting to a new situation requires the brain to identify the control state that is appropriate for the current situation. This control state will determine how information is processed, maintained, and used to guide behavior in that situation. Although a large number of high-dimensional control states would allow for precise control of behavior, optimizing it for the current situation, this would also make it difficult to identify the “best” control state. In contrast, if the brain uses a small, low-dimensional set of control states, then it will be easier to find the best one of the set. However, such low-dimensional control states will be necessarily coarse and, so, imperfect. This suggests that there is a trade-off between high-dimensional control states, which would be accurate but slow to adapt, and low-dimensional control states, which would be flexible, yet suboptimal. Given this, the brain may choose to sample a limited number of control states that can balance precision and flexibility (MacDowell et al., 2022).

SUMMARY

Working memory is central to cognition, acting as a workspace on which thoughts are stored and manipulated. Given its importance, it is no surprise that multiple mechanisms have evolved to support the maintenance of working memory. Classic results showing the sustained representations of items in working memory are not wrong, they are just an incomplete picture. Mounting evidence points to other mechanisms and emergent properties. The neural basis of working memory is complex and dynamic, just as Mark Stokes told us. It is in these dynamics that we have gained insight into both how we hold items “in mind” and how those thoughts are controlled. Working memory is not yet solved, but the work of Mark Stokes showed us a path to a deeper level of understanding.

Acknowledgments

This work was supported by NIMH R01MH115042 (T. J. B.) and ONR N00014-22-1-2453, NEI 1R01EY033430, and The JPB Foundation (E. K. M.).

Reprint requests should be sent to Timothy J. Buschman, Princeton Neuroscience Institute and Department of Psychology, Princeton University, PNI 256, Washington Road, Princeton, NJ 08544, or via e-mail: tbuschma@princeton.edu.

Author Contributions

Timothy J. Buschman: Conceptualization; Writing—Original draft; Writing—Review & editing. Earl K. Miller: Conceptualization; Writing—Original draft; Writing—Review & editing.

Funding Information

Earl K. Miller, National Eye Institute (<https://dx.doi.org/10.13039/1000000053>), grant number: 1R01EY033430. Earl K. Miller, JPB Foundation (<https://dx.doi.org/10.13039/100007457>). Earl K. Miller, Office of Naval Research (<https://dx.doi.org/10.13039/1000000006>), grant number: N00014-22-1-2453. Timothy J. Buschman, National Institute of Mental Health (<https://dx.doi.org/10.13039/1000000025>), grant number: R01MH115042.

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience (JoCN)* during this period were $M(\text{an})/M = .407$, $W(\text{oman})/M = .32$, $M/W = .115$, and $W/W = .159$, the comparable proportions for the articles that these authorship teams cited were $M/M = .549$, $W/M = .257$, $M/W = .109$, and $W/W = .085$ (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance. The authors of this article report its proportions of citations by gender category to be as follows: $M/M = .829$; $W/M = .098$; $M/W = .073$; $W/W = 0$.

REFERENCES

Antzoulatos, E. G., & Miller, E. K. (2014). Increases in functional connectivity between prefrontal cortex and striatum during category learning. *Neuron*, 83, 216–225. <https://doi.org/10.1016/j.neuron.2014.05.005>, PubMed: 24930701

Antzoulatos, E. G., & Miller, E. K. (2016). Synchronous beta rhythms of frontoparietal networks support only behaviorally relevant representations. *eLife*, 5, e17822. <https://doi.org/10.7554/eLife.17822>, PubMed: 27841747

Bastos, A. M., Loonis, R., Kornblith, S., Lundqvist, M., & Miller, E. K. (2018). Laminar recordings in frontal cortex suggest distinct layers for maintenance and control of working memory. *Proceedings of the National Academy of Sciences, U.S.A.*, 115, 1117–1122. <https://doi.org/10.1073/pnas.1710323115>, PubMed: 29339471

Bastos, A. M., Lundqvist, M., Waite, A. S., Kopell, N., & Miller, E. K. (2020). Layer and rhythm specificity for predictive routing. *Proceedings of the National Academy of Sciences, U.S.A.*, 117, 31459–31469. <https://doi.org/10.1073/pnas.2014868117>, PubMed: 33229572

Beukers, A. O., Buschman, T. J., Cohen, J. D., & Norman, K. A. (2021). Is activity silent working memory simply episodic memory? *Trends in Cognitive Sciences*, 25, 284–293. <https://doi.org/10.1016/j.tics.2021.01.003>, PubMed: 33551266

Bouchacourt, F., & Buschman, T. J. (2019). A flexible model of working memory. *Neuron*, 103, 147–160. <https://doi.org/10.1016/j.neuron.2019.04.020>, PubMed: 31103359

