

Opinion

Working memory needs pointers

Edward Awh ^{1,*} and Edward K. Vogel¹

Cognitive neuroscience has converged on a definition of working memory (WM) as a capacity-limited system that maintains highly accessible representations via stimulus-specific neural patterns. We argue that this standard definition may be incomplete. We highlight the fundamental need to recognize specific instances or tokens and to bind those tokens to the surrounding context. We propose that contextual binding is supported by spatiotemporal ‘pointers’ and that pointers are the source of neural signals that track the number of stored items, independent of their content. These content-independent pointers may provide a productive perspective for understanding item-based capacity limits in WM and the role of WM as a gateway for long-term storage.

A gap in the standard definition of WM

WM is an online memory system that is a core component of intelligent behavior. Individual differences in WM capacity predict a wide range of cognitive abilities, including fluid intelligence, scholastic achievement, and attentional control [1–4]. Nevertheless, we highlight a fundamental question regarding our definition of WM storage. What are the necessary and sufficient conditions for concluding that an item is stored in WM? A consensus on this basic question will be essential to determine the neural mechanisms for storage and to understand how and when WM capacity limits will affect performance. We highlight a critical facet of WM storage: the binding of item representations to the surrounding event context. This contextual binding, sometimes referred to as ‘tokenization’ [5,6], involves the spatiotemporal tracking of discrete items through an ongoing event. Although the importance of this operation has been elucidated both in prominent models of memory [6–10] and in theories of dynamic visual cognition [11,12], most neural studies of WM storage have focused primarily on stimulus-specific activity patterns (e.g., [12–14]) that may be functionally dissociated from item-based binding signals. Moreover, many cognitive and neural models have argued that capacity in WM is limited entirely by content-based interference [15–17], without considering whether binding is a limiting factor. Thus, careful consideration of the distinction between the representation of each item’s content and the binding of those representations to context may reveal distinct classes of storage-related neural activity [9,18–20], one of which may underlie item-based or ‘chunk-based’ limits in WM performance [21–26].

Defining basic constructs such as WM is challenging. Thus, the literature has offered a wide variety of definitions that differ in important ways [27]. Nevertheless, three specific cognitive and neural assumptions feature in virtually all accounts within cognitive neuroscience. First, this memory system is strongly capacity-limited. Although there is debate regarding the specific reasons for limited capacity in WM [17,28], it is clear that there are sharp limits on concurrent storage in WM. Second, cognitive theories emphasize the heightened accessibility of WMs [29]. This idea of online availability motivates the view that WM is critical for complex cognition such as mental arithmetic or problem solving. Finally, a third facet of the modal definition was prompted by early observations of stimulus-specific neural activity that tracks the content of items stored in WM [13]. These item-specific neural traces index the identity [30] or location [31] of behaviorally relevant items, even during blank delay periods when the items are no longer physically present.

Highlights

Within cognitive neuroscience, the definition of storage in working memory (WM) has focused on the cognitive accessibility of the stored representations and the stimulus-specific neural patterns that track them. We argue that this definition may be incomplete.

We propose a refined definition that acknowledges the fundamental need to track entities through time and space and to bind those representations to the surrounding context. In line with past theories, we postulate the operation of a spatiotemporal ‘pointer’ system that enables this tracking process.

This pointer operation may be the basis of neural signals that track the number of individuated items stored in WM while generalizing across variations in both the type and number of visual features stored. Limits on the number of pointers that can be deployed may help to explain why capacity in WM is item-based rather than determined by information load.

¹University of Chicago, Chicago, IL, USA*Correspondence:
awh@uchicago.edu (E. Awh).

These observations, plus a wide variety of conceptual replications [e.g., 24–27], have established stimulus-specific neural patterns as the dominant operational definition of WM storage within the neuroscience literature. Thus, although recent work has challenged the assumption that WM storage requires persistent or unbroken patterns of neural activity (Box 1), stimulus-specific neural patterns are widely regarded as a key ‘fingerprint’ of WM storage. To summarize, cognitive neuroscience has converged on a standard definition of WM as a capacity-limited system that maintains highly accessible representations of information that are represented via stimulus-specific neural patterns.

Challenges for the standard definition

Although there is continued debate regarding whether WM storage is limited by the number of items stored [22,28] or by the total amount of information [15,17], there is a strong consensus that WM storage is sharply limited. Indeed, this is one of the most compelling functional differences between WM and long-term memory (LTM). Because any successful definition will acknowledge WM capacity limits, we focus our critique on the other two key features of the standard definition: cognitive accessibility and stimulus-specific neural activity.

Cognitive accessibility

The standard definition’s emphasis on heightened accessibility is motivated by the notion that rapid access is critical for complex cognition. But a key challenge for this as a defining feature of WM storage is that heightened accessibility is not unique to WM. In their seminal paper [29], Ericsson and Kintsch discussed the vital role of LTM in complex tasks, arguing that skilled

Box 1. Activity-silent WM

Is persistent neural activity necessary for the retention of information in WM? Although neural activity measured with EEG and fMRI tracks the contents of WM and predicts individual differences in WM capacity [58,60,82], in some cases, temporary interruptions of these neural signals may have only minor consequences for later access to the stored information [83,84]. Although it is sometimes difficult to tell when all neural signals have been interrupted [85], such results have motivated the hypothesis that WM storage can be accomplished via rapid changes in the synaptic connectivity, without requiring persistent spiking activity within the cellular assemblies that represent stored information [86,87]. This phenomenon has been labeled ‘activity-silent’ WM.

An alternative to activity-silent WM is that observers offload information into episodic LTM and then retrieve that information when it is required [88–90]. This could explain how observers maintain cognitive access to those memories without postulating activity-silent WM storage. Although most extant work cannot rule out this possibility, a recent study showed that under conditions known to interrupt neurally active signatures of WM storage, observers could still maintain about four memory representations that were resistant to proactive interference (PI) [91]. Given that both capacity limits and resistance to PI are key functional signatures of WM storage, these findings lend further support to the possibility of activity-silent WM. This may appear at odds with the central thesis of this review. After all, we are postulating that spatiotemporal pointers – indexed by neurally active patterns of EEG activity – are a defining feature of WM storage. Can these ideas be reconciled? One possible reconciliation is that the neural pointer operation that we are highlighting may be essential for ‘short-term consolidation’ (STC), the process of transitioning between an initial sensory representation and more durable storage in WM [92,93]. Indeed, ‘durable storage’ is operationally defined by resistance to masking and interference from subsequent stimuli, and these properties of consolidated memories depend on a precise index of the time and place of encoding. Past estimates of the time needed for STC vary widely, from hundreds of milliseconds to well over 1 s [92,93], in line with the latency and duration of neural pointer signals. Thus, while this hypothesis still highlights an essential role for pointers during WM encoding, it allows for lingering accessible memory representations as long as a neurally active process of STC has completed.

