

# Visual working memory capacity: from psychophysics and neurobiology to individual differences

Steven J. Luck<sup>1</sup> and Edward K. Vogel<sup>2</sup>

<sup>1</sup> Center for Mind and Brain and Department of Psychology, University of California, Davis, 267 Cousteau Place, Davis, CA 95618, USA

<sup>2</sup> Department of Psychology, University of Oregon, Straub Hall, Eugene, OR 97403, USA

**Visual working memory capacity is of great interest because it is strongly correlated with overall cognitive ability, can be understood at the level of neural circuits, and is easily measured. Recent studies have shown that capacity influences tasks ranging from saccade targeting to analogical reasoning. A debate has arisen over whether capacity is constrained by a limited number of discrete representations or by an infinitely divisible resource, but the empirical evidence and neural network models currently favor a discrete item limit. Capacity differs markedly across individuals and groups, and recent research indicates that some of these differences reflect true differences in storage capacity whereas others reflect variations in the ability to use memory capacity efficiently.**

## The rise of visual working memory

For many decades, the concept of working memory capacity has played a central role in large-scale theories of cognition [1,2] and in explaining individual differences in cognitive ability [3]. Originally, research on working memory was dominated by verbal paradigms, such as digit span tasks, which require participants to repeat back a series of digits, and complex span tasks, in which participants must alternate between two tasks that require processing of information and storing of items in memory [4]. However, the past 15-year period has seen an explosion of research on visual working memory (VWM, as defined in Box 1; see Glossary).

In this article, we review some of the key findings of this research, focusing on the cognitive and neural mechanisms of VWM capacity and on individual and group differences in VWM capacity. We begin by asking why vision needs a working memory system. We then discuss whether capacity is constrained by a limit on the number of discrete items that can be represented or by a limit on a resource that can be divided among large numbers of items. We then discuss how and why capacity varies among individuals and between groups. Finally, we discuss the neural mechanisms that may determine VWM capacity. Our overall perspective is that limits on VWM capacity reflect the broader

problem of maintaining multiple active representations in networks of interconnected neurons. This problem can be solved by maintaining a limited number of discrete representations, which then impacts almost every aspect of cognitive function.

## Why study visual working memory?

There are at least four major reasons for the explosion of research on VWM capacity. First, studies of change blindness in the 1990s (Figure 1A) provided striking examples of the limitations of VWM capacity in both the laboratory and the real world [5,6].

Second, the change detection paradigm (Figure 1B–D) was popularized to provide a means of studying the same basic phenomenon with more precisely controlled visual displays [7–9]. This paradigm made it possible to quantify VWM capacity and to link VWM to the enormous body of

## Glossary

**Cell assembly:** set of neurons that together represent a single item (e.g., a set of red-selective neurons and a set of vertical-selective neurons that together represent a red vertical bar).

**Complex span tasks:** tasks that assess working memory capacity by requiring subjects to switch back and forth between a memory encoding task and a processing task. In the operation span task, for example, subjects see a simple mathematics problem followed by a word that is to be stored in memory. In each trial, a sequence of these pairs is presented and subjects are asked to recall the words at the end of the trial. Memory span is quantified in terms of the number of words that can be recalled at the end of the trial.

**Contralateral delay activity:** sustained ERP response during the delay period of a visual working memory task that is observed over the hemisphere contralateral to the items being maintained in memory.

**Event-related potential:** ERPs are specific event-related brain responses that are embedded within the electroencephalogram (EEG). They arise from the summed postsynaptic potentials of many thousands of neurons and are conducted through the brain and skull to the scalp, where they can be recorded noninvasively via surface electrodes.

**Functional magnetic resonance imaging:** this technique takes advantage of the different magnetic properties of oxygenated and deoxygenated hemoglobin to localize changes in blood flow that are triggered by changes in neural activity.

**$K$ :** number of items stored in working memory on a given trial type.

**$K_{max}$ :** maximum number of items that a given individual can store in working memory. This is a measure of working memory capacity.

**$P_{mem}$ :** probability that a given item is present in memory at the time of a test.

**Set size:** number of items in an array.

**Single-unit recordings:** recordings of the action potentials of individual neurons via the tip of an electrode placed just outside the cell body of a neuron.

**Standard deviation:** measure of the spread of a distribution of values. In VWM research, standard deviation is used to quantify the distribution of memory errors and is inversely related to the precision of the memory representation.

**Visual working memory:** active maintenance of visual information to serve the needs of ongoing tasks.

Corresponding author: Luck, S.J. (sjluck@ucdavis.edu).

1364-6613/\$ – see front matter

© 2013 Elsevier Ltd. All rights reserved. <http://dx.doi.org/10.1016/j.tics.2013.06.006>

### Box 1. What is visual working memory?

The term working memory is used in many different ways and it is therefore important for researchers to define exactly what they mean when they use this term. In this article, we define VWM as the active maintenance of visual information to serve the needs of ongoing tasks. We are not suggesting that this is the only valid definition of VWM. Instead, this definition is intended to provide a clear statement of the scope of the memory system that we address in this review. We also believe that it reflects the typical usage of the term by researchers who come from a vision science perspective.

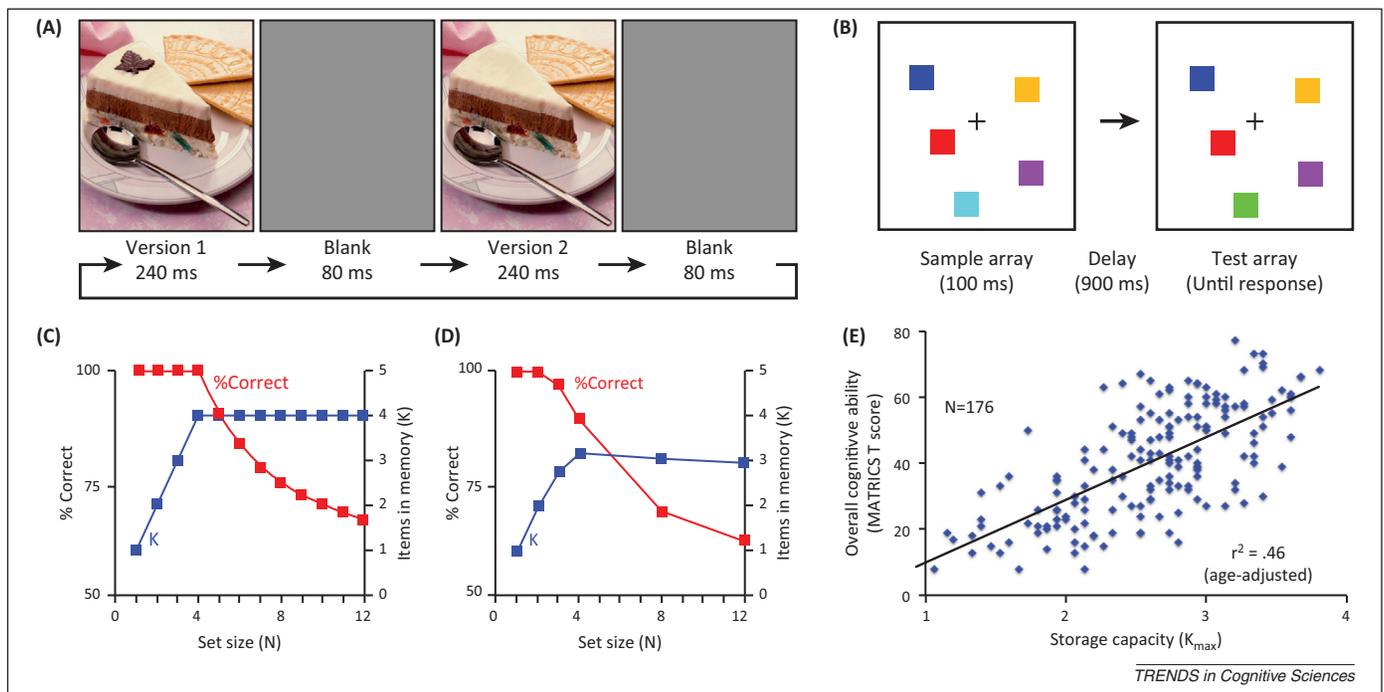
Our definition has three key components. First, to qualify as VWM, it is not sufficient that the information was acquired through the visual modality; the representation of the information must be visual in nature. If the observer stores a verbal or amodal conceptual representation of the sensory input, we no longer consider it to be a visual memory. Second, VWM is based on active maintenance. That is, a VWM representation is maintained by a change in sustained, energy-requiring neural activity rather than by a change in synaptic strength (which can be verified with physiological recordings). This distinguishes VWM representations from passively stored, longer-term memories. Third, the representations must be used in the service of broader cognitive tasks. This is the “*workin*” part of VWM.

research on vision [10]. Moreover, this task is so simple that even pigeons can do it [11].