Buschman, T. J., Denovellis, E. L., Diogo, C., Bullock, D., & Miller, E. K. (2012). Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. *Neuron*, 76, 838–846. <https://doi.org/10.1016/j.neuron.2012.09.029>, PubMed: 23177967

Christophel, T. B., Klink, P. C., Spitzer, B., Roelfsema, P. R., & Haynes, J.-D. (2017). The distributed nature of working memory. *Trends in Cognitive Sciences*, 21, 111–124. <https://doi.org/10.1016/j.tics.2016.12.007>, PubMed: 28063661

DuBrow, S., Rouhani, N., Niv, Y., & Norman, K. A. (2017). Does mental context drift or shift? *Current Opinion in Behavioral Sciences*, 17, 141–146. <https://doi.org/10.1016/j.cobeha.2017.08.003>, PubMed: 29335678

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology*, 61, 331–349. <https://doi.org/10.1152/jn.1989.61.2.331>, PubMed: 2918358

Fuster, J. M., & Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science*, 173, 652–654. <https://doi.org/10.1126/science.173.3997.652>, PubMed: 4998337

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences*, 16, 129–135. <https://doi.org/10.1016/j.tics.2011.11.014>, PubMed: 22209601

Harvey, C. D., Coen, P., & Tank, D. W. (2012). Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature*, 484, 62–68. <https://doi.org/10.1038/nature10918>, PubMed: 22419153

Hussar, C., & Pasternak, T. (2010). Trial-to-trial variability of the prefrontal neurons reveals the nature of their engagement in a motion discrimination task. *Proceedings of the National Academy of Sciences, U.S.A.*, 107, 21842–21847. <https://doi.org/10.1073/pnas.1009956107>, PubMed: 21098286

King, J.-R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: The temporal generalization method. *Trends in Cognitive Sciences*, 18, 203–210. <https://doi.org/10.1016/j.tics.2014.01.002>, PubMed: 24593982

Kozachkov, L., Tauber, J., Lundqvist, M., Brincat, S. L., Slotine, J.-J., & Miller, E. K. (2022). Robust and brain-like working memory through short-term synaptic plasticity. *bioRxiv*. <https://doi.org/10.1101/2022.01.09.475558>

Libby, A., & Buschman, T. J. (2021). Rotational dynamics reduce interference between sensory and memory representations. *Nature Neuroscience*, 24, 715–726. <https://doi.org/10.1038/s41593-021-00821-9>, PubMed: 33821001

Lundqvist, M., Bastos, A. M., & Miller, E. K. (2020). Preservation and changes in oscillatory dynamics across the cortical hierarchy. *Journal of Cognitive Neuroscience*, 32, 2024–2035. https://doi.org/10.1162/jocn_a_01600, PubMed: 32573380

Lundqvist, M., Herman, P., Warden, M. R., Brincat, S. L., & Miller, E. K. (2018). Gamma and beta bursts during working memory readout suggest roles in its volitional control. *Nature Communications*, 9, 394. <https://doi.org/10.1038/s41467-017-02791-8>, PubMed: 29374153

Lundqvist, M., Rose, J., Herman, P., Brincat, S. L., Buschman, T. J., & Miller, E. K. (2016). Gamma and beta bursts underlie