Finally, a hypothesis that preserves a link between persistent neural activity and WM storage is that memories that are resistant to PI may not be unique to WM. One possibility is that observers can form rapid associations in LTM with adequate contextual cues to resist PI [94]. Thus, even if WM storage depends on persistent neural activity, this hypothesis could explain how memories resistant to PI could still be retrieved when those neural signals have been silenced. One attractive feature of this argument is that it preserves the common proposition that WMs are ‘in mind.’ Given that extant demonstrations of activity-silent WM have all involved disengagement of attention from the stored items, the activity-silent construct presumes a disconnect between WM storage and awareness of what is stored. Thus, refining our understanding of the boundaries between WM and LTM will be crucial for the effort to understand the role of neurally active representations in WM.

performance often relies on highly efficient access to information stored in LTM rather than WM. Although earlier theorists had presumed that LTM storage and retrieval were too slow to serve such a computational role, the efficiency of these operations was typically measured using novel sequences of arbitrarily associated materials. Although this approach helps to control for uncontrolled influences of past knowledge and skill with the memory materials, it fails to acknowledge the transformational importance of past knowledge and skill with the memory materials! Thus, with sufficient expertise, people exhibit rapid access to far more information than can be maintained in WM alone. For example, chess experts can apprehend and subsequently retrieve complex chess positions far more rapidly than traditional LTM experiments suggest. Likewise, prolonged training enabled a dedicated subject to increase his digit span performance from around seven to more than 80 digits [32], a feat that relied on rapid LTM access of specific configurations of digit stimuli. Thus, LTM can be highly accessible – that is, working – when subjects have sufficient experience with the formation and retrieval of memories from a specific domain.

In line with this observation, recent work highlights how LTM can be a powerful guiding force for various ‘online’ tasks such as visual search [33,34] or the selection of targets among distractors [35]. On the one hand, these demonstrations have elucidated the constant interplay between WM and LTM. On the other hand, they show that heightened accessibility does not suffice as a defining feature for storage in WM.

Stimulus-specific neural activity

Measurements of stimulus-specific activity provided a key advance in our understanding of how WM is enacted in the human brain [36–39]. Reconstruction of stimulus identities on the basis of neural activity increases confidence that this activity supports the WM representation itself rather than other task-related but non-mnemonic processes. That said, stimulus-specific patterns of neural activity may not be sufficient to conclude that an item has been encoded into WM. This possibility is underscored by various demonstrations of late selection, the selection of items to be encoded – or rejected – from WM after stimulus identification [40–43]. For example, past work measured the relative speeds of ‘perceiving a scene’ versus ‘remembering a scene’ by presenting subjects with rapid serial visual presentations (RSVPs) of real-world scenes at different speeds [44]. Recognition memory performance was poor even for the slower (i.e., 333 ms) presentation rates. However, target detection performance was very good even for the faster presentation rates (e.g., 167 ms). Thus, the processes involved in ‘perceiving a scene’ are much faster than those needed to successfully store those scenes in memory.

Results such as this demonstrate that extensive perceptual processing – including determining the ‘gist’ of a scene – can occur despite the subject having poor access to those items in memory. In fact, recent work with macaque monkeys viewing RSVP streams found that they could decode the identity of each stimulus in the sequence using the local field potentials in the prefrontal cortex even when the presentation rate was 10 items per second [45]. Likewise, electroencephalogram (EEG) studies with humans have revealed category and stimulus-specific patterns of neural activity evoked from RSVP streams with up to a 20-Hz presentation rate [46], even though memory encoding at such speeds is very poor. These are not simply low-level neural patterns reflecting precategorical sensory activity. For example, magnetoencephalography recordings show that the long-term memorability of an irrelevant scene (presented during a task in which observers were searching for faces) can be decoded within about 150 ms of stimulus onset in an ultrafast RSVP display (34 ms/item) [47]. Although this neural signature of memorability was sustained across four or five subsequent image presentations in the RSVP stream, observers had no memory of distractor scenes even after 30 repetitions. These findings show that intrinsic memorability – a relatively high-level property of a visual item – can be robustly encoded without storing that item in WM.

Finally, although seminal neural studies of WM focused on spatially specific signals that were sustained across a blank delay period [31], there is growing evidence that such signals can be generated without encoding into WM. For example, when observers are instructed to direct covert attention toward varying numbers of locations, sustained spatially specific signatures of attention are observed in the absence of well-established neural signatures of WM storage [48]. Thus, spatial attention and WM gating may represent distinct facets of attentional control, a dichotomy that highlights important boundaries for our definition of the pointer construct (Box 2). Moreover, these findings provide another clear example of sustained stimulus-specific neural patterns without storage in WM.

Box 2. Content-independent pointers and spatial attention: distinct forms of attentional control

Our emphasis on how pointers mediate spatiotemporal tracking raises important questions about the relationship between pointers and covert spatial attention. Although we have argued that pointer assignment may enable the binding of items to context, there are long-standing claims that spatial attention may mediate binding [95–97]. Moreover, there is substantial evidence that spatial attention is actively deployed during storage in visual WM [38,98] and that attention-based rehearsal enhances WM storage [99,100]. These findings notwithstanding, recent work has suggested a clear dissociation between the deployment of covert spatial attention and the encoding of items into WM. For instance, while a robust, set-size dependent CDA is observed when observers are asked to remember the locations of two or four simple objects, the CDA is largely absent when observers instead direct spatial attention to two or four locations in anticipation of an upcoming target item [48]. This implies that WM encoding (measured by the CDA) is not obligated when attention is directed toward an empty region of space, even though this cueing task required precise memory for the cued positions across a blank delay period. Thus, neural signals indexing encoding into WM are tracking a process that goes beyond the deployment of covert spatial attention (and the maintenance of information *per se*).