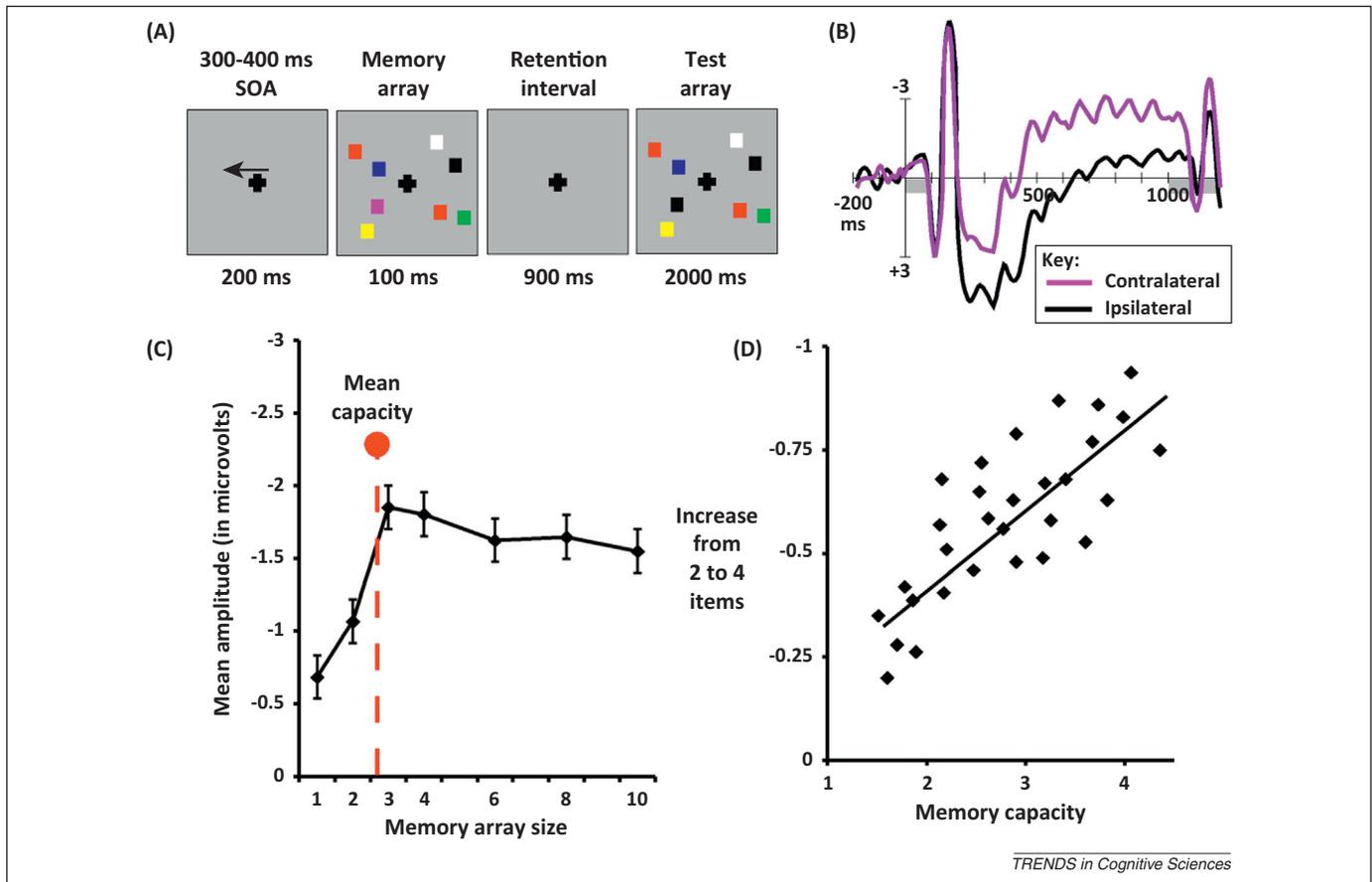
Third, estimates of VWM capacity have excellent psychometric properties when optimal methods are used [12,13]. For example, a 10-min change localization task yielded test-retest reliability of 0.77 for testing episodes separated by 1.5 years [13]. In addition, VWM capacity is highly correlated with measures of broad cognitive function, accounting for

43% of individual differences in a global fluid intelligence [14] and 46% of individual differences in overall performance on a broad battery of cognitive tasks (Figure 1E) [13]. These high correlations are particularly striking given that the change detection paradigm provides a relatively simple measure of VWM capacity, with little or no impact of long-term memory when canonical task parameters are used. Specifically, there is little or no effect of proactive interference [15] (but see [16,17]) or medial temporal lobe lesions [18] (but see [19]) with canonical parameters.

Finally, researchers have discovered neural correlates of VWM maintenance that are strongly correlated with individual differences in VWM capacity. In studies of VWM in non-human primates, neurons in several brain areas exhibit elevated firing rates and increased synchrony during the delay interval [20,21]. In human event-related potential (ERP) studies, an analogous sustained change in voltage is observed during the delay interval in change detection tasks [22,23]. This effect is called contralateral delay activity (CDA) because it is found in the hemisphere contralateral to a set of lateralized objects that are being remembered over a delay period (Figure 2A). CDA amplitude increases as the set size increases, reaching an asymptote at the capacity limit (typically three or four items). This is true both at the group level and the single-subject level, with very strong correlation between an individual’s behaviorally measured VWM capacity and that individual’s CDA asymptote point (Figure 2D). An analogous effect can be seen in functional magnetic resonance imaging



**Figure 1.** (A) Example of a change blindness task. Many cycles are required before an observer notices the difference between the two images. Reprinted from [10] by permission of Oxford University Press, USA. (B) Example of a change detection task [8]. A brief sample array is followed by a blank delay and then a test array. The test array is either identical to the sample array or differs in one feature of one of the objects, and the observer indicates whether a change is present. In the change localization variant, a change is always present and the subject indicates which item has changed [47,60,79]. (C) Hypothetical results for an observer with a capacity ( $K_{\max}$ ) of four items, assuming a slot model. Accuracy (% correct) is perfect when the set size ( $M$ ) is less than  $K_{\max}$  (assuming that changes in color are very large, when present). When  $N > K_{\max}$ , the changed item will be present in memory for  $N/K_{\max}$  trials, and subjects will fail to detect the change when the changed item is not in memory. Accuracy will therefore decrease systematically as  $N$  increases above  $K_{\max}$ . By taking into account guessing, it is possible to estimate the number of items that the observer must have had in memory ( $K$ ) for each set size [80–83]. (D) Data from an actual experiment with college student subjects [8]. (E) Scatter plot of the relationship between storage capacity ( $K_{\max}$ ) measured in a 10-min change localization task and a measure of broad cognitive function (the T score from the MATRICS battery) in a sample of subjects including both schizophrenia patients and matched controls [13]. The correlations were similar in both groups, justifying an aggregated analysis.



