

Evidence for a fixed capacity limit in attending multiple locations

Edward F. Ester · Keisuke Fukuda · Lisa M. May · Edward K. Vogel · Edward Awh

Published online: 12 November 2013
© Psychonomic Society, Inc. 2013

Abstract A classic question concerns whether humans can attend multiple locations or objects at once. Although it is generally agreed that the answer to this question is “yes,” the limits on this ability are subject to extensive debate. According to one view, attentional resources can be flexibly allocated to a variable number of locations, with an inverse relationship between the number of selected locations and the quality of information processing at each location. Alternatively, these resources might be quantized in a “discrete” fashion that enables concurrent access to a small number of locations. Here, we report a series of experiments comparing these alternatives. In each experiment, we cued participants to attend a variable number of spatial locations and asked them to report the orientation of a single, briefly presented target. In all experiments, participants’ orientation report errors were well-described by a model that assumes a fixed upper limit in the number of locations that can be attended. Conversely, report errors were poorly described by a flexible-resource model that assumes no fixed limit on the number of locations that can be attended. Critically, we showed that these discrete limits were predicted by cue-evoked neural activity elicited before the onset of the target array, suggesting that performance was limited by selection processes that began prior to subsequent encoding and

memory storage. Together, these findings constitute novel evidence supporting the hypothesis that human observers can attend only a small number of discrete locations at an instant.

Keywords Attention · Working memory · ERP

As a doctoral student in the laboratory that was codirected by Edward E. Smith and John Jonides, the senior author of this article was introduced to the field of cognitive neuroscience when it was still an emerging trend rather than the dominant approach for understanding cognition. Ed Smith was the perfect advisor during this exciting time. He had always been celebrated for his ability to see the broad conceptual connections between different domains of psychology, and that same vision fostered his transition from an eminent career in traditional cognitive psychology to a prominent place in the field of cognitive neuroscience. As a thesis committee member, Ed Smith had a substantial influence on E.A.’s doctoral work examining the links between working memory and attention. Indeed, a core theme of that work is still evident in this article: We report that—as with visual working memory—capacity limits in visual selective attention are best described by discrete rather than flexible resource allocation. Moreover, we show that capacity limits for internal storage in working memory are strongly correlated with the number of positions to which covert attention could be allocated. We are grateful to have benefited from Ed Smith’s lasting and powerful contributions to psychology and neuroscience, and we hope that this work can serve as a small token of our gratitude for his guidance, energy, and vision.

The human visual system has a limited processing capacity. Consequently, mechanisms of selective attention are needed to prioritize stimuli that are relevant to the current behavioral goals. Converging evidence from behavioral

E. F. Ester · K. Fukuda · L. M. May · E. K. Vogel · E. Awh
Department of Psychology, University of California San Diego,
La Jolla, CA, USA

L. M. May · E. K. Vogel · E. Awh
Department of Psychology, University of Oregon,
Eugene, Oregon, USA

E. F. Ester (✉)
Department of Psychology, University of California San Diego,
9500 Gilman Drive, La Jolla, CA 92093, USA
e-mail: edward.ester01@gmail.com

(e.g., Alvarez & Cavanagh, 2005; Alvarez, Gill, & Cavanagh, 2012; Awh & Pashler, 2000; Franconeri, Alvarez, & Enns, 2007; Kramer & Hahn, 1995), electrophysiological (e.g., Anderson, Vogel, & Awh, 2011, 2013; Drew & Vogel, 2008; Ester, Drew, Klee, Vogel, & Awh, 2012; Malinowski, Fuchs, & Müller, 2007; Müller, Malinowski, Grube, & Hillyard, 2003), and neuroimaging (e.g., McMains & Somers, 2004, 2005) studies have suggested that human observers can select at least two noncontiguous locations at once. Although it is generally agreed that there is some limit in the number of locations or stimuli that can be concurrently attended, the nature of this limit is unclear. The goal of the present work was to compare two broad theoretical perspectives on this issue.

Before introducing these perspectives, we should first explain exactly what we mean by “selection.” A typical visual scene contains a multitude of sensory signals, only a handful of which may be relevant to the current behavioral goals. To facilitate processing, the visual system must isolate behaviorally relevant signals from irrelevant signals (and also isolate these relevant signals from one another). Thus, the process of “selecting” a stimulus (or location) requires generating an *individuated* representation of that stimulus that is perceptually and/or neurally segregated from other representations (e.g., Kahneman, Treisman, & Gibbs, 1992; Xu & Chun, 2009). Here, our goal was to examine capacity limits in selection as it pertains to this individuation process.