Another recent study that reveals a dissociation between spatial attention and WM storage [19] compared EEG signatures of spatial attention and WM storage using ‘dot cloud’ stimuli that afforded independent manipulations of the number of items stored and the breadth of covert spatial attention within the display (Figure 1). Variations in the size of the attended regions were precisely tracked by spatially specific alpha activity. These changes in attentional breadth, however, had no impact on a robust multivariate measure of the number of items encoded into WM. At the same time, variations in the number of individuated items stored in WM had no effect on alpha measures of the locus and breadth of covert attention. Thus, distinct neural patterns track spatial attention and WM storage, and separate factors modulate these disparate neural signals. Finally, neural signals tracking spatial attention and WM storage explain distinct variance in WM performance [101], and they have different time courses for their response to irrelevant onsets [84]. Thus, while there is clearly a tight collaboration between covert spatial attention and WM storage [98,102,103], they are dissociable aspects of attentional control that likely have separable computational roles in memory.

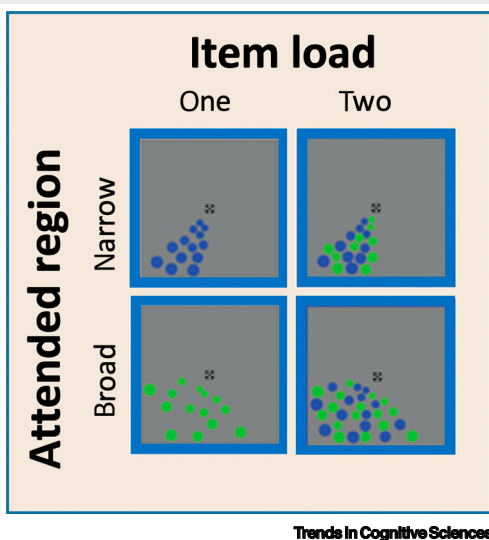


Figure 1. Independent manipulation of spatial attention and pointer deployment. Observers stored the position and shape of all presented dot clouds. Because the clouds could overlap with minimal interference, the design afforded independent manipulation of the breadth of spatial attention and the number of items stored in working memory.

We are not arguing that stimulus-specific neural activity is unimportant for storage in WM. To the contrary, this class of neural activity is likely essential for online memories. Nevertheless, we have highlighted various cases in which stimulus-specific neural patterns can be observed even when the stimulus has not been encoded into WM. Thus, stimulus-specific activity is necessary but not sufficient to hold items in an online state in WM.

Content-independent pointers

We have argued that cognitive accessibility and stimulus-specific neural representations may be necessary but not sufficient to conclude that an item has been encoded into WM. What is the missing ingredient in the standard definition? Here, we propose that contextual binding, the registration of an item's representation to the surrounding event context, may be a defining feature of WM storage [6,7]. This argument draws support from the discovery of neural signals that track the number of items stored in WM, independent of the specific features contained within the stored items. We propose that these signals reflect the deployment of content-independent pointers that support the contextual binding of stored items and serve as a springboard for the creation of episodic memories.

The need for abstract indexing

Long-standing theories of dynamic visual cognition [11,12] have noted that perception and action depend critically on the observer's ability to continuously track selected items through time and space. In turn, continuous tracking requires an indexing operation that will be robust even when the visible properties of tracked items change over time (e.g., due to changes in lighting, viewing angle). Thus, 'object files' [12] were proposed as a mechanism for registering specific items or tokens [5,49] in the visual field and forming a continuous episodic representation of those items within the surrounding event. Likewise, 'fingers of instantiation' (FINSTs) were proposed to describe the process of indexing or 'pointing at' relevant tokens in a scene, enabling them to be tracked through space and time despite changes in appearance or spatial position [11]. Thus, both object files and FINSTs refer to a spatiotemporal 'pointer' operation that tracks tokens through time and space, even when the visible properties of the tracked items are unstable. Critically, both theories also acknowledge a distinction between the indexing operation supported by pointers and the maintenance of the featural details associated with each token. This is why pointer deployment is sometimes described as an 'abstract indexing' operation. Here, there is a useful analogy between content-independent pointers and demonstrative words such as 'this' or 'that' [11]; demonstrative words hold no meaning by themselves, but they serve to refer to other words. Likewise, pointers may refer to specific tokens in a visual scene, while parallel but distinct operations maintain the contents of those items.

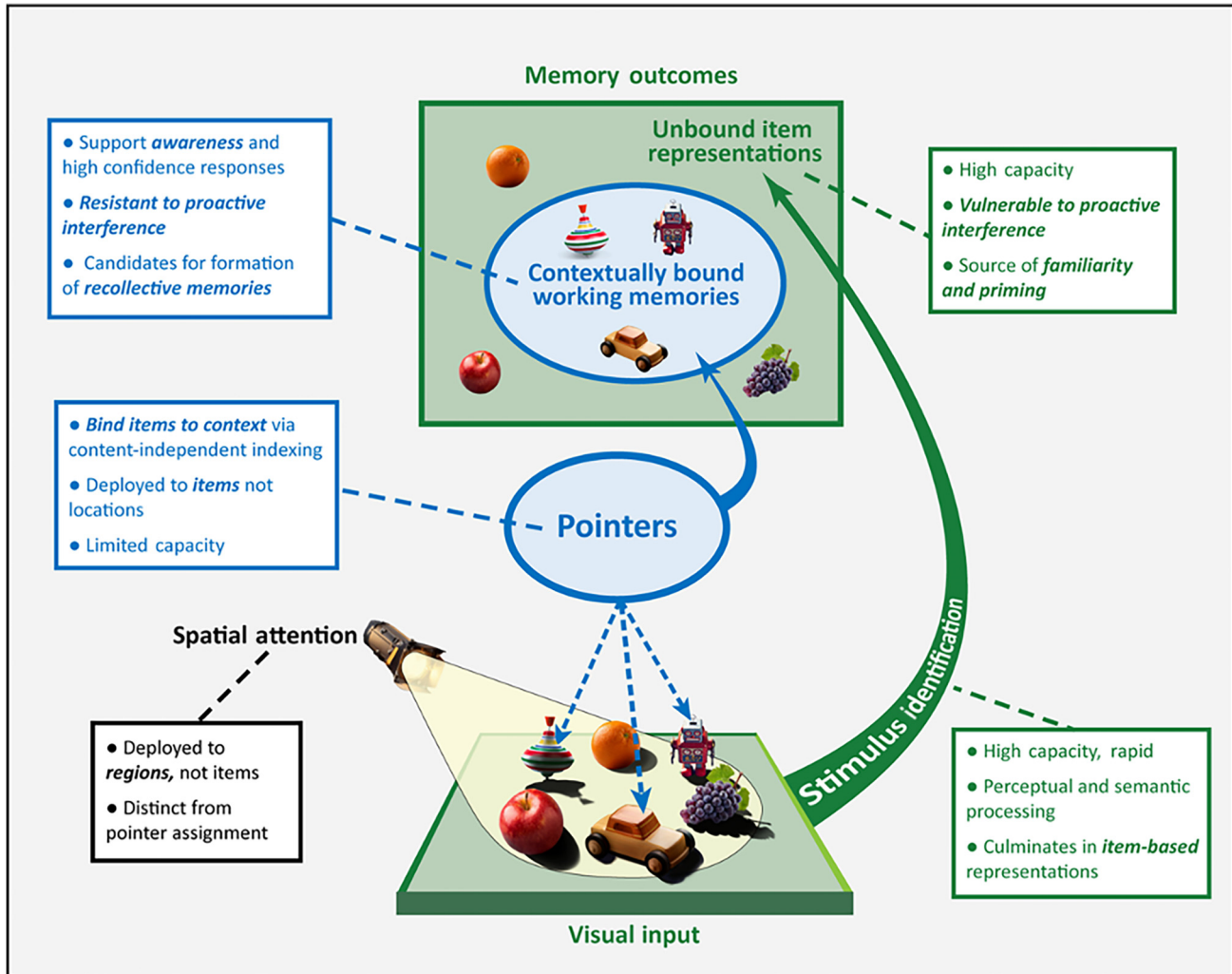
Neural evidence for content-independent pointers