**Figure 2.** (A) Event-related potential (ERP) paradigm for recording contralateral delay activity (CDA) [22]. Subjects are instructed to remember the colors of the items on the side indicated by the arrow and report whether a color has changed on that side in the test array. (B) ERP waveforms from ipsilateral versus contralateral electrode sites relative to the side of the array that was encoded into memory. Time zero is the onset of the test array, and the CDA is the difference in voltage between the ipsilateral and contralateral waveforms during the delay period. Note the negative direction of the y-axis. (C) CDA amplitude as a function of the number of items on the side to be remembered, averaged over subjects. Note that CDA amplitude reaches an asymptote near the average working memory capacity limit. (D) Scatter plot for individual subjects, showing that individual differences in working memory capacity ( $K_{max}$ ) are correlated with differences in the CDA asymptote (quantified as the difference in CDA amplitude between set sizes of  $N=2$  and  $N=4$ ).

(fMRI) studies, where the signal in the intraparietal sulcus (IPS) during the delay period increases as the set size increases, reaching an asymptote at the individual subject's VWM capacity [24,25].

### The role of working memory in vision

Visual information is typically acquired during short periods of fixation (usually 200–500 ms in duration) separated by saccadic eye movements that temporarily suppress processing and then shift the retinal image. Some kind of memory is needed to bridge the temporal gaps and spatial shifts created by eye movements [26]. Seminal research by Irwin and colleagues [27,28] demonstrated that iconic memory does not survive eye movements but that VWM can be used to link the pre-saccade representation of an object at one retinal location with the post-saccade representation of that object in a different retinal location. More recent research has shown that the target of an upcoming eye movement is automatically stored in VWM, and after the eye movement this VWM representation is compared with the newly fixated object [29]. In addition, eye movements may be biased toward objects that match the current contents of VWM [30], and even the simplest saccades are faster if the saccade target matches the current contents of VWM [31].

VWM also plays a key role in higher-level visual tasks, demonstrating that it is truly a working memory. For example, when the target for a visual search task is cued in a trial-by-trial manner, the cue is stored in VWM, leading to a CDA in the interval between the cue and the search array [32]. In this situation, search performance is impaired if VWM is filled to capacity by a secondary object memory task [33]. However, after several trials of searching for the same target, the CDA disappears [32]. Moreover, search performance is no longer impaired by a concurrent VWM load when the target remains the same on trial after trial [34]. These results indicate that the 'search template' is transferred from VWM into a longer-term memory store when the target remains constant over several trials.

It is natural to assume that the sole purpose of working memory is to store items that are no longer present, but recent research indicates that the same system is also used to represent information that is currently visible. For example, both VWM capacity limits and neural indices of VWM activation have been observed in tasks in which the items remain visible throughout the trial, such as visual search [35,36] and multiple object tracking [37,38]. A recent study took this a step further and showed that when observers were asked to remember the colors of

items in a static array that remained visible for the entire trial, their VWM capacity was indistinguishable from trials in which the items disappeared during a 1-s retention period [39]. Moreover, the CDA was the same whether or not the stimuli were visible during the retention period, indicating that the same neural mechanisms are used to represent visual objects whether or not they are currently visible. Together, these findings suggest that VWM may not really be a memory system *per se*, but may instead be a general-purpose visual representation system that can, when necessary, maintain information over short delays.

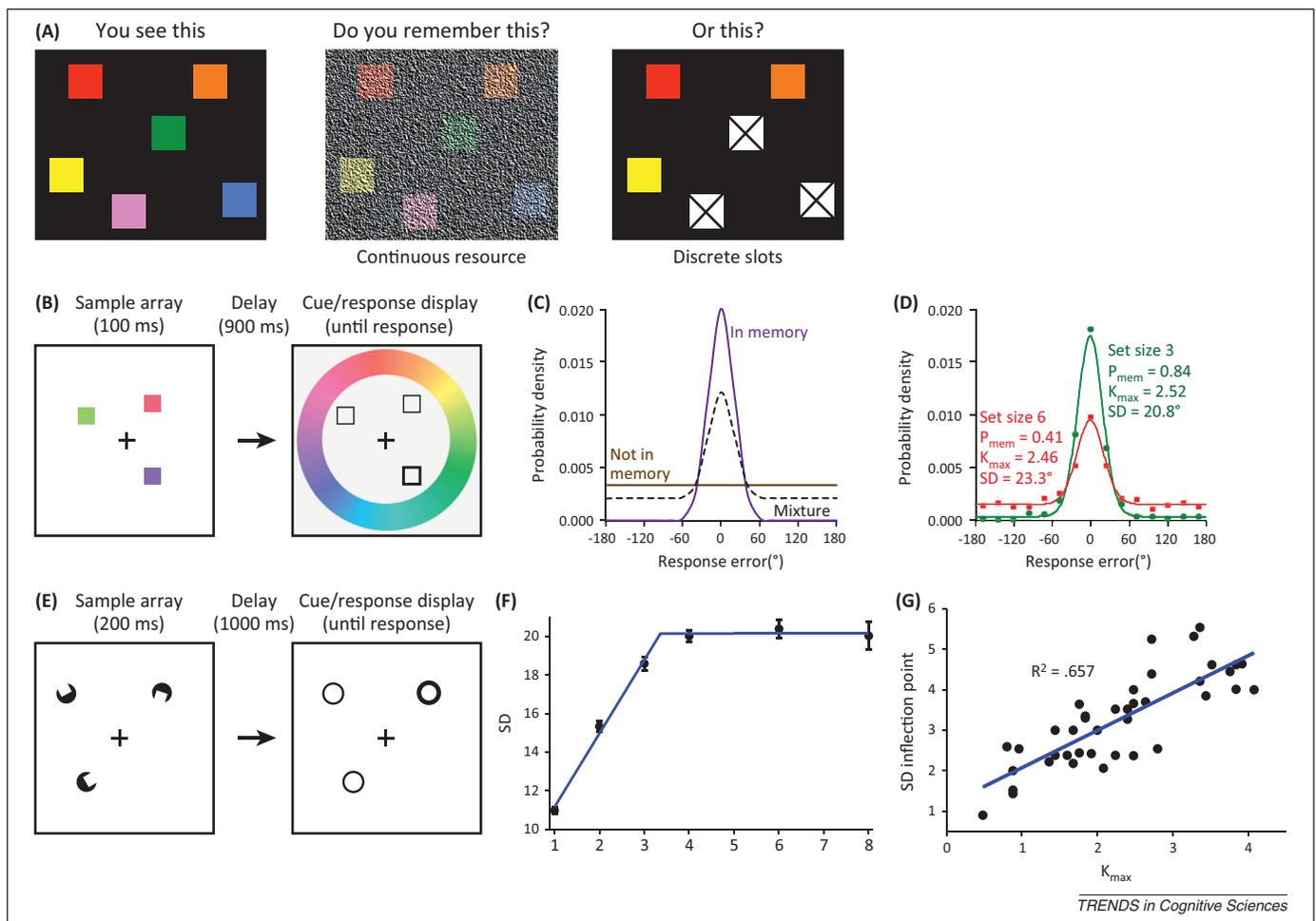
### The nature of VWM capacity limitations: slots versus resources

When memory for simple, highly discriminable colored squares is tested, the typical college student has a capacity of only three to four objects' worth of information [40]. What is the nature of this limit? Object complexity plays a clear role: task performance is less accurate for complex objects than for simple objects in most cases [41,42].

However, complexity is not a very well-defined term [10,43], and it is much easier to understand capacity limits for simple, unidimensional features.

Two main classes of theories of VWM capacity have been proposed, a discrete slots class and a continuous resource class [10,44–46]. Figure 3A shows how a display of six colored squares would be represented in VWM according to these two theory classes. Slot-based theories assume that a limited number of items,  $K_{\max}$ , can be stored in VWM; if the number of items in the sensory input is greater than  $K_{\max}$ , then  $K_{\max}$  of the items are stored in VWM and no information about the other items is stored in VWM. Note, however, that internal and external sources of variability will cause each representation to be imperfect and may cause  $K_{\max}$  to vary from trial to trial.

