Cambridge University Press
978-1-108-70644-5 — Competition and Control during Working Memory
Anastasia Kiyonaga , Mark D'Esposito
Excerpt
More Information

1 Introduction

As humans, we often need to keep several thoughts and goals in mind while we also face immediate demands in our external environment. In other words, we need to "maintain" pertinent information with working memory while we are busy doing (and remembering) other things. We know from everyday experience that this can be difficult. Imagine rehearsing an unfamiliar phone number, which becomes virtually impossible when interrupted by a few moments of conversation or, worse, a stream of numbers on the radio. However, we can also successfully juggle our internal maintenance and external demands in many circumstances. For instance, we can remind ourselves of a shopping list while operating a car and navigating traffic. These conflicting observations, of either fragile or robust working memory, fuel an ongoing debate about whether working memory and (externally oriented) attention overlap in their structure and function. This Element will address this debate by examining when and why internal working memory maintenance competes with externally oriented attention, as well as how goal-directed control processes and whole-brain network reconfiguration can reduce that competition.

Many of the terms that will be used in this Element can take on several meanings, and they often encompass multiple concepts, abilities, or processes. For instance, "attention" refers to something that can be oriented toward its subject, sustained across time, or serve as a filter for what information gets processed. It can operate over locations, objects, or features and be driven by internal goals or environmental cues. Yet across all of these functions, a fundamental property of attention is that it is capacity limited. As a result, attention must be selective. Accordingly, the term "attention" will be used here to mean the selective focusing on a subset of information for further processing. This is achieved by modulation of the neural activity associated with the information being processed, often in a goal-oriented – or "top-down" – manner. That is, perception of a stimulus in the outside world results in a neural response pattern in sensory cortex that depicts the stimulus, and we refer to this depiction as a sensory "representation." When attention is focused on a particular external stimulus, its associated sensory representation is enhanced. Thus, perception activates sensory representations, and those representations are modulated by attention. Here, the term "perceptual attention" will be used to describe attention directed toward external stimuli that are currently present in the environment.

Attention can also be focused internally to maintain representations of recently perceived stimuli or to endogenously activate long-term memory representations for use in the short term. Here, the term "working memory maintenance" will be used to describe attention that is directed internally to modulate representations of information that is unavailable to the senses. Working memory (WM) retains temporary representations that serve many purposes. For instance, representations of the recent sensory past may bridge moments of experience, while more abstract goals may guide behavior throughout an episode. Like "attention," the term "working memory" is often loosely applied and can be used alternatingly to mean a system for short-term retention, a process of short-term maintenance, or a storage location for temporary representations. WM operates across the sensory modalities (verbal, visual, etc.) and wide ranges of task manipulations are used as tests of WM. These tasks place variable demands on processes such as perceptual encoding of the information to be remembered, updating or manipulation of the internal representations, binding of item features, resolution of interference from distractions or disruptions, as well as memory retrieval and decision making at the time when memory is probed. These many cognitive processes can likewise engage distinct neural processes, which rely on varied brain regions and networks. This variability likely contributes to ongoing debates about how WM operates and how WM information is stored. Here, "WM" will refer to the attentional process of maintenance, and the maintained representations that are stored with WM will be referred to as "WM content." This content can span a range including (but not limited to) low-level stimulus features, abstract categorical information, and complex task rules.

Finally, while goal-oriented perceptual attention and WM maintenance are both achieved by top-down modulation of relevant representations, that modulatory process can face competition from additional demands. For instance, visual perceptual attention can be taxed by irrelevant interfering stimuli in the visual field, and WM maintenance can be taxed by a requirement to perform some operation on the WM content. Moreover, perceptual attention can be taxed by concurrent WM maintenance demands, and vice versa. When there is competition among content (e.g., several stimuli to attend) or processes (e.g., internal maintenance vs. external selection), attentional control must be engaged to reduce the competition. That is, there can be local-level competition for stimulus representation as well as systems-level competition among task processes. Here we will use the umbrella term "attentional control" to describe the processes that regulate this competition. That is, while perceptual attention and WM maintenance are themselves modulatory processes that guide ongoing behavior, here we will use "attentional control" to describe the additional modulatory processes that prioritize and apply goal-relevant representations in the face of competition.

This Element will examine local-level properties of WM content to illuminate the storage format of WM, as well as systems-level properties of WM maintenance to illuminate the attentional processes underlying WM. In the first half of the Element, we will address two related issues about the nature of WM: (1) whether WM maintenance activates the same sensory representations as perceptual attention (i.e., whether WM content is "stored" in sensory cortices) and (2) whether WM maintenance and perceptual attention are analogous processes that are achieved via the same neural mechanisms. A central argument against the perspective that WM and perceptual attention share representations and mechanisms is that WM performance is often unimpaired by concurrent perceptual attentional demands (and vice versa). We will first describe evidence for a gradient of competition between WM and perceptual representations that resembles established competitive effects in perceptual attention. This gradient should only exist if WM and perception activate overlapping representations. We will also examine how WM maintenance and perceptual attention can be modulated to limit competition between them, clarifying how these cognitive processes might not always compete (despite activating shared representations). In the second half of the Element, we will examine the brain-wide processes underlying WM maintenance as well as the attentional control processes that are engaged when maintenance is taxed. There is currently widespread debate about the storage substrates and processes that support working memory. Yet the research that sustains this debate has tended to focus on isolated brain regions or maintenance mechanisms, rather than integrating across the brain and across levels of analysis. Therefore, the second half of the Element will address the functional roles of the many brain regions beyond sensory cortex that are involved in WM maintenance, and how those relate to sensory WM representations. We will integrate both cognitive and neuroscientific accounts to propose a multi-level network architecture of WM and perceptual attention that provides robustness to competition and allows flexible attentional control in line with goals.

2 Competition Reveals Mechanisms of WM Storage and Processing

There is now broad agreement that WM and perceptual attention are related functions, engaging many similar mechanisms and brain regions (Awh & Jonides, 2001; Chun, 2011; Chun & Johnson, 2011; Gazzaley & Nobre, 2012; Theeuwes, Kramer, & Irwin, 2011). However, whether WM and perceptual attention truly operate by the same processes and representational substrates is still under debate (Bae & Luck, 2018; Harrison & Bays, 2018; Hollingworth &

Cambridge University Press
978-1-108-70644-5 — Competition and Control during Working Memory
Anastasia Kiyonaga , Mark D'Esposito
Excerpt
More Information

Maxcey-Richard, 2012; Leavitt, Mendoza-Halliday, & Martinez-Trujillo, 2017; Mendoza-Halliday & Martinez-Trujillo, 2017; Myers, Stokes, & Nobre, 2017; Nee & Jonides, 2009; van Kerkoerle, Self, & Roelfsema, 2017; Woodman & Luck, 2010; Xu, 2017). For instance, one perspective on WM storage proposes that representations of the remembered content are maintained in prefrontal and parietal brain regions (Leavitt et al., 2017; Xu, 2017), while another proposes dedicated storage "buffers" that are specific to the type of content that is being maintained (e.g., Baddeley & Hitch, 2018; Yue, Martin, Hamilton, & Rose, 2019). These perspectives share the proposal that WM information is represented separately from perceptual stimuli. A clear advantage of such an organization is that it would protect WM content from interference by incoming perceptual input. However, another prominent view asserts that WM activates the same sensory cortical representations as perception (D'Esposito & Postle, 2015; Lara & Wallis, 2015; Pasternak & Greenlee, 2005; Postle, 2006; Serences, 2016; Sreenivasan, Curtis, & D'Esposito, 2014). This framework is appealing in its efficiency, as the same cortical territory can be multipurposed for several kinds of processing, but the theory must reconcile how WM withstands simultaneous perceptual input. The debate between these perspectives can be informed by examining how interactions within and between WM and perceptual content influence behavioral and neural responses. The following sections will therefore examine two hypotheses: (1) If WM activates sensory representations for storage, then perceptual input should influence WM content (and task performance) proportional to the degree of cortical overlap between representations and (2) if WM maintenance operates by the same top-down modulation processes as perceptual attention, then it should be associated with comparable patterns of behavior and be taxed by concurrent perceptual attentional demands.

2.1 WM Maintenance and Perceptual Attention Share Representations

2.1.1 Perceptual Input Influences WM Maintenance

Perceptual attention is limited by the amount and nature of information that it must process. With greater numbers of items and greater similarity between items (e.g., in either location or stimulus features), both behavioral performance and neural response patterns associated with individual items are degraded (Desimone & Duncan, 1995; Pelli & Tillman, 2008; Reddy, Kanwisher, & VanRullen, 2009). That is, items that are represented closer to each other in cortical space are more likely to produce competing neural representations and

Cambridge University Press
978-1-108-70644-5 — Competition and Control during Working Memory
Anastasia Kiyonaga , Mark D'Esposito
Excerpt
More Information

therefore inhibit each other behaviorally (Franconeri, Alvarez, & Cavanagh, 2013). Consequently, the number of spatial locations that can be attentionally selected is increased when the separation between target locations is increased (Franconeri, Jonathan, & Scimeca, 2010). If WM maintenance and perceptual attention engage the same cortical substrate, WM content and perceptual content should also compete, and in the same manner as perceived stimuli. That is, there should be increased competition between WM and perceptual representations when there is greater overlap between the features of the WM and perceptual content. We can therefore learn about the structure of WM representations by examining how WM performance is influenced by concurrent perceptual input, as described throughout this section.

At the broadest level, when two WM tasks are interleaved, concurrent WM content from the same sensory modality – also sometimes referred to as the WM "domain" – impairs performance more than when the two tasks require memory of content from different modalities. For instance, during verbal WM maintenance for a string of letters or words, a concurrent verbal WM task is more disruptive to performance than a visual WM task (Baddeley, 2003). Although this empirical finding has historically been interpreted as evidence for domain-specific WM storage buffers (e.g., the phonological loop or visuo-spatial sketchpad), it is equally compatible with a revised model of WM storage wherein there is less competition between WM content from different sensory modalities because those representations overlap less in sensory cortex. If perceptual input (rather than concurrent WM demand) were to produce a similar pattern of interference with WM maintenance, it would provide compelling evidence that WM and perceptual content activate shared representations (i.e., interference occurs in sensory cortices rather than dedicated WM buffers).

In keeping with this hypothesis, WM performance is in fact slower and less accurate when distractors presented during WM maintenance are from the same category as the WM content (as opposed to a different category). For instance, during WM delayed recognition for images of faces – when participants were asked to remember a set of images and determine whether a test image matched any items from the WM sample set – performance was impaired more by distractor images of other faces during the WM delay than by images from different categories (e.g., scenes or shoes; Jha & Kiyonaga, 2010; Sreenivasan & Jha, 2007; Yoon, Curtis, & D'Esposito, 2006). Moreover, when WM for faces is tested with continuous recall (rather than recognition), WM recall precision is biased to be more similar to distractor faces that were shown during the delay (Mallett, Mummaneni, & Lewis-Peacock, 2020). Low-level stimuli that vary on a continuous dimension (e.g., orientations or directions of motion) also display a similar gradient of distractor influence, demonstrating that competition between representations of WM content and perceptual input occurs at both the category and exemplar levels, for complex objects as well simple visual features (Magnussen & Greenlee, 1992; Magnussen, Greenlee, Asplund, & Dyrnes, 1991; Nemes, Parry, Whitaker, & McKeefry, 2012; Pasternak & Greenlee, 2005; Rademaker, Bloem, De Weerd, & Sack, 2015; Teng & Kravitz, 2019).

Neural response patterns are also sensitive to the relationship between WM content and perceptual input, illuminating the source of the competition that manifests in behavior. For instance, WM maintenance activity is observed with functional magnetic resonance imaging (fMRI) in the same cortical regions that respond to perception of stimuli from the WM category (Druzgal & D’Esposito, 2001; Ranganath, DeGutis, & D’Esposito, 2004). This finding provides rudimentary evidence that WM storage recruits sensory representations. That interpretation is further corroborated by observations that sensory cortical activity is modulated by the similarity between perceptual input and concurrent WM content. For instance, when a series of faces is presented for comparison to a face item that is being maintained with WM, activity amplitude in face-processing visual regions increases with increasing feature similarity between the WM content and the perceptual input (Sreenivasan, Gratton, Vytlacil, & D’Esposito, 2014). This suggests that WM maintenance may be achieved by feedback signals to perceptual processing and representation regions, which thereby modulate sensory cortical responses to perceptual input. Consistent with this proposed feedback maintenance mechanism, WM for face stimuli is associated with correlated fMRI activity (i.e., functional connectivity) between the lateral prefrontal cortex (PFC) and face-responsive visual regions, suggesting that information is maintained in WM by coactivation between PFC and sensory regions. In other words, lateral PFC may exert top-down modulatory signals that bias or tune sensory representations for WM maintenance.

Maintenance-related coactivation between PFC and face-processing regions is also disrupted by face (but not scene) distractor stimuli, suggesting that this maintenance signal is affected by perceptual input, but only when it competes with the WM content (Postle, Druzgal, & D’Esposito, 2003; Yoon et al., 2006). However, this disruption of presumed communication between PFC and sensory regions during distraction may serve an adaptive purpose – to protect WM from interfering content. For instance, if WM modulation of sensory regions promotes increased responsivity to similar perceptual input (which can impede WM representations), then it would be beneficial to dampen this activity to limit processing of potentially interfering input. Consistent with that interpretation, when distractor face images are presented during a WM maintenance delay

CAMBRIDGE

Cambridge University Press
978-1-108-70644-5 — Competition and Control during Working Memory
Anastasia Kiyonaga , Mark D'Esposito
Excerpt
More Information

period, the face-sensitive N170 event-related potential (ERP) response to those faces is reduced when WM is also for face stimuli (as compared to a different WM category; Sreenivasan & Jha, 2007). This suggests that attentional processing of interfering distractors may be reduced to curtail their impact on WM. Therefore, task-relevant perceptual input can enhance neural signals associated with WM maintenance, while task-irrelevant and detrimental input can disrupt those signals, and both of these modulating effects depend on the degree of similarity between WM and perceptual content.

Going further, multivariate neuroimaging analysis techniques have recently been developed to gauge the pattern of population responses across a set of measurements. These stand in comparison to analysis methods that gauge the mean response amplitude of a region. For instance, for multivariate analyses, electroencephalogram (EEG) data can be analyzed according to the pattern of responses across multiple scalp electrodes, or fMRI data can be analyzed according to the pattern of activity across a selection of voxels (i.e., cubes, or volumetric pixels, that are the MRI unit of measurement in three-dimensional brain space). Machine-learning methods have been applied to such locally distributed patterns of neural activity to either classify (i.e., decode) or reconstruct stimulus representations across a population (Norman, Polyn, Detre, & Haxby, 2006; Sprague, Saproo, & Serences, 2015). These methods have shown that information about WM content features can be extracted from the sensory regions that respond to perception of that type of content (**Figure 1a**). This WM content information can be detected even when sustained average activation of the regions is absent or fails to distinguish between features (Harrison & Tong, 2009; Kiyonaga, Dowd, & Egner, 2017; Lorenc, Lee, Chen, & D’Esposito, 2015; Riggall & Postle, 2012; Serences, Ester, Vogel, & Awh, 2009; Sreenivasan, Curtis et al., 2014; Sreenivasan, Vytlacil, & D’Esposito, 2014). The consistent detection of WM feature information across sensory cortices has provided compelling support for the idea that WM representations are stored in the same cortical regions that process sensory input (Postle, 2015). Moreover, these multivariate classification and reconstruction techniques can detect the features of what is being remembered using classifiers that are trained on the activity patterns measured during stimulus perception. This provides further evidence that perceptual content and WM content are represented by shared activation substrates.

Population-level multivariate measures can also serve as an index of stimulus representation quality (Ester, Anderson, Serences, & Awh, 2013) and therefore reveal a more precise description of how WM and perceptual content compete. One example of this approach used an inverted encoding model analysis method to reconstruct the angle of remembered visual orientation stimuli under

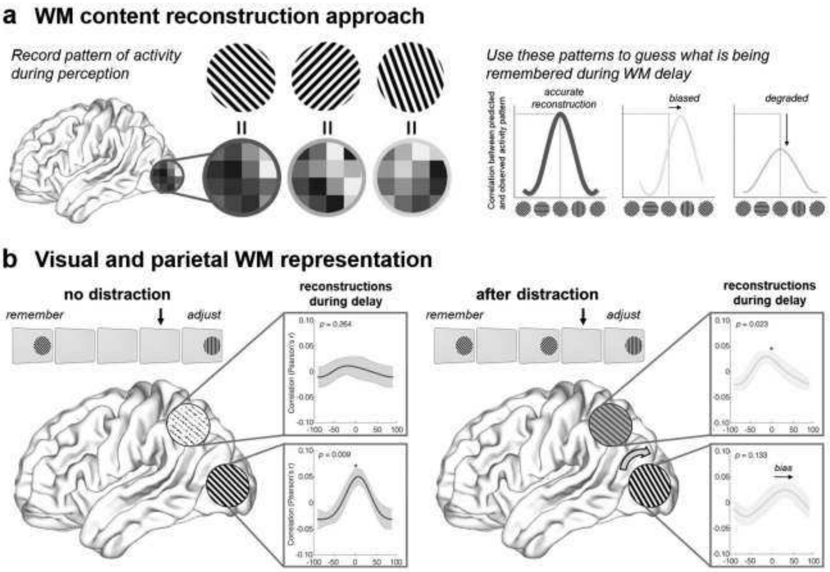

Figure 1 Perceptual input competes with WM content. (a) Left panel illustrates how different patterns of locally distributed activity, in this case across visual cortex voxels, correspond to different stimulus features, in this case the angle of oriented gratings. Right panel shows example orientation tuning functions or "reconstructions" of population-level stimulus representations during WM maintenance. (b) Results reproduced from Lorenc et al. (2018) illustrating WM stimulus reconstructions across visual and parietal regions during WM maintenance with and without distraction. After distraction (right), visual cortex representations were biased toward the angle of the distractor stimulus, while unbiased representations were detected in parietal cortex that were absent when there was no distraction (left).

different WM delay conditions (see **Figure 1**; see Sprague et al., 2018 for an overview of this method). When distractor orientation stimuli were presented during the delay, WM orientation reconstructions from visual cortical activity patterns were biased toward the angle of the distractor (Lorenc, Sreenivasan, Nee, Vandenbroucke, & D'Esposito, 2018). For instance, if an irrelevant distractor stimulus were oriented 45° clockwise relative to the WM stimulus, the visual cortical activity pattern associated with the WM representation looked like an activity pattern for a stimulus that was rotated ~15° clockwise relative to the WM sample. Thus, the feature information in the memory-related activity pattern was pulled in the direction of the distractor (see **Figure 1b**). This finding suggests that the memory representation itself can be corrupted (or modified) by perceptual distraction. This shift in neural WM stimulus reconstructions was

CAMBRIDGE

Cambridge University Press
978-1-108-70644-5 — Competition and Control during Working Memory
Anastasia Kiyonaga , Mark D'Esposito
Excerpt
More Information

also accompanied by an error in continuous report WM behavior: on average, participants recalled an angle biased slightly toward the angle of the distractor stimulus, suggesting that distraction may influence WM behavior by contaminating sensory WM representations.

However, we know from everyday experience that we can often simultaneously perceive and remember things successfully. How is this feasible if WM maintenance and perceptual input activate shared representations? One possibility may be the context-specific recruitment of additional regions to support WM when necessary. As an example from the orientation WM task described earlier, when a distractor was presented during WM maintenance, an unbiased (though less precise) WM content representation could be reconstructed from a region around the intraparietal sulcus (IPS; see **Figure 1b**). Yet such WM feature information was not discernible from IPS when distraction was absent (Lorenc et al., 2018; see also Bettencourt & Xu, 2016; Christophel, Iamshchinina, Yan, Allefeld, & Haynes, 2018; Rademaker, Chunharas, & Serences, 2019). Thus, WM may successfully rely on maintenance via high-fidelity representations in sensory cortex when no perceptual competition exists but may recruit additional regions and maintenance formats to protect WM content during perceptual competition (cf. Gayet, Paffen, & Stigchel, 2018). Indeed, WM-relevant representations at different levels of abstraction (e.g., precise stimulus identity, broad category, task-rule) are widely distributed across the cortex (Christophel, Klink, Spitzer, Roelfsema, & Haynes, 2017). WM representations in higher-order cortical regions may be coarser but able to bridge disruptions and reinstate high-fidelity sensory representations via top-down signals when necessary (Scimeca, Kiyonaga, & D'Esposito, 2018). Therefore, the representation of WM content in sensory cortex in no way precludes the representation of WM-relevant information in other regions. Instead, WM representations at several levels of the processing stream may integrate to promote successful WM maintenance.

Together, these findings demonstrate that perceptual input influences WM neural representations and behavior. The pattern of distractor influence on WM maintenance suggests that WM content is represented in the same cortical regions as perceptual input of the same type. WM is influenced more by perceptual input from the same sensory domain, input from the same category within the same sensory domain, and input with the same features within the same category – all suggesting that representations of both WM and perceptual content rely on the same cortical content maps, which interfere more when they overlap more. However, WM content information is represented in additional cortical regions when distracting perceptual input competes for representation

(Lorenc et al., 2018), and neural responses during WM maintenance are sensitive to the current task goals (Sreenivasan, Gratton et al., 2014; Yoon et al., 2006). These findings suggest that attentional control processes can adaptively manage simultaneous WM and perceptual processing demands in line with goals, and those control processes will be further examined in section 2.1.3 and in the second half of the Element.

2.1.2 WM Maintenance Influences Perceptual Attention

The previous section described how perceptual input can impact WM behavior and neural representations. If this does in fact occur because WM maintenance activates the same sensory representations as perceptual attention, then we should also observe the converse effect whereby WM content impacts perceptual processing. In support of that hypothesis, a great deal of evidence demonstrates that WM maintenance can increase perceptual sensitivity and attentional capture for features in the environment that match the WM content. Consider that you are sitting in your car at a stoplight, rehearsing your to-do list, envisioning the green shirt you need to pick up at the dry cleaner (imagine "green shirt, green shirt, green shirt"). If a person walks through the crosswalk wearing a green shirt, you will likely notice that person immediately, while others in blue suits or yellow dresses will not capture your attention. This can occur without consciously looking for a green shirt, and possibly at the expense of the more pressing task of monitoring the traffic. This sort of occurrence may also be beneficial, however, if the active memory content overlaps with currently relevant goals – for instance, if your immediate task were to pick up a green shirt–clad friend at the airport. In the laboratory, when a simple visual WM stimulus is maintained and a visual search is performed during the WM delay, search performance is best if the search target matches the features of the WM content and worst if the WM content matches search distractors (Dowd, Kiyonaga, Beck, & Egner, 2015; Dowd, Kiyonaga, Egner, & Mitroff, 2015; Kiyonaga, Egner, & Soto, 2012; Soto et al., 2008). For instance, if the task is to remember a yellow square for a WM test, a concurrent visual search during WM maintenance will be fastest when the target is located within a yellow square and slowest when a distractor is located within a yellow square (compared with a neutral condition wherein no yellow square appears; **Figure 2**). This type of modulation by WM occurs in the early stages of processing (Gayet et al., 2017; Hollingworth, Matsukura, & Luck, 2013; Saad & Silvanto, 2013; Soto, Wriglesworth, Bahrami-Balani, & Humphreys, 2010), which provides compelling evidence that the WM content is being represented in the same regions that process the sensory input.