We argue that pointers support an item-based indexing operation that is independent from the maintenance of featural content. Thus, the neural signature of pointer deployment should remain stable despite strong variations in the types of visual features stored or in the number of features contained within the stored items. Indeed, multiple studies have documented neural activity that tracks the number of individuated items stored in WM, independent of the featural content of those items [18,19,50–56]. For example, EEG recordings show that storing visual items in WM yields a sustained negative slow wave in posterior electrodes that increases in amplitude with each additional item stored [50]. This contralateral delay activity (CDA) plateaus at a set size of about three items, in line with behavioral estimates of WM capacity for visual items [21–24,57], and the shape of the CDA function across set sizes is a robust predictor of individual differences in WM capacity [58]. Importantly, the link between CDA amplitude and set size appears to be independent of the number of features stored from each object [59], suggesting that this neural signal tracks the number of individuated items stored rather than the content of those items.

More recently, we have shown that multivariate analysis of EEG activity enables robust decoding of the number of items stored in WM, even at the single-trial level [19,20,55,60]. In line with CDA findings, this multivariate signature of WM load tracks the number of items stored, independent of the content of those memories. For example, neural signatures of WM load generalize across memory conditions that vary in terms of both the type (e.g., color, orientation, motion coherence) and number (one versus two) of relevant features contained in each item [20,55]. Moreover, when multiple items are perceptually grouped, the decoded load matches the number of perceived items, not the number of attended positions or elements [55,61]. At the same time, these studies also found clear evidence for feature-specific neural activity that tracked the features stored in WM [20,55], verifying that observers selectively stored different features across conditions. But, critically, these feature-specific responses explain distinct variance in ongoing EEG activity compared with the content-independent signals that track pointer deployment [19,20]. Thus, we are not denying the role of stimulus-specific activity in WM storage. Instead, we are highlighting the existence of a separate class of neural signals that tracks item-based load but not content.

fMRI studies have also identified content-independent load signals. For example, past work has shown that activity in the inferior parietal cortex tracks the number of items stored in visual WM, regardless of the complexity of the stored items, while activity in distinct brain regions is sensitive to the total information load within each item [18]. Findings such as this motivated a ‘neural object file’ account in which objects are selected and individuated during an initial stage of processing, followed by processing of the precise identity of the selected objects [9]. Our proposal embraces a similar distinction between the selection of discrete items into WM and parallel processing of the features contained within those items, but there are also important points of divergence. Although the former account portrays object selection and individuation as a necessary precursor to full stimulus identification, we argue – motivated by the late selection evidence reviewed earlier – that full stimulus identification can be completed for items that are never encoded into WM. Thus, content-independent load activity does not necessarily precede stimulus identification, and indeed load activity may not be generated at all for items that are identified and deemed irrelevant.

Our proposal (Figure 1) is that these load signals reflect the deployment of content-independent pointers that enable the spatiotemporal tracking of the selected items. Item-based pointer assignment explains why these signals are indifferent to changes in both the type and number of visual features contained within the stored items. Our emphasis on spatiotemporal indexing falls in line with WM models that identify binding as a key limiting factor for storage in WM [6,7,62,63]. Here it is useful to distinguish between two types of binding operations. On the one hand, the stimulus identification typically requires binding of multiple distinct features (e.g., the color and shape of an object) into an item-based representation. On the other hand, it is also essential to bind those item-based representations to the surrounding event context. Our hypothesis is that pointer deployment is critical for the latter form of binding, but not for the binding of an object’s constituent features. In line with this, the late selection evidence described above has shown that multifeatured items can be fully identified without being encoded into WM, even though stimulus identification requires the binding of an object’s constituent features. Thus, we argue that pointer deployment is required specifically for the contextual binding of tokens, not the construction of multifeatured item representations. Moreover, we hypothesize that pointer assignment is a necessary step for the encoding of items into LTM (Box 3). To conclude, we argue that the assignment of content-independent pointers – combined with enhanced behavioral accessibility and the parallel encoding of stimulus-specific neural representations – may provide a productive view of the necessary and sufficient conditions for storage in visual WM.



Trends In Cognitive Sciences

Figure 1. The role of pointers in the formation of bound representations in working memory. The task is to store the toys and ignore the fruits. Spatial attention is thus deployed to all candidate objects to facilitate discrimination of targets and distractors. Rapid stimulus identification processes produce item-based representations – for both target and distractors – that will evoke priming and familiarity effects. Pointers are assigned to target items, enabling the formation of bound representations in working memory. These contextually bound representations support awareness and high confidence reports and may be required for the formation of recollective long-term memories.

Pointers may explain chunk-based capacity limits in WM

There has been persistent debate about the ‘units of storage’ in WM. On the one hand, WM capacity could be ‘chunk-based’ such that storage is limited by the total number of discrete entities stored [28,64]. On the other hand, WM capacity could be limited by the number of features or the total ‘information load’ stored in WM [17,65]. In this section, we argue that limits on pointer deployment may explain the evidence for chunk-based capacity limits in WM.