Resource-based theories assume that VWM capacity is a flexibly divisible resource that can be spread among all the items in the display, but with fewer resources per item and therefore reduced precision as the set size increases. These theories can also be framed in terms of an increase neural noise as the set size increases.



**Figure 3.** (A) Essence of the continuous resource and discrete slot model classes. (B) Example of a continuous report task with color stimuli. The cue (thicker box) indicates which item should be reported by clicking on the color wheel. (C) Hypothetical distribution of response errors (difference between actual color and reported color) according to the slot model [48]. If the cued item is present in memory (violet line), the errors will be normally distributed around the correct value (the Von Mises distribution is used for circular dimensions such as hue). If the cued item is not remembered (brown line), errors will be random (a uniform distribution). The actual data consist of a weighted sum of these two distributions (black line). (D) Data observed for set sizes of  $N=3$  and  $N=6$ , and estimates of the parameters of the underlying distributions [48]. (E) Continuous report task for orientation [49]. The sample array contains circles with gaps; when the test display appears, the subject reports the orientation remembered for the gap in the item that is cued by the thicker circle; the orientation is reported by clicking on the corresponding location on the cue circle. (F) Standard deviation (SD) of the distribution of response errors in the task shown in (E) as a function of set size. The group data fit well to a function that increases linearly, has an inflection point at the average  $K_{\max}$ , and is then flat. (G) Inflection point as a function of  $K_{\max}$  for individual subjects, showing that the point at which the SD reaches an asymptote for a given subject is predicted by that subject's visual working memory capacity.

### Evidence of discrete slots from continuous report experiments

The most obvious way to distinguish between these classes of theories is to determine whether VWM representations become less precise as the set size increases, and this is most easily assessed with the continuous report paradigm shown in Figure 3B [47,48]. Each trial begins with a sample array, followed by a retention interval and then a test display. The test display includes a cue at one location, and the observer clicks on a color wheel to report the color remembered for the cued item. If the cued item is present in memory, the color reported should be close to the actual color (i.e., the magnitude of the error will be small), and the distribution of error magnitudes over many trials would be expected to be approximately normal (Figure 3C). The width of this distribution (quantified as the standard deviation, SD) is inversely related to the precision of the memory representation. The same method can be used for other feature dimensions, such as orientation [49] (but see [50]) and shape [48].

When the set size exceeds the number of items that can be stored in VWM ( $K_{\max}$ ), the cued item will not be present in memory on a subset of trials. When this happens, the observer will guess randomly, leading to a uniform distribution of errors (Figure 3C). Because the data from a set of trials may contain a mixture of in-memory and out-of-memory trials, the observed distribution of errors will consist of a mixture of a normal distribution (from in-memory trials) and a uniform distribution (from out-of-memory trials). This is equivalent to a normal distribution that has been shifted vertically by an amount that is related to the probability that the cued item was absent from memory. From the observed mixture, it is possible to derive two key VWM parameters:  $P_{\text{mem}}$ , the probability that the cued item was present in memory; and SD, the width of the normal distribution. The number of items being held in memory for a given set size ( $K$ ) is simply  $P_{\text{mem}} \times \text{set size}$ .

Figure 3D shows that increasing the set size from  $N = 3$  to  $N = 6$  leads to a large vertical shift in the distribution of responses, with no substantial change in the width of the normal portion of the distribution [48]. Thus, an increase in the set size appears to decrease the probability that the cued item is present in memory ( $P_{\text{mem}}$ ) without changing the precision (SD) of the representation. These data are consistent with discrete slot models and are incompatible with most resource-based models.

Although precision did not change significantly between set sizes of  $N = 3$  and  $N = 6$ , precision did improve when the set size was reduced to less than three items. This is best illustrated by a study in which subjects remembered orientation information rather than color information (Figure 3E) [49]. SD increased linearly as the set size increased up to a point and then reached an asymptote (Figure 3F). Interestingly, the inflection point in this bilinear function was closely related to VWM capacity (Figure 3G). These results suggest that resources can be shared among items until a maximum number of items ( $K_{\max}$ ) is reached [49,51].

### Evidence of continuous resources

Evidence against discrete slots and in favor of continuous resources was provided by Bays and Husain [46] using variations on the spatial memory paradigm shown in

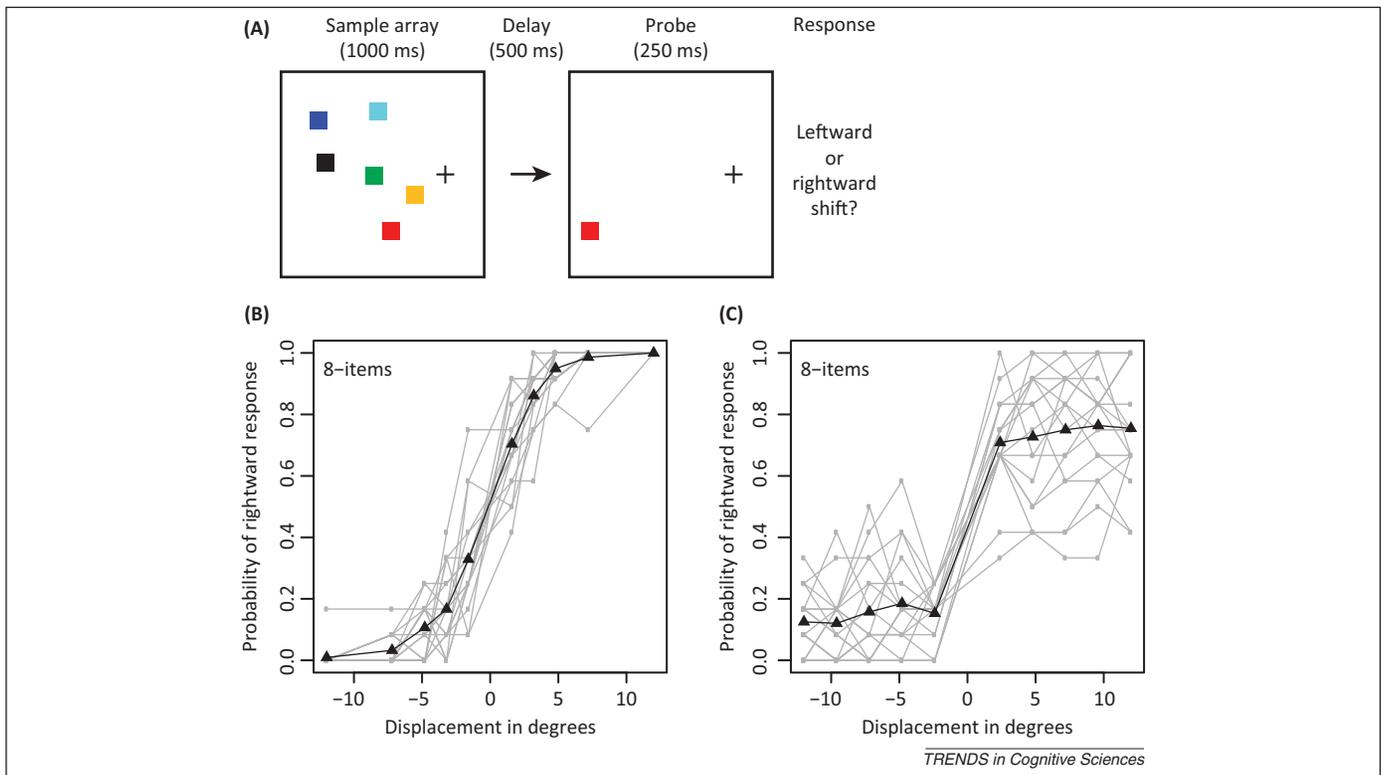
Figure 4A. A sample array of colored squares was presented, followed after a delay by a probe stimulus, and the task was to report whether the probe stimulus was displaced to the left or to the right of the corresponding item in the sample array. Slot-based models predict that observers should make errors for large set sizes, even when the displacements are very large (when no information about a given item is present in VWM). By contrast, resource-based models predict that performance should remain near perfect for sufficiently large displacements. The latter pattern was found (Figure 4B). Moreover, the precision of the representations declined monotonically as the set size increased.