One possibility is that capacity limits in visual selective attention are determined by a “flexible resource.” This framework proposes that attentional resources can be allocated to as many or as few locations (or items) as an observer wishes, with the caveat that as more locations are selected, these resources are spread more thinly and the efficiency of visual processing at each selected location is reduced. A previous study (Franconeri et al., 2007) was consistent with this prediction. In one experiment, participants were cued to monitor two, three, four, five, or six locations. After a brief retention interval, a search array was presented, and participants were instructed to search exclusively through cued locations in order to determine whether a target was present (foils were presented at noncued locations). When dense (24 total items) search arrays were presented, the results suggested that participants could select two or three locations at most. Conversely, when sparse (12 total items) arrays were presented, participants could select upward of five or six locations. This result was interpreted as follows: When dense search arrays were presented, participants needed to adopt fine-grained selection regions in order to avoid accidentally selecting information from neighboring locations, and fewer locations could be selected. However, when sparse search arrays were presented, coarse selection regions could be utilized, and participants were thus able to select a larger number of locations. This inverse relationship between capacity and granularity is considered a hallmark of flexible resource allocation.

The flexible resource view also predicts that stimuli of greater complexity (relatively speaking) will consume a greater proportion of attentional resources. Consequently, observers may be able to select many “simple” objects, but only a few “complex” objects. Indeed, it is well-known that detecting a singleton target among uniform distractors (e.g., a green disk among red disks) is trivial, but detecting a conjunction target (i.e., a target defined by two or more features) in a heterogeneous display is much more difficult (e.g., Treisman & Gelade, 1980). More recently, Scharff, Palmer, and Moore (2011a, 2011b) asked observers to report either the location of a higher-contrast disk (presented amongst three lower-contrast distractors) or the location of a word that matched a prespecified semantic category (presented among three non-target words). On some trials, the stimuli were presented sequentially (e.g., two items in each of two successive displays), whereas on other trials, all four stimuli were presented simultaneously. Assuming that words are more “complex” than disks, the flexible-resource view predicts that observers should be able to select fewer items in the word discrimination task than in the contrast discrimination task. Consequently, performance on the word discrimination task should benefit from sequential (relative to simultaneous) displays, whereas performance on the contrast discrimination task should be roughly equal across these displays. This is precisely what Scharff and colleagues observed: Performance in the contrast discrimination task was virtually identical for simultaneous and sequential displays, but performance in the word discrimination task was substantially lower during simultaneous (relative to sequential) displays.

Another possibility is that capacity limits in visual selective attention are determined by a “discrete resource.” Here, attentional resources are quantized in a “slot-like” fashion that precludes selection beyond a small (and fixed) number of locations or items (typically estimated at around three or four). Once this limit has been surpassed, observers can obtain no information about additional items. Evidence consistent with this possibility has been reported in the subitizing literature. For example, Trick and Pylyshyn (1993, 1994) asked subjects to rapidly enumerate the number of items present in a display and found that responses were fast and accurate for arrays containing up to four targets. However, response latencies and error rates increased monotonically with target numerosity once this range was exceeded. This profile was well-approximated by a bilinear function, and static performance at set sizes 1–4 was interpreted as evidence for a limited-capacity selection mechanism that enables concurrent access to a small number of stimuli.

How could a discrete resource account for the findings favoring unlimited-capacity (or variable-capacity) selection in simple search? We suspect that in many cases, the apparent selection of large numbers of items may actually reflect the selection of a relatively small number of *groups* of items.

Consider a case in which an observer is asked to report the presence or absence of a green disk among a variable number of red disks. Prior work (e.g., Treisman & Gelade, 1980) suggested that performance (e.g., response latency) on this task is equivalent for displays containing 10 or 100 red disks (e.g., Treisman & Gelade, 1980). However, this does not necessarily mean that the observer can select (i.e., individuate) 100 items. Instead, one alternative explanation is that these displays are perceived as a much smaller number of perceptual *groups*. For instance, observers might perceive the green target as a single item, and the array of red distractors (grouped by the shared color red) as a second item; indeed, this is precisely the explanation that Duncan and Humphreys (1989) offered to explain the phenomenon of “pop out” (Duncan & Humphreys, 1989; Rensink & Enns, 1995). Moreover, it is well-known that humans are able to rapidly segregate and discriminate textures that share similar first- and second-order statistics (e.g., Bergen & Julesz, 1983; Sagi & Julesz, 1985). Thus, it should be easy to detect a set of rightward-tilted bars amongst a larger array of leftward-tilted bars, but considerably more difficult to detect a single target word amongst nontarget words. Note that in both of these examples, observers need not individuate all (or even most) of the unique elements in a given display. Instead, rapid discrimination can be guided by knowledge of local image statistics and/or other Gestalt grouping cues.