Behavioral evidence for chunk-based capacity

There are various demonstrations of chunk-based limits on WM performance. For example, WM performance is far better for letters that are part of meaningful words than for equal number of random letters, and WM performance with random words is equivalent to WM performance

Box 3. Pointers provide a gateway to LTM

Assigning a pointer to an object during encoding may provide a distinctive tag that can be used to increase the probability of later access in both WM and LTM tasks. In this way, pointer-assigned objects could support high-threshold memory representations because the index itself provides source information that may aid retrieval. However, past demonstrations that encoding into WM can be determined after full perceptual analysis [44,104] highlight the point that items may be fully identified without having a pointer assigned. In these instances, stimulus-specific activity may still be sufficient to form some residual memory of items that do not enter WM. However, the lack of a unique spatiotemporal index would mean that those representations would be highly susceptible to proactive and retroactive interference from other similar items and would primarily contribute to mnemonic judgments that could be driven by familiarity alone. Given that PI is based on confusions between items encountered at different points in time, we propose that binding an item to its spatiotemporal context may provide the basis for resistance to PI in WM. Thus, although there is some debate on whether WM is immune to PI [105], our proposal that WM storage depends on pointer assignment aligns with repeated observations of minimal PI on representations held in WM [91,106,107].

The control process of pointer assignment may help illuminate the overlap between the encoding and retrieval mechanisms for WM and LTM. Historically, these two memory systems were proposed to operate sequentially, with WM maintenance being the exclusive gateway for encoding into LTM [108]. Our view is more consistent with dual encoding models of memory that propose that WM and LTM encode information concurrently [10,109,110]. Here, we are specifying that the common encoding process between the two memory systems is the pointer allocation process. From this view, when presented with a display of many objects, a limited subset of items can have a pointer allocated. These items will be available for aiding immediate judgments in WM tasks and are more likely to be subsequently accessible from LTM. Items that receive no pointer assignment, by contrast, may not be explicitly accessible but may still impact behavioral performance if stimulus identification was nevertheless completed. For example, letters that cannot be reported due to the attentional blink nonetheless yielded reliable priming of subsequent responses, even though they were apparently denied access to WM [111]. One possibility is that the same representations that enable priming – but not robust conscious access – underlie various ‘familiarity’ effects that have been argued to be distinct from recollection-based memory phenomena. This view aligns with a recently proposed model in which memory judgments in WM are made on the basis of a mixture of high-threshold representations and a general factor of memory strength [10].

with an equal number of prelearned word pairs [25]. Within the visual domain, a key study argued that storage in visual WM was limited by the number of objects stored rather than by the total number of features contained within those objects [57]. Although subsequent research has made it clear that performance declines when additional features are added to each item [23,63,64], it is also clear that the number of individuated objects – independent of the number of features within each object – is a clear limiting factor for WM performance [23,66]. Likewise, although some studies had suggested that increasing stimulus complexity reduces the number of items that can be stored [65,67], follow-up studies showed that increased memory errors with complex items are due to confusions between targets and lures rather than a failure to store the targets in WM [3,68,69]. When those confusions are minimized, behavioral data reveal a common capacity limit for complex and simple objects [3,68,69]. Thus, while storage of all features is not guaranteed, item-based limits on WM performance are well established.

The link between late selection and chunk-based capacity limits

We argue that chunk-based capacity limits fall in line with the evidence for late selection into WM. Late selection, by definition, allows complete stimulus identification – and all of the low-level binding and LTM retrieval that this entails – prior to encoding into memory. Thus, if pointers are assigned to chunks, then the perceptual processing and memory retrieval that are required to recognize a chunk must occur prior to pointer assignment. Accordingly, words such as ‘cat’ evoke an EEG load signature commensurate with a single colored square (instead of three individuated letters), despite the fact that changing any one of the word’s three letters can create an entirely distinct word [70]. Thus, each of the three letters could be independently identified without requiring a separate pointer. Likewise, visual objects can be understood as a conjunction of features that have been chunked into a single item by virtue of Gestalt cues such as proximity and collinearity. In this way, postperceptual assignment of a limited number of pointers can explain why visual WM capacity is limited by the number of individuated objects that are perceived,

regardless of the number of features contained within those objects. Pointers support the contextual binding of recognized items, not features.

Here it is critical to consider how to define a 'chunk.' We prefer a hierarchical model of how a scene is parsed into items or chunks. For example, while a single word evokes a neural load signature commensurate with a single colored square, three random letters evoke the neural load associated with three colored squares [70]. In this case, associative learning enables a skilled reader to encode a word as a single chunk. However, this depends on the perspective that is required by the task at hand. If the observer were required to report the constituent letters of this word, it would require active retrieval of the letters from LTM [71–73] and the deployment of new pointers to each retrieved item. Thus, the observer's goals and library of preexisting associations will strongly shape how a scene is parsed into individuated chunks or items.

Given that pointers are deployed in an item-based fashion, a pointer limit would constrain the number of items that could be contextually bound, regardless of the content or complexity of those items. Why would WM performance be dependent on contextual binding? Typical WM tasks employ designs that include high levels of proactive interference (PI) due to the repeated use of a small set of possible stimuli. Thus, it is critical for the observer to respond to a test display only on the basis of information from the current trial. Otherwise, PI from other trials will undermine performance. Accordingly, it is known that using trial-unique stimuli that minimize PI yields apparent capacity limits that are far higher than in typical WM procedures [74], most likely because of contributions from familiarity-based representations in LTM.

This focus on contextual binding highlights an important difference between the pointer construct and the 'slots' metaphor that has often been used to illustrate the discrete character of WM capacity. While the concept of a 'slot' may be effective at illustrating how maximum storage limits are disconnected from the contents of each item, this metaphor lumps together the processes of contextual binding and the representation of stimulus-specific features. That is, the slots metaphor implies that slots contain the stored content rather than binding that content to the current context. By contrast, the pointer metaphor illustrates an abstract indexing operation that can be more readily distinguished from the maintenance of each item's featural content.

Finally, past work examining individual differences in visual WM has revealed a dissociation between the number of items an individual can hold in WM and the precision or fidelity of the stored representations [3,68]. Likewise, while perceptual expertise robustly enhances the precision of visual WM, it does not impact the number of representations that can be simultaneously stored [69]. These studies motivate a taxonomy of WM ability that distinguishes between the ability to maintain precise representations and the ability to control which representations are given access to WM. This taxonomy aligns with our working hypothesis that a limited number of content-independent pointers may determine maximum capacity limits in WM, while separate processes determine the fidelity of those memories. From this view, individual differences in apparent WM capacity may instead reflect how well one can control the assignment of pointers to the items in a display rather than reflecting differences in maximal storage space. This is well in line with current models that propose that individual differences in WM capacity reflect differences in attention control [75,76]. For example, individuals with low WM are known to have difficulty in preventing task-irrelevant items from being unnecessarily stored in WM [77–79]. From the current perspective, this unnecessary storage may be the consequence of errors in assigning pointers to the task-irrelevant items along with the task-relevant items. Thus, the efficiency of control over pointer deployment may be a fundamental driver of individual differences in WM capacity.