Although this appears to be strong evidence of continuous resources, it appears to reflect a guessing strategy. If, as shown in Figure 4A, the probe is near the left edge of the display, the subject can guess that it was a leftward shift even if the corresponding sample item was not stored in memory. Indeed, a subsequent study [52] showed that near-perfect performance can be obtained for large displacements when this guessing strategy is possible (Figure 4B), but performance does not reach a ceiling when the task is modified to prevent this strategy (Figure 4C).

A more compelling challenge to slot-based models was provided by van den Berg *et al.* [44], who proposed a new resource-based model in which the precision varies randomly from trial to trial. According to this model, the distribution of errors in continuous report tasks (Figure 3D) reflects the average of many different normal distributions with different widths (SDs). A careful analysis showed beyond doubt that a single normal distribution systematically misfits the actual distribution of responses and that a mixture of multiple normal distributions more accurately fits the observed distribution.

This model also proposes that the amount of variation in the SD increases with the set size. For large set sizes, memories will sometimes be so imprecise that very large errors will occur, making it seem as if observers are guessing randomly. Consequently, the frequent occurrence of extreme errors for a set size of  $N = 6$  in Figure 3D may reflect very poor memory precision for a subset of trials, and not the complete absence of a representation of the item tested. However, it remains to be seen whether this variable-precision resource model fits the data better than a slot-based model in which precision is allowed to vary from trial to trial (as would be expected in any imperfect storage system) but does not increase with set size.

Sims *et al.* proposed a very different resource-based model in which VWM capacity can be conceived in terms of classic information theory [53]. In this model, the sensory input is optimally recoded so that it can be represented in terms of a specific number of bits of information. The model predicts that because of optimal recoding, observers will be able to retain more precise information when the range of possible values is small than when it is large, and this prediction was confirmed. If visual information can be arbitrarily recoded in abstract bits, this naturally brings up the question of whether the representations are still visual. It remains to be seen whether the storage of this recoded information occurs in visual cortex or instead occurs in a more generic, amodal working memory system [54].



**Figure 4.** (A) Spatial visual working memory (VWM) paradigm of Bays and Husain [46]. Observers report whether the probe is displaced leftward or rightward relative to the corresponding sample item. (B) Results from a set size of  $N = 8$  in a replication experiment [52]. The x-axis shows the displacement of the probe relative to the original item, with negative values indicating leftward and positive values indicating rightward displacement. The y-axis shows the probability that the subject reports rightward displacement. When the displacement was large, subjects were nearly perfect: they nearly always reported rightward displacement for a large rightward displacement and almost never reported rightward displacement for a large leftward displacement. Bays and Husain argued that this nearly perfect memory for large displacements for a set size of  $N = 8$  is strong evidence against the slot model and in favor of the resource model, but later research showed that these results could be explained by a guessing strategy [52]. (C) Results when the task was changed slightly to eliminate the guessing strategy. Observers were no longer nearly perfect for large displacements. Panels (B) and (C) are reprinted from [52] with kind permission from Springer Science and Business Media.

Another possibility is that VWM capacity limits are a by-product of competition between similar representations [55,56]. However, this would predict that capacity would be lower when the items to be remembered are similar to each other, which is inconsistent with the available evidence [41,57].

#### Additional evidence of discrete slots

Converging evidence from multiple experimental paradigms will be needed to distinguish between the broad classes of slot-based and resource-based models. The data from set size manipulations such as those shown in Figures 3 and 4 are not yet conclusive, but three additional approaches have provided evidence in favor of discrete slots.

First, Rouder *et al.* recorded confidence judgments from observers in a change detection task so that receiver operating characteristic (ROC) curves could be constructed [45]. When large change magnitudes are used, the all-or-none storage posited by slot models should lead to a linear ROC curve, whereas low-resolution memory representations would lead to a bowed ROC curve. The ROC curves observed were close to linear, supporting the slot model assumption of all-or-none memory encoding.

Second, Zhang and Luck combined the color wheel paradigm shown in Figure 3B with a spatial cuing manipulation [48]. The sample array contained two items along with a spatial cue, which could be valid, invalid, or neutral.

Resource-based models would predict that precision would be very high on valid trials, intermediate on neutral trials, and very low on invalid trials (because the invalidly cued objects should have 'just a few drops' of resources). Slot-based models, however, predict that precision should be the same on neutral and invalid trials (because it is not possible to have 'just a few drops' of resources in these models). Instead, the probability of having any representation at all should decline for invalid trials. This is exactly what was found.

Third, a more recent study tested whether observers could trade precision for capacity, increasing the number of items stored in VWM beyond the typical  $K_{\max}$  by decreasing the precision of the representations [58]. Observers were never able to increase  $K_{\max}$  by reducing precision, even when given monetary incentives to do so. This is strong evidence against the idea that resources can be allocated flexibly to increase the number of VWM representations.

Fourth, if observers devoted all of their resources to the items in the display, regardless of whether one or 20 items were present, then it is difficult to explain why ERP and fMRI measures of VWM delay activity increase as the set size increases from one up to the individual observer's  $K_{\max}$  and then reach an asymptote [22,24]. Resource models would instead predict that delay period activity should be constant as long as observers are devoting all their resources to the task.

Multiple sources of evidence therefore support the idea of discrete slots. However, state-of-the-art resource models [44,53] have not yet had an opportunity to explain these other results. Thus, significant progress is being made and many specific models have been ruled out, but additional research will be needed to provide definitive evidence distinguishing between these two broad classes of model.

### Individual and group differences in visual working memory capacity

As discussed earlier, VWM capacity is a stable individual difference and is impressively correlated with measures of higher cognitive function (Figure 1E). In addition, substantial differences in VWM capacity can be observed across groups. For example,  $K_{\max}$  is lower in people with schizophrenia than in healthy control subjects [59,60], with a very large effect size (Cohen's  $d$ ) of 1.11 in a study of 99 patients and 77 controls [13]. Moreover, the reduction in  $K_{\max}$  accounted for approximately 40% of the impairment for a measure of broad intellectual function in the patient group. Assuming that reduced capacity actually causes reduced intellectual function (Box 2), this finding suggests that 40% of the cognitive deficit in people with schizophrenia could be eliminated by a treatment that normalizes their VWM capacity.

Several studies have attempted to determine the causes underlying individual and group differences in VWM capacity. Among healthy young adults, Vogel and colleagues found that differences in VWM capacity can be attributed to differences in attentional processes that are responsible for filtering out irrelevant information. For example, when given arrays containing both relevant and irrelevant objects, low-capacity individuals tend to encode irrelevant information into VWM to a greater extent than high-capacity individuals do [23]. An fMRI study indicated that this reflects impaired connectivity among prefrontal cortex, the

basal ganglia, and parietal cortex [61]. In addition, recovery is slower in low-capacity individuals than in high-capacity individuals after attention is captured by irrelevant information [62]. Thus, differences in  $K_{\max}$  among healthy young adults may not reflect the capacity of VWM *per se*, but may instead reflect variations in filtering processes that control the encoding of information into VWM.

Although attentional differences explain the bulk of the between-subject variability in VWM capacity among healthy college students, other factors appear to contribute to group differences. For example,  $K_{\max}$  is lower in patients with Parkinson's disease than in healthy control subjects, but only part of this can be explained by impaired filtering; these patients also appear to have lower VWM storage capacity *per se* [63]. Similarly,  $K_{\max}$  declines in aging, as does filtering efficiency, but the timing of the filtering efficiency differences between younger and older adults is not the same as that for differences between low- and high- $K_{\max}$  younger adults [64,65].