This general line of reasoning may suffice to explain the findings reported by Scharff et al. (2011a, 2011b). For example, during the contrast discrimination task, observers may have treated the three identical distractors as a background texture, thereby allowing rapid and efficient identification of the target. However, the same strategy could not be used to solve the word discrimination task; different words contain a multitude of different features with different local statistics, and are thus not easily grouped. Following this general logic, the discrete-resource model can provide a plausible alternative explanation for many demonstrations that appear to favor an unlimited- or variable-capacity selection mechanism (though see Huang & Pashler, 2005, for one possible exception to this trend).

The goal of the present study was to compare flexible- and discrete-resource models of attentional selection in the same task. In Experiment 1, we asked participants to monitor a variable number of cued locations and to report the orientation of a single, briefly presented and masked, target (see Fig. 1). We then fit participants' orientation report errors (i.e., the difference between the reported and actual target orientations) with quantitative functions that encapsulated key predictions of the flexible- and discrete-resource views. The key question here was whether or not participants could provide nonzero information about large numbers of items (as predicted by the flexible-resource view) or whether the proportion of “guessing” responses would be large, indicating that participants

were incapable of selecting all elements within a display (as predicted by the discrete-resource view).

Experiment 1

Method

Participants A group of 17 undergraduate students participated in Experiment 1 in exchange for course credit. All of the participants gave both written and oral informed consent, and all reported normal or corrected-to-normal visual acuity and color vision. The data from one participant could not be modeled (visual inspection of this participant's report error histograms revealed a flat function over orientation space, suggesting that he or she simply guessed on every trial); the data reported here reflect the remaining 16 participants.

Stimuli and apparatus The stimuli were generated in MATLAB (Mathworks, Natick, MA) using the Psychophysics Toolbox software (Brainard, 1997; Pelli, 1997) and rendered on a medium-gray background via a CRT monitor cycling at 120 Hz. Participants were seated approximately 60 cm from the monitor (head position was unconstrained) and instructed to maintain fixation on a small dot (subtending 0.2°) in the center of the display for the duration of both tasks. Each participant completed five blocks of 60 trials.

Procedure The selection task is shown in Fig. 1. Each trial began with a cue array consisting of eight circles (2° radius) spaced equally along the perimeter of an imaginary circle (5° radius) centered at fixation. On each trial, four or all eight of these circles were rendered in red. On set size 4 trials, the remaining circles were rendered in black. During set size 4 trials, we cued either the upper or lower two positions (as in Fig. 1) or the leftmost and rightmost two positions (i.e., the locations of the black circles in Fig. 1) in the display.¹ After 2 s, the target array was presented. As is shown in Fig. 1, the target array contained a single radial target (whose position was unknown to the participant in advance) and seven diametrical distractors. All stimuli were embedded within four-dot contours (see Fig. 1). After 100 ms, all display elements were removed except for the four-dot contour at the target's

¹ Note that our cued locations were always arranged contiguously. One might wonder whether our results would generalize to a situation in which cued locations were randomly arranged or interleaved between uncued locations. The present results cannot speak directly to this question. However, in pilot experiments not reported here, we obtained similar findings (i.e., the discrete-resource model always outperformed the flexible-resource model) when we cued random locations in the left and right hemifields on each trial (with the constraint that each hemifield contained two cued positions).

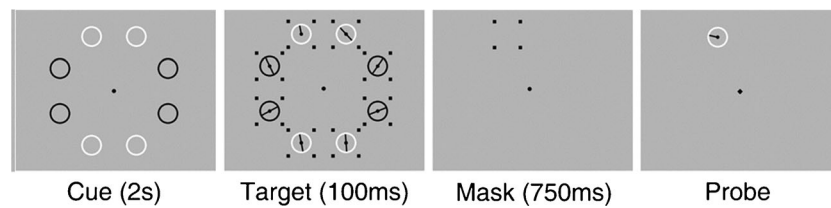


Fig. 1 Selection task. Each trial began with a cue array containing four or eight red circles (shown here in white) arranged around the circumference of an imaginary circle at fixation. The target array then appeared for 100 ms; participants were instructed to discriminate the orientation of a single radial target (top left of target panel) that appeared in one of the cued locations. The target array was followed by a 750-ms substitution