- working memory. *Neuron*, 90, 152–164. <https://doi.org/10.1016/j.neuron.2016.02.028>, PubMed: 26996084
- Lundqvist, M., Brincat, S. L., Rose, J., Warden, M. R., Buschman, T., Miller, E. K., et al. (2022). Spatial computing for the control of working memory. *bioRxiv*. <https://doi.org/10.1101/2020.12.30.424833>
- MacDowell, C. J., Tafazoli, S., & Buschman, T. J. (2022). A Goldilocks theory of cognitive control: Balancing precision and efficiency with low-dimensional control states. *Current Opinion in Neurobiology*, 76, 102606. <https://doi.org/10.1016/j.conb.2022.102606>, PubMed: 35870301
- Mendoza-Halliday, D., Major, A. J., Lee, N., Lichtenfeld, M., Carlson, B., Mitchell, B., et al. (2022). A ubiquitous spectrolaminar motif of local field potential power across the primate cortex. *bioRxiv*. <https://doi.org/10.1101/2022.09.30.510398>
- Mensink, G.-J., & Raaijmakers, J. G. (1988). A model for interference and forgetting. *Psychological Review*, 95, 434–455. <https://doi.org/10.1037/0033-295X.95.4.434>
- Meyers, E. M., Freedman, D. J., Kreiman, G., Miller, E. K., & Poggio, T. (2008). Dynamic population coding of category information in inferior temporal and prefrontal cortex. *Journal of Neurophysiology*, 100, 1407–1419. <https://doi.org/10.1152/jn.90248.2008>, PubMed: 18562555
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167–202. <https://doi.org/10.1146/annurev.neuro.24.1.167>, PubMed: 11283309
- Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience*, 16, 5154–5167. <https://doi.org/10.1523/JNEUROSCI.16-16-05154.1996>, PubMed: 8756444
- Miller, E. K., Lundqvist, M., & Bastos, A. M. (2018). Working memory 2.0. *Neuron*, 100, 463–475. <https://doi.org/10.1016/j.neuron.2018.09.023>, PubMed: 30359609
- Mongillo, G., Barak, O., & Tsodyks, M. (2008). Synaptic theory of working memory. *Science*, 319, 1543–1546. <https://doi.org/10.1126/science.1150769>, PubMed: 18339943
- Oberauer, K., & Awh, E. (2022). Is there an activity-silent working memory? *Journal of Cognitive Neuroscience*, 34, 2360–2374. https://doi.org/10.1162/jocn_a_01917, PubMed: 36122353
- O'Reilly, R. C., & Frank, M. J. (2006). Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*, 18, 283–328. <https://doi.org/10.1162/089976606775093909>, PubMed: 16378516
- Panichello, M. F., & Buschman, T. J. (2021). Shared mechanisms underlie the control of working memory and attention. *Nature*, 592, 601–605. <https://doi.org/10.1038/s41586-021-03390-w>, PubMed: 33790467
- Pinotsis, D. A., & Miller, E. K. (2022). Beyond dimension reduction: Stable electric fields emerge from and allow representational drift. *Neuroimage*, 253, 119058. <https://doi.org/10.1016/j.neuroimage.2022.119058>, PubMed: 35272022
- Rainer, G., Rao, S. C., & Miller, E. K. (1999). Prospective coding for objects in primate prefrontal cortex. *Journal of Neuroscience*, 19, 5493–5505. <https://doi.org/10.1523/JNEUROSCI.19-13-05493.1999>, PubMed: 10377358
- Roesch, M. R., & Olson, C. R. (2005). Neuronal activity dependent on anticipated and elapsed delay in macaque prefrontal cortex, frontal and supplementary eye fields, and premotor cortex. *Journal of Neurophysiology*, 94, 1469–1497. <https://doi.org/10.1152/jn.00064.2005>, PubMed: 15817652
- Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosses, O., Starrett, M. J., Meyering, E. E., et al. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science*, 354, 1136–1139. <https://doi.org/10.1126/science.1227011>, PubMed: 27934762
- Spaak, E., Watanabe, K., Funahashi, S., & Stokes, M. G. (2017). Stable and dynamic coding for working memory in primate prefrontal cortex. *Journal of Neuroscience*, 37, 6503–6516. <https://doi.org/10.1523/JNEUROSCI.3364-16.2017>, PubMed: 28559375
- Stokes, M. G. (2015). ‘Activity-silent’ working memory in prefrontal cortex: A dynamic coding framework. *Trends in Cognitive Sciences*, 19, 394–405. <https://doi.org/10.1016/j.tics.2015.05.004>, PubMed: 26051384
- Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., & Duncan, J. (2013). Dynamic coding for cognitive control in prefrontal cortex. *Neuron*, 78, 364–375. <https://doi.org/10.1016/j.neuron.2013.01.039>, PubMed: 23562541
- Sutterer, D. W., Foster, J. J., Serences, J. T., Vogel, E. K., & Awh, E. (2019). Alpha-band oscillations track the retrieval of precise spatial representations from long-term memory. *Journal of Neurophysiology*, 122, 539–551. <https://doi.org/10.1152/jn.00268.2019>, PubMed: 31188708
- Vijayraghavan, S., Wang, M., Birnbaum, S. G., Williams, G. V., & Arnsten, A. F. T. (2007). Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. *Nature Neuroscience*, 10, 376–384. <https://doi.org/10.1038/nn1846>, PubMed: 17277774
- Wang, X.-J. (1999). Synaptic basis of cortical persistent activity: The importance of NMDA receptors to working memory. *Journal of Neuroscience*, 19, 9587–9603. <https://doi.org/10.1523/JNEUROSCI.19-21-09587.1999>, PubMed: 10531461
- Warden, M. R., & Miller, E. K. (2007). The representation of multiple objects in prefrontal neuronal delay activity. *Cerebral Cortex*, 17(suppl_1), i41–i50. <https://doi.org/10.1093/cercor/bhm070>, PubMed: 17726003
- Watanabe, K., & Funahashi, S. (2014). Neural mechanisms of dual-task interference and cognitive capacity limitation in the prefrontal cortex. *Nature Neuroscience*, 17, 601–611. <https://doi.org/10.1038/nn.3667>, PubMed: 24584049
- Wolff, M. J., Jochim, J., Akyürek, E. G., & Stokes, M. G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nature Neuroscience*, 20, 864–871. <https://doi.org/10.1038/nn.4546>, PubMed: 28414333