In addition, the lower  $K_{\max}$  exhibited by schizophrenia patients does not appear to reflect impaired filtering at all. First, schizophrenia patients show the same ability to exclude irrelevant distractor items as control subjects do [60]. Second, CDA amplitude is actually greater in schizophrenia patients than in control subjects when a single object is stored in memory, but is lower in patients compared to controls when three or five items must be stored [66]. This difference was observed even in subsets of patients and controls with equivalent  $K_{\max}$  values. The patient impairment may reflect a tendency to hyperfocus on a small number of items, directly reducing VWM capacity. Schizophrenia patients may also tend to hyperfocus on perceptually salient information: they exhibit impaired filtering [67] and slowed disengagement [68] when faced with high-salience distractors.

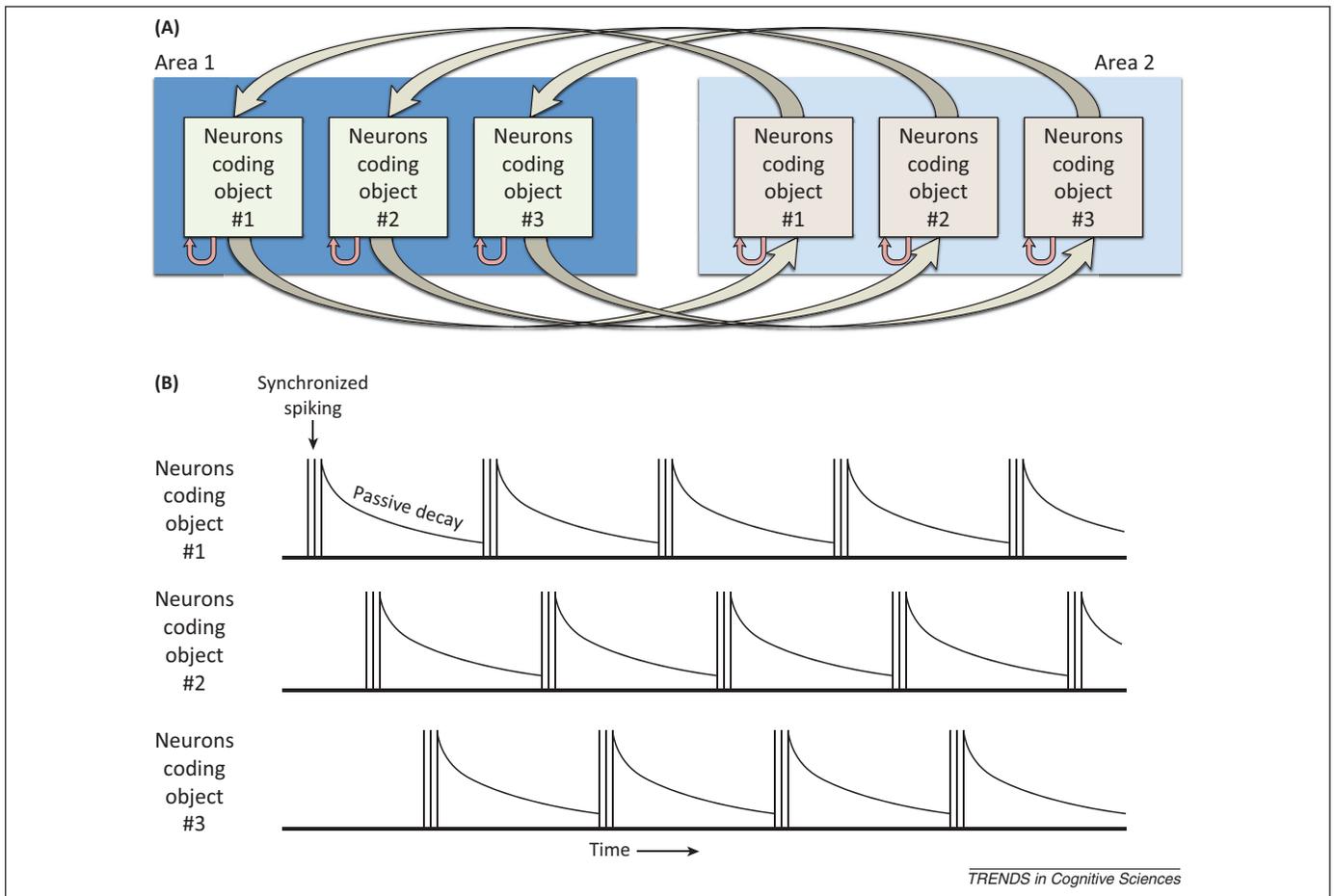
### Neural mechanisms of visual working memory

The simplicity of the change detection task makes VWM amenable to neural network modeling and electrophysiological recordings. Most neural network models assume that VWM representations are maintained by recurrent feedback loops, in which information flows from one set of neurons to another and then back again (Figure 5A) [69]. This explains the persistence of memories, the increased neural activity observed during the retention interval, and the finding of oscillations as the activity bounces back and forth among neural populations.

A recurrent feedback loop can easily maintain a single item, but it is more difficult to keep representations of multiple different items from collapsing into a single representation. To solve this problem, all the neurons that represent a given item are linked together in a synchronously firing cell assembly, and only one cell assembly fires at a given moment in time (Figure 5B). Synchrony within a cell assembly helps to maintain recurrent activation, and asynchrony between different cell assemblies avoids interference between the representations of different items. A synchronous cell assembly effectively serves as a slot in VWM. The cell assemblies are formed dynamically, combining whatever set of neurons is necessary to represent a given object.

#### Box 2. Outstanding questions

- Debate is continuing about whether VWM is best conceived as a set of discrete slot-like representations or as a flexible continuous resource. Substantial progress has been made and many specific models have been ruled out by the data. However, creative new experimental designs and analytical procedures are needed before we can definitively distinguish between these broad theory classes.
- Most models assume that trial-by-trial variations in memory (e.g., the distribution of errors shown in Figure 2D, in main text) reflect neural noise. However, much of this variance could instead reflect systematic differences across trials (e.g., differences in the specific stimuli being remembered). The source of this variance is therefore an important issue for future research.
- Studies of individual differences typically assume that having more memory capacity causes people to perform better on broader tests of cognitive ability. However, the direction of causation may actually be in the opposite direction. That is, smarter people may figure out better ways to perform working memory tasks. Determining the actual direction of causality will be vitally important in future research.
- If variations in VWM capacity actually cause variations in overall cognitive ability, then the next obvious question is whether it is possible to improve VWM capacity and thereby improve overall cognitive ability. This is currently a hot topic among working memory researchers, but no clear answer has yet emerged.



**Figure 5.** Neural representation of three cell assemblies (groups of neurons coding separate objects in visual working memory, VWM). Each cell assembly consists of a group of neurons from one or more cortical areas. In some models, neurons are recruited to a specific cell assembly at the moment of encoding to represent the features of the object being encoded, and a given neuron may be allocated to different cell assemblies depending on the information being stored in memory. **(A)** Groups of neurons coding a given object form local recurrent loops within an area (small U-shaped arrows) and long-range recurrent loops between areas (large arrows). The recurrent connections cause the activity to be maintained over time, and the activity oscillates as it bounces back and forth between neurons (both within and between cortical areas). Most models include only one or two cortical areas (e.g., inferotemporal and prefrontal cortex), but many different areas are likely synchronized in this manner. **(B)** The neurons in a given cell assembly spike together briefly (represented by vertical lines) and then the activity decays. The different cell assemblies spike at different times, minimizing interference between them. However, a given cell assembly must spike again before it decays too far (in which case the cell assembly stops firing and the VWM representation is lost). This limits the number of cell assemblies that can be simultaneously active without either interfering with each other or decaying into oblivion.