mask and the appearance of a randomly oriented probe at the target's location. Participants manipulated the orientation of the probe using the arrow keys on a standard keyboard until it matched their percept of the target's orientation. Note that the stimuli are not drawn to scale; see the Method section for information about the stimulus size, eccentricity, and so forth

location. Under these circumstances, the contour would typically act as a “substitution mask” that prohibits observers from accessing information at the masked location (e.g., Enns & DiLollo, 1997). We included masks to discourage participants from consulting a sensory (i.e., iconic) representation of the target array that might persist for a few hundred milliseconds after its offset. Substitution masks are a natural choice because—unlike metacontrast, backward, or lateral masks—they are rendered ineffective when spatial attention is directed to the masked location prior to target offset (Enns & Di Lollo, 1997). Thus, our assumption was that if participants selected a subset of the (cued) locations prior to the onset of the target array, then information appearing at these locations would escape masking. Conversely, information at unselected locations should be highly susceptible to masking, and this should prevent participants from querying a lingering sensory representation of items that they failed to select. Note that only the target location was masked; thus, the mask also functioned as a 100 %-valid postcue regarding the target's location. The mask display was presented for 750 ms and was followed by the presentation of a randomly oriented probe at the target's location. Participants were allowed to adjust the orientation of the target (via the left and right arrow keys on a standard keyboard) until it matched their percept of the target, entering their final response by pressing the space bar. Participants were instructed to respond as precisely as possible, and no response deadline was imposed.

It is now well-known that visual selective attention and visual working memory (WM) share common neural substrates (see, e.g., Awh & Jonides, 2001, for a review). Thus, one might also ask whether putative flexible or discrete limits in visual selective attention also manifest during the temporary retention of visual information. To examine this possibility, each participant also completed a WM task whose design was similar to that of the selection task. Here, the cue array remained on the screen for 1,000 ms, and was followed by a 1,000-ms sample array. Radial line segments were presented inside of each cued location (no stimuli were presented in noncued circles), and participants were instructed to remember the orientations of these stimuli over a subsequent

1,000-ms blank interval. No masks were presented. At the end of each trial, a probe stimulus appeared at one of the cued locations. Participants adjusted the orientation of this probe using the keyboard until it matched the orientation of the target that had appeared in that position. Again, participants were instructed to respond as precisely as possible, and no response deadline was imposed.

Data analysis For each task, we computed a distribution of orientation report errors (i.e., the angular difference between the reported and actual target orientations) for each participant and each set size. We then attempted to describe these distributions using quantitative functions that would encapsulate the key predictions of flexible- and discrete-resource models. First, consider first a hypothetical case in which a participant is presented with a single radial target. Under these conditions, the probability of observing report error x (where $-\pi \leq x \leq \pi$) is given by a von Mises distribution (the circular analogue of a standard normal distribution) with mean μ (uniquely determined by the perceived target orientation θ) and concentration k (uniquely determined by σ and corresponding to the precision of the observer's representation; see Eq. 1):

$$p(x|\theta, \sigma) = \frac{e^{k \cos(x-\mu)}}{2\pi I_0(k)}, \quad (1)$$

where I_0 is the modified Bessel function of the first kind of order 0. In the absence of any systematic perceptual biases (i.e., if θ is a reliable estimator of the target's orientation), then estimates of μ should take values near 0° , and observers' performance should be limited primarily by noise (σ).

In contrast, discrete-resource models predict a fixed upper limit on the number of locations that can be selected. Once this limit is exceeded, participants should be unable to obtain any information about additional items in a display. Consider the case of a participant who can select a maximum of three locations. During set size 4 trials, we assume that this participant will select a random subset of three cued locations. Thus, on 75 % of trials the target will appear at a selected location,

and the participant will obtain some nonzero amount of information regarding its orientation. Here, we would once again expect the participant's responses to be normally distributed around μ (i.e., 0° response error), with relatively few high-magnitude errors. However, on the remaining 25 % of trials, the target will appear at an unselected location. If no information can be gleaned from unselected locations, then the participant will have to guess. Across many trials, these guesses will manifest as a uniform distribution from $-\pi$ to π . Because participants will sometimes select and sometimes fail to select the target's location, the empirically observed distribution of response errors can be expressed as the weighted sum of a uniform and a von Mises distribution (Zhang & Luck, 2008, Eq. 2):

$$p(x|\theta, \sigma, nr) = (1-nr) \frac{e^{k \cos(x-\mu)}}{2\pi I_0(k)} + \frac{nr}{2\pi}, \quad (2)$$

where μ and k are the mean and concentration of a von Mises distribution (as in Eq. 1