Concluding remarks

Cognitive neuroscience has made great progress in studying how the brain generates the stimulus-specific neural patterns that represent content in WM. Nevertheless, we argue that it is important to complement those efforts with a stronger focus on the binding of item representations to the current context [6,7,63]. In line with theories of dynamic visual cognition [11,12], we argue that items are contextually bound via the assignment of content-independent pointers, while parallel processes maintain the featural details associated with those items. This proposed separation between contextual binding and the maintenance of content resonates with other models that have asserted fundamental divisions between the stimulus-specific representations that determine the content of one's thoughts and abstract control processes that enable the manipulation and gating of those thoughts. For example, recent work has offered a 'spatial computing' model in which control processes that support the manipulation of WM representations are implemented independently of the neural activity that represents the details of the stored items themselves [80]. Likewise, others have argued that distinct neural pathways can be identified that encode the abstract structure of items and events on the one hand and pathways that encode the specific content of the encoded items on the other hand [81], suggesting that this kind of architecture may be necessary to apply learned rules to novel content. Thus, while further work is needed to directly examine whether spatiotemporal tracking is the core computation that limits storage in visual WM (see also [Outstanding questions](#)), we are optimistic about the broad utility of a distinction between the control processes that gate memory and the operations that maintain stimulus-specific content.

Acknowledgments

This work was supported by National Institute of Mental Health Grant No. RO1MH087214 and Office of Naval Research Grant No. N00014-12-1-0972.

Declaration of interests

The authors have no interests to declare.

References

- Cowan, N. *et al.* (2005) On the capacity of attention: its estimation and its role in working memory and cognitive aptitudes. *Cognit. Psychol.* 51, 42–100
- Engle, R.W. *et al.* (1999) Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *J. Exp. Psychol. Gen.* 128, 309–331
- Fukuda, K. *et al.* (2010) Quantity, not quality: the relationship between fluid intelligence and working memory capacity. *Psychon. Bull. Rev.* 17, 673–679
- Unsworth, N. *et al.* (2014) Working memory and fluid intelligence: capacity, attention control, and secondary memory retrieval. *Cognit. Psychol.* 71, 1–26
- Kanwisher, N.G. (1987) Repetition blindness: type recognition without token individuation. *Cognition* 27, 117–143
- Swan, G. and Wyble, B. (2014) The binding pool: a model of shared neural resources for distinct items in visual working memory. *Atten. Percept. Psychophys.* 76, 2136–2157
- Hedayati, S. *et al.* (2022) A model of working memory for latent representations. *Nat. Hum. Behav.* 6, 709–719
- Oberauer, K. (2019) Working memory capacity limits memory for bindings. *J. Cogn.* 2, 40
- Xu, Y. and Chun, M.M. (2009) Selecting and perceiving multiple visual objects. *Trends Cogn. Sci.* 13, 167–174
- Yonelinas, A.P. (2024) The role of recollection and familiarity in visual working memory: a mixture of threshold and signal detection processes. *Psychol. Rev.* 131, 321–348
- Pylyshyn, Z.W. (2009) Perception, representation, and the world: the FINST that binds. In *Computation, Cognition, and Pylyshyn* (Dedrick, D. and Trick, L., eds), pp. 3–48, MIT Press
- Kahneman, D. *et al.* (1992) The reviewing of object files: object-specific integration of information. *Cognit. Psychol.* 24, 175–219
- Fuster, J.M. and Alexander, G.E. (1971) Neuron activity related to short-term memory. *Science* 173, 652–654
- Postle, B.R. (2006) Working memory as an emergent property of the mind and brain. *Neuroscience* 139, 23–38
- Bays, P.M. (2014) Noise in neural populations accounts for errors in working memory. *J. Neurosci.* 34, 3632–3645
- Schurigin, M.W. *et al.* (2020) Psychophysical scaling reveals a unified theory of visual memory strength. *Nat. Hum. Behav.* 4, 1156–1172
- Ma, W.J. *et al.* (2014) Changing concepts of working memory. *Nat. Neurosci.* 17, 347–356
- Xu, Y. and Chun, M.M. (2006) Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature* 440, 91–95
- Jones, H.M. *et al.* (2024) Electroencephalogram decoding reveals distinct processes for directing spatial attention and encoding into working memory. *Psychol. Sci.* 35, 1108–1138
- Jones, H. *et al.* (2024) Cortically disparate visual features evoke content-independent load signals during storage in working memory. *J. Neurosci.* 44, e0448242024
- Zhang, W. and Luck, S.J. (2008) Discrete fixed-resolution representations in visual working memory. *Nature* 453, 233–235
- Adam, K.C.S. *et al.* (2017) Clear evidence for item limits in visual working memory. *Cognit. Psychol.* 97, 79–97
- Ngiam, W.X.Q. *et al.* (2024) Object-based encoding constrains storage in visual working memory. *J. Exp. Psychol. Gen.* 153, 86–101

Outstanding questions

Are there distinct pointer operations for distinct sensory modalities? On the one hand, various demonstrations of multisensory integration suggest that common indices can be deployed to features from different sensory modalities. On the other hand, multiple studies have found minimal costs of concurrently storing items from different sensory modalities, suggesting the operation of distinct memory systems.

Is the same pointer system deployed when information is retrieved from LTM? Most models of retrieval from LTM presume that representations are retrieved into WM. Does retrieval from LTM recruit the same pointer system that mediates storage in WM?

What is the functional relationship between spatial attention and pointer assignment? We have reviewed evidence that they reflect distinct aspects of cognitive control, but important questions remain about their interactions. For example, can a pointer be assigned to a spatially unattended item?

What limits the maximum number of pointers that can be assigned at once? We have proposed that pointers support spatiotemporal indexing that binds items to context. Although we have not proposed a specific biophysical explanation for why this process may be limited, specifying binding as a limiting factor provides a more concrete and testable target for future investigations of why capacity in WM is limited.

24. Ngiam, W.X.Q. *et al.* (2023) Distinguishing guesses from fuzzy memories: further evidence for item limits in visual working memory. *Atten. Percept. Psychophys.* 85, 1695–1709
25. Chen, Z. and Cowan, N. (2009) Core verbal working-memory capacity: the limit in words retained without covert articulation. *Q. J. Exp. Psychol.* 62, 1420–1429
26. Miller, G.A. (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63, 81–97
27. Cowan, N. (2017) The many faces of working memory and short-term storage. *Psychon. Bull. Rev.* 24, 1158–1170
28. Fukuda, K. *et al.* (2010) Discrete capacity limits in visual working memory. *Curr. Opin. Neurobiol.* 20, 177–182
29. Ericsson, K.A. and Kintsch, W. (1995) Long-term working memory. *Psychol. Rev.* 102, 211–245
30. Fuster, J.M. and Jervey, J.P. (1981) Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science* 212, 952–955
31. Funahashi, S. *et al.* (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* 61, 331–349
32. Chase, W.G. and Ericsson, K.A. (2013) Skilled memory. In *Cognitive Skills and their Acquisition*, pp. 141–189, Psychology Press
33. Carlisle, N.B. *et al.* (2011) Attentional templates in visual working memory. *J. Neurosci.* 31, 9315–9322
34. Wolfe, J.M. (2020) Visual search: how do we find what we are looking for? *Annu. Rev. Vis. Sci.* 6, 539–562
35. Woodman, G.F. and Chun, M.M. (2006) The role of working memory and long-term memory in visual search. *Vis. Cogn.* 14, 808–830
36. Serences, J.T. *et al.* (2009) Stimulus-specific delay activity in human primary visual cortex. *Psychol. Sci.* 20, 207–214
37. Harrison, S.A. and Tong, F. (2009) Decoding reveals the contents of visual working memory in early visual areas. *Nature* 458, 632–635
38. Foster, J.J. *et al.* (2016) The topography of alpha-band activity tracks the content of spatial working memory. *J. Neurophysiol.* 115, 168–177
39. Kamitani, Y. and Tong, F. (2005) Decoding the visual and subjective contents of the human brain. *Nat. Neurosci.* 8, 679–685
40. Moray, N. (1959) Attention in dichotic listening: affective cues and the influence of instructions. *Q. J. Exp. Psychol.* 11, 56–60
41. Duncan, J. (1980) The locus of interference in the perception of simultaneous stimuli. *Psychol. Rev.* 87, 272–300
42. Potter, M.C. (1975) Meaning in visual search. *Science* 187, 965–966
43. Shiffrin, R.M. and Schneider, W. (1977) Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychol. Rev.* 84, 127–190
44. Potter, M.C. (1976) Short-term conceptual memory for pictures. *J. Exp. Psychol. [Hum. Learn.]* 2, 509–522
45. Bellet, J. *et al.* (2022) Decoding rapidly presented visual stimuli from prefrontal ensembles without report nor post-perceptual processing. *Neurosci. Conscious.* 2022, niac005
46. Grootswagers, T. *et al.* (2019) The representational dynamics of visual objects in rapid serial visual processing streams. *NeuroImage* 188, 668–679
47. Mohsenzadeh, Y. *et al.* (2019) The perceptual neural trace of memorable unseen scenes. *Sci. Rep.* 9, 6033
48. Hakim, N. *et al.* (2019) Dissecting the neural focus of attention reveals distinct processes for spatial attention and object-based storage in visual working memory. *Psychol. Sci.* 30, 526–540
49. Chun, M.M. (1997) Types and tokens in visual processing: a double dissociation between the attentional blink and repetition blindness. *J. Exp. Psychol. Hum. Percept. Perform.* 23, 738–755
50. Vogel, E.K. and Machizawa, M.G. (2004) Neural activity predicts individual differences in visual working memory capacity. *Nature* 428, 748–751
51. Todd, J.J. and Marois, R. (2004) Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* 428, 751–754
52. Cowan, N. *et al.* (2011) A neural region of abstract working memory. *J. Cogn. Neurosci.* 23, 2852–2863
53. Majerus, S. *et al.* (2016) Cross-modal decoding of neural patterns associated with working memory: evidence for attention-based accounts of working memory. *Cereb. Cortex* 26, 166–179
54. Balaban, H. *et al.* (2019) Neural evidence for an object-based pointer system underlying working memory. *Cortex* 119, 362–372
55. Thyer, W. *et al.* (2022) Storage in visual working memory recruits a content-independent pointer system. *Psychol. Sci.* 33, 1680–1694
56. Balaban, H. *et al.* (2023) Dissociable online integration processes in visual working memory. *Cereb. Cortex* 33, 11420–11430
57. Luck, S.J. and Vogel, E.K. (1997) The capacity of visual working memory for features and conjunctions. *Nature* 390, 279–281
58. Luria, R. *et al.* (2016) The contralateral delay activity as a neural measure of visual working memory. *Neurosci. Biobehav. Rev.* 62, 100–108
59. Woodman, G.F. and Vogel, E.K. (2008) Selective storage and maintenance of an object's features in visual working memory. *Psychon. Bull. Rev.* 15, 223–229
60. Adam, K.C.S. *et al.* (2020) Multivariate analysis reveals a generalizable human electrophysiological signature of working memory load. *Psychophysiology* 57, e13691
61. Diaz, G.K. *et al.* (2021) Perceptual grouping reveals distinct roles for sustained slow wave activity and alpha oscillations in working memory. *J. Cogn. Neurosci.* 33, 1354–1364
62. Baddeley, A.D. *et al.* (2017) Binding in visual working memory: the role of the episodic buffer. In *Exploring Working Memory*, pp. 312–331, Routledge
63. Oberauer, K. and Lin, H.-Y. (2017) An interference model of visual working memory. *Psychol. Rev.* 124, 21–59
64. Cowan, N. (2001) The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–114
65. Alvarez, G.A. and Cavanagh, P. (2004) The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychol. Sci.* 15, 106–111
66. Hardman, K.O. and Cowan, N. (2015) Remembering complex objects in visual working memory: do capacity limits restrict objects or features? *J. Exp. Psychol. Learn. Mem. Cogn.* 41, 325–347
67. Eng, H.Y. *et al.* (2005) Visual working memory for simple and complex visual stimuli. *Psychon. Bull. Rev.* 12, 1127–1133
68. Awh, E. *et al.* (2007) Visual working memory represents a fixed number of items regardless of complexity. *Psychol. Sci.* 18, 622–628
69. Scolar, M. *et al.* (2008) Perceptual expertise enhances the resolution but not the number of representations in working memory. *Psychon. Bull. Rev.* 15, 215–222
70. Rajsic, J. *et al.* (2019) Contralateral delay activity tracks the storage of visually presented letters and words. *Psychophysiology* 56, e13282
71. Huang, L. and Awh, E. (2018) Chunking in working memory via content-free labels. *Sci. Rep.* 8, 23
72. Ngiam, W.X.Q. *et al.* (2019) 'Memory compression' effects in visual working memory are contingent on explicit long-term memory. *J. Exp. Psychol. Gen.* 148, 1373–1385
73. Musfeld, P. *et al.* (2023) Repetition learning is neither a continuous nor an implicit process. *Proc. Natl. Acad. Sci. U. S. A.* 120, e2218042120
74. Endress, A.D. and Potter, M.C. (2014) Large capacity temporary visual memory. *J. Exp. Psychol. Gen.* 143, 548–565
75. Engle, R.W. (2002) Working memory capacity as executive attention. *Curr. Dir. Psychol. Sci.* 11, 19–23
76. Martin, J.D. *et al.* (2021) The visual arrays task: visual storage capacity or attention control? *J. Exp. Psychol. Gen.* 150, 2525–2551
77. Vogel, E.K. *et al.* (2005) Neural measures reveal individual differences in controlling access to working memory. *Nature* 438, 500–503
78. McNab, F. and Klingberg, T. (2008) Prefrontal cortex and basal ganglia control access to working memory. *Nat. Neurosci.* 11, 103–107

79. Fukuda, K. and Vogel, E.K. (2009) Human variation in overriding attentional capture. *J. Neurosci.* 29, 8726–8733
80. Lundqvist, M. *et al.* (2023) Working memory control dynamics follow principles of spatial computing. *Nat. Commun.* 14, 1429
81. O'Reilly, R.C. *et al.* (2022) The structure of systematicity in the brain. *Curr. Dir. Psychol. Sci.* 31, 124–130
82. Todd, J.J. and Marois, R. (2005) Posterior parietal cortex activity predicts individual differences in visual short-term memory capacity. *Cogn. Affect. Behav. Neurosci.* 5, 144–155
83. Lewis-Peacock, J.A. *et al.* (2012) Neural evidence for a distinction between short-term memory and the focus of attention. *J. Cogn. Neurosci.* 24, 61–79
84. Hakim, N. *et al.* (2020) Perturbing neural representations of working memory with task-irrelevant interruption. *J. Cogn. Neurosci.* 32, 558–569
85. Christophel, T.B. *et al.* (2018) Cortical specialization for attended versus unattended working memory. *Nat. Neurosci.* 21, 494–496
86. Stokes, M.G. (2015) 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. *Trends Cogn. Sci.* 19, 394–405
87. Lundqvist, M. *et al.* (2018) Working memory: delay activity, yes! Persistent activity? Maybe not. *J. Neurosci.* 38, 7013–7019
88. Rhodes, S. and Cowan, N. (2018) Attention in working memory: attention is needed but it yearns to be free. *Ann. N. Y. Acad. Sci.* 1424, 52–63
89. Beukers, A.O. *et al.* (2021) Is activity silent working memory simply episodic memory? *Trends Cogn. Sci.* 25, 284–293
90. Foster, J.J. *et al.* (2024) 13 Working memory as persistent neural activity. In *The Oxford Handbook of Human Memory, Two Volume Pack: Foundations and Applications, online edn* (Kahana, M.J. and Wagner, A.D., eds), pp. 353–370, Academic, Oxford
91. Oberauer, K. and Awh, E. (2022) Is there an activity-silent working memory? *J. Cogn. Neurosci.* 34, 2360–2374
92. Jolicoeur, P. and Dell'Acqua, R. (1998) The demonstration of short-term consolidation. *Cognit. Psychol.* 36, 138–202
93. Ricker, T.J. and Hardman, K.O. (2017) The nature of short-term consolidation in visual working memory. *J. Exp. Psychol. Gen.* 146, 1551–1573
94. Cowan, N. and Chen, Z. (2008) How chunks form in long-term memory and affect short-term memory limits. In *Interactions Between Short-Term and Long-Term Memory in the Verbal Domain*, pp. 98–119, Psychology Press
95. Treisman, A.M. and Gelade, G. (1980) A feature-integration theory of attention. *Cognit. Psychol.* 12, 97–136
96. Reynolds, J.H. and Desimone, R. (1999) The role of neural mechanisms of attention in solving the binding problem. *Neuron* 24, 19–29
97. Robertson, L.C. (2003) Binding, spatial attention and perceptual awareness. *Nat. Rev. Neurosci.* 4, 93–102
98. Awh, E. and Jonides, J. (2001) Overlapping mechanisms of attention and spatial working memory. *Trends Cogn. Sci.* 5, 119–126
99. Awh, E. *et al.* (1998) Rehearsal in spatial working memory. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 780–790
100. Williams, M. *et al.* (2013) Visual-spatial attention aids the maintenance of object representations in visual working memory. *Mem. Cogn.* 41, 698–715
101. Fukuda, K. *et al.* (2015) α Power modulation and event-related slow wave provide dissociable correlates of visual working memory. *J. Neurosci.* 35, 14009–14016
102. Jha, A.P. (2002) Tracking the time-course of attentional involvement in spatial working memory: an event-related potential investigation. *Cogn. Brain Res.* 15, 61–69
103. Foster, J.J. *et al.* (2017) Alpha-band activity reveals spontaneous representations of spatial position in visual working memory. *Curr. Biol.* 27, 3216–3223.e6
104. Luck, S.J. *et al.* (1996) Word meanings can be accessed but not reported during the attentional blink. *Nature* 383, 616–618
105. Ralph, A. *et al.* (2011) Immunity to proactive interference is not a property of the focus of attention in working memory. *Mem. Cogn.* 39, 217–230
106. Cowan, N. *et al.* (2005) Capacity limits in list item recognition: evidence from proactive interference. *Memory* 13, 293–299
107. Oberauer, K. *et al.* (2017) The role of long-term memory in a test of visual working memory: proactive facilitation but no proactive interference. *J. Exp. Psychol. Learn. Mem. Cogn.* 43, 1–22
108. Atkinson, R.C. and Shiffrin, R.M. (1968) Human memory: a proposed system and its control processes. *Psychol. Learn. Motiv.* 2, 89–195
109. Yonelinas, A.P. (2002) The nature of recollection and familiarity: a review of 30 years of research. *J. Mem. Lang.* 46, 441–517
110. Rugg, M.D. and Yonelinas, A.P. (2003) Human recognition memory: a cognitive neuroscience perspective. *Trends Cogn. Sci.* 7, 313–319
111. Shapiro, K. *et al.* (1997) Priming from the attentional blink: a failure to extract visual tokens but not visual types. *Psychol. Sci.* 8, 95–100