In these models, a cell assembly passively decays after each time it fires, and the representation will be lost if too much time passes before it fires again. Consequently, the number of items that can be maintained is limited by the need to keep multiple cell assemblies from firing at the same time, while also preventing long delays between successive firings of a given cell assembly so that it does not decay too far [70,71]. Realistic biophysical parameters lead to an average capacity of three or four discrete objects, with some stochastic variation in the number of items stored on each trial [71]. This model can also explain the fact that multidimensional objects can be remembered as easily as single-dimension objects [8], because the neurons coding different dimensions can be synchronized into a single cell assembly [71,72]. In general, models of this nature can explain how slot-like behavior can arise from the dynamics of a continuous neural network [56].

It is difficult to test these models from neural recordings because of the difficulty of recording from dozens of individual neurons at the same time and determining how they are linked together. Nonetheless, electrophysiological recordings from both humans and non-human primates have provided evidence that synchronized spikes

and gamma-band oscillations represent the individual cell assemblies [73], which are then sequenced by means of coupling to theta-band oscillations [74–77]. Alpha-band oscillations may also play an important role. For example, asymmetric modulations of alpha amplitude may contribute to sustained slow waves such as the CDA [78].

### Putting it all together

Differences in VWM capacity among healthy individuals are strongly predictive of broad cognitive abilities [13,14], and impairments in VWM capacity in patient groups may provide an important key to understanding their real-world cognitive impairments [13]. It is remarkable that memory for simple stimuli such as colored squares is so strongly predictive of broader measures of cognitive ability and so clearly impaired in a variety of groups. However, this is fortunate, because VWM for simple colored squares is amenable to rigorous psychophysical measurement, neural network modeling, ERP and fMRI experiments in humans, and invasive measures of neural activity in animals. Consequently, we are rapidly gaining a detailed mechanistic understanding of the factors that determine VWM capacity, and this may in turn lead to major

advances in understanding individual and group differences in broader cognitive function.

### Acknowledgments

This study was supported by grants from the National Institute of Mental Health (R01MH076226 and R01MH065034 to S.J.L. and R01MH087214 to E.K.V.) and the Office of Naval Research (N000141210972 to E.K.V.). We thank Ed Awh and Wei Ji Ma for many interesting discussions of these issues and Andy Yonelinas for comments on the manuscript.

### References

- 1 Anderson, J.R. *et al.* (1997) ACT-R: a theory of higher level cognition and its relation to visual attention. *Hum. Comput. Interact.* 12, 439–462
- 2 Meyer, D.E. and Kieras, D.E. (1997) A computational theory of executive cognitive processes and multiple-task performance: part 1. Basic mechanisms. *Psychol. Rev.* 104, 3–65
- 3 Kyllonen, P.C. and Christal, R.E. (1990) Reasoning ability is (little more than) working-memory capacity?! *Intelligence* 14, 389–433
- 4 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
- 5 Simons, D.J. and Levin, D.T. (1997) Change blindness. *Trends Cogn. Sci.* 1, 261–267
- 6 Simons, D.J. and Rensink, R.A. (2005) Change blindness: past, present, and future. *Trends Cogn. Sci.* 9, 16–20
- 7 Phillips, W.A. (1974) On the distinction between sensory storage and short-term visual memory. *Percept. Psychophys.* 16, 283–290
- 8 Luck, S.J. and Vogel, E.K. (1997) The capacity of visual working memory for features and conjunctions. *Nature* 390, 279–281
- 9 Irwin, D.E. (1991) Information integration across saccadic eye movements. *Cogn. Psychol.* 23, 420–456
- 10 Luck, S.J. (2008) Visual short-term memory. In *Visual Memory* (Luck, S.J. and Hollingworth, A., eds), pp. 43–85, Oxford University Press
- 11 Gibson, B. *et al.* (2011) Qualitative similarities in the visual short-term memory of pigeons and people. *Psychonom. Bull. Rev.* 18, 979–984
- 12 Kyllingsbaek, S. and Bundesen, C. (2009) Changing change detection: improving the reliability of measures of visual short-term memory capacity. *Psychonom. Bull. Rev.* 16, 1000–1010
- 13 Johnson, M.K. *et al.* (2013) The relationship between working memory capacity and broad measures of cognitive ability in healthy adults and people with schizophrenia. *Neuropsychology* 27, 220–229
- 14 Fukuda, K. *et al.* (2010) Quantity, not quality: the relationship between fluid intelligence and working memory capacity. *Psychonom. Bull. Rev.* 17, 673–679
- 15 Lin, P.-H. and Luck, S.J. (2012) Proactive interference does not meaningfully distort visual working memory capacity estimates in the canonical change detection task. *Front. Psychol.* 3, 42
- 16 Hartshorne, J.K. (2008) Visual working memory capacity and proactive interference. *PLoS ONE* 3, e2716
- 17 Shipstead, Z. and Engle, R.W. (2013) Interference within the focus of attention: working memory tasks reflect more than temporary maintenance. *J. Exp. Psychol. Learn. Mem. Cogn.* 39, 277–289
- 18 Jeneson, A. *et al.* (2012) Visual working memory capacity and the medial temporal lobe. *J. Neurosci.* 32, 3584–3589
- 19 Olson, I.R. *et al.* (2006) Working memory for conjunctions relies on the medial temporal lobe. *J. Neurosci.* 26, 4596–4601
- 20 Fuster, J.M. and Jervey, J.P. (1982) Neuronal firing in the inferotemporal cortex of the monkey in a visual memory task. *J. Neurosci.* 2, 361–375
- 21 Miller, E.K. and Desimone, R. (1991) A neural mechanism for working and recognition memory in inferior temporal cortex. *Science* 254, 1377–1379
- 22 Vogel, E.K. and Machizawa, M.G. (2004) Neural activity predicts individual differences in visual working memory capacity. *Nature* 428, 748–751
- 23 Vogel, E.K. *et al.* (2005) Neural measures reveal individual differences in controlling access to working memory. *Nature* 438, 500–503
- 24 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
- 25 Todd, J.J. and Marois, R. (2004) Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* 428, 751–754
- 26 Henderson, J.M. (2008) Eye movements and visual memory. In *Visual Memory* (Luck, S.J. and Hollingworth, A., eds), pp. 87–121, Oxford University Press
- 27 Irwin, D.E. (1992) Memory for position and identity across eye movements. *J. Exp. Psychol. Learn. Mem. Cogn.* 18, 307–317
- 28 Irwin, D.E. and Andrews, R.V. (1996) Integration and accumulation of information across saccadic eye movements. In *Attention and Performance XVI* (Inui, T. and McClelland, J.L., eds), pp. 125–155, MIT Press
- 29 Hollingworth, A. *et al.* (2008) Understanding the function of visual short-term memory: transsaccadic memory, object correspondence, and gaze correction. *J. Exp. Psychol. Gen.* 137, 163–181
- 30 Mannan, S.K. *et al.* (2010) Early oculomotor capture by new onsets driven by the contents of working memory. *Vis. Res.* 50, 1590–1597
- 31 Hollingworth, A. *et al.* (2013) Visual working memory modulates rapid eye movements to simple onset targets. *Psychol. Sci.* 24, 790–796
- 32 Carlisle, N.B. *et al.* (2011) Attentional templates in visual working memory. *J. Neurosci.* 31, 9315–9322
- 33 Woodman, G.F. *et al.* (2007) The role of working memory representations in the control of attention. *Cereb. Cortex* 17, i118–i124
- 34 Woodman, G.F. *et al.* (2001) Visual search remains efficient when visual working memory is full. *Psychol. Sci.* 12, 219–224
- 35 Luria, R. and Vogel, E.K. (2011) Visual search demands dictate reliance on working memory storage. *J. Neurosci.* 31, 6199–6207
- 36 Anderson, D.E. *et al.* (2013) A common discrete resource for visual working memory and visual search. *Psychol. Sci.* 24, 929–938
- 37 Drew, T. *et al.* (2011) Delineating the neural signatures of tracking spatial position and working memory during attentive tracking. *J. Neurosci.* 31, 659–668
- 38 Drew, T. and Vogel, E.K. (2008) Neural measures of individual differences in selecting and tracking multiple moving objects. *J. Neurosci.* 28, 4183–4191
- 39 Tsubomi, H. *et al.* (2013) Neural limits to representing objects still within view. *J. Neurosci.* 33, 8257–8263
- 40 Vogel, E.K. and Awh, E. (2008) How to exploit diversity for scientific gain: using individual differences to constrain cognitive theory. *Curr. Dir. Psychol. Sci.* 17, 171–176
- 41 Awh, E. *et al.* (2007) Visual working memory represents a fixed number of items regardless of complexity. *Psychol. Sci.* 18, 622–628
- 42 Alvarez, G.A. and Cavanagh, P. (2004) The capacity of visual short-term memory is set both by information load and by number of objects. *Psychol. Sci.* 15, 106–111
- 43 Brady, T.F. *et al.* (2011) A review of visual memory capacity: beyond individual items and toward structured representations. *J. Vis.* 11, 4
- 44 van den Berg, R. *et al.* (2012) Variability in encoding precision accounts for visual short-term memory limitations. *Proc. Natl. Acad. Sci. U.S.A.* 109, 8780–8785
- 45 Rouder, J.N. *et al.* (2008) An assessment of fixed-capacity models of visual working memory. *Proc. Natl. Acad. Sci. U.S.A.* 105, 5975–5979
- 46 Bays, P.M. and Husain, M. (2008) Dynamic shifts of limited working memory resources in human vision. *Science* 321, 851–854
- 47 Wilken, P. and Ma, W.J. (2004) A detection theory account of change detection. *J. Vis.* 4, 1120–1135
- 48 Zhang, W. and Luck, S.J. (2008) Discrete fixed-resolution representations in visual working memory. *Nature* 453, 233–235
- 49 Anderson, D.E. *et al.* (2011) Precision in visual working memory reaches a stable plateau when individual item limits are exceeded. *J. Neurosci.* 31, 1128–1138
- 50 Anderson, D.E. and Awh, E. (2012) The plateau in mnemonic resolution across large set sizes indicates discrete resource limits in visual working memory. *Atten. Percept. Psychophys.* 74, 891–910
- 51 Xu, Y. and Chun, M.M. (2006) Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature* 440, 91–95
- 52 Thiele, J.E. *et al.* (2011) On perfect working-memory performance with large numbers of items. *Psychonom. Bull. Rev.* 18, 958–963
- 53 Sims, C.R. *et al.* (2012) An ideal observer analysis of visual working memory. *Psychol. Rev.* 119, 807–830
- 54 Saults, J.S. and Cowan, N. (2007) A central capacity limit to the simultaneous storage of visual and auditory arrays in working memory. *J. Exp. Psychol. Gen.* 136, 663–684
- 55 Franconeri, S.L. *et al.* (2013) Flexible cognitive resources: competitive content maps for attention and memory. *Trends Cogn. Sci.* 17, 134–141

- 56 Wei, Z. *et al.* (2012) From distributed resources to limited slots in multiple-item working memory: a spiking network model with normalization. *J. Neurosci.* 32, 11228–11240
- 57 Lin, P.-H. and Luck, S.J. (2008) The influence of similarity on visual working memory representations. *Vis. Cogn.* 17, 356–372
- 58 Zhang, W. and Luck, S.J. (2011) The number and quality of representations in working memory. *Psychol. Sci.* 22, 1434–1441
- 59 Gold, J.M. *et al.* (2003) Working memory for visual features and conjunctions in schizophrenia. *J. Abnorm. Psychol.* 112, 61–71
- 60 Gold, J.M. *et al.* (2006) Intact attentional control of working memory encoding in schizophrenia. *J. Abnorm. Psychol.* 115, 658–673
- 61 McNab, F. and Klingberg, T. (2008) Prefrontal cortex and basal ganglia control access to working memory. *Nat. Neurosci.* 11, 103–107
- 62 Fukuda, K. and Vogel, E.K. (2009) Human variation in overriding attentional capture. *J. Neurosci.* 29, 8726–8733
- 63 Lee, E.-Y. *et al.* (2010) Visual working memory deficits in patients with Parkinson's disease are due to both reduced storage capacity and impaired ability to filter out irrelevant information. *Brain* 133, 2677–2689
- 64 Jost, K. *et al.* (2011) Are old adults just like low working memory young adults? Filtering efficiency and age differences in visual working memory. *Cereb. Cortex* 21, 1147–1154
- 65 Cashdollar, N. *et al.* (2013) Prolonged disengagement from attentional capture in normal aging. *Psychol. Aging* 28, 77–86
- 66 Leonard, C.J. *et al.* (2013) Toward the neural mechanisms of reduced working memory capacity in schizophrenia. *Cereb. Cortex* 23, 1582–1592
- 67 Hahn, B. *et al.* (2010) Failure of schizophrenia patients to overcome salient distractors during working memory encoding. *Biol. Psychiatry* 68, 603–609
- 68 Mayer, J.S. *et al.* (2012) Impaired contingent attentional capture predicts reduced working memory capacity in schizophrenia. *PLoS ONE* 7, e48586
- 69 Deco, G. and Rolls, E.T. (2008) Neural mechanisms of visual memory: a neurocomputational perspective. In *Visual Memory* (Luck, S.J. and Hollingworth, A., eds), pp. 247–289, Oxford University Press
- 70 Lisman, J.E. and Idiart, M.A.P. (1995) Storage of  $7 \pm 2$  short-term memories in oscillatory subcycles. *Science* 267, 1512–1515
- 71 Raffone, A. and Wolters, G. (2001) A cortical mechanism for binding in visual working memory. *J. Cogn. Neurosci.* 13, 766–785
- 72 Luck, S.J. and Vogel, E.K. (1998) Response from Luck and Vogel. *Trends Cogn. Sci.* 2, 78–80
- 73 Pesaran, B. *et al.* (2002) Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nat. Neurosci.* 5, 805–811
- 74 Liebe, S. *et al.* (2012) Theta coupling between V4 and prefrontal cortex predicts visual short-term memory performance. *Nat. Neurosci.* 15, 456–462 S451–452
- 75 Lee, H. *et al.* (2005) Phase locking of single neuron activity to theta oscillations during working memory in monkey extrastriate visual cortex. *Neuron* 45, 147–156
- 76 Sauseng, P. *et al.* (2009) Brain oscillatory substrates of visual short-term memory capacity. *Curr. Biol.* 19, 1846–1852
- 77 Siegel, M. *et al.* (2009) Phase-dependent neuronal coding of objects in short-term memory. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21341–21346
- 78 van Dijk, H. *et al.* (2013) Modulations in oscillatory activity with amplitude asymmetry can produce cognitively relevant event-related responses. *Proc. Natl. Acad. Sci. U.S.A.* 107, 900–905 10.1073/pnas.0908821107
- 79 Hyun, J.-S. *et al.* (2009) The comparison of visual working memory representations with perceptual inputs. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1140–1160
- 80 Pashler, H. (1988) Familiarity and visual change detection. *Percept. Psychophys.* 44, 369–378
- 81 Cowan, N. *et al.* (2005) On the capacity of attention: its estimation and its role in working memory and cognitive aptitudes. *Cogn. Psychol.* 51, 42–100
- 82 Rouder, J.N. *et al.* (2011) How to measure working memory capacity in the change detection paradigm. *Psychonom. Bull. Rev.* 18, 324–330
- 83 Morey, R.D. and Morey, C.C. (2011) WoMMBAT: A user interface for hierarchical Bayesian estimation of working memory capacity. *Behav. Res. Methods* 43, 1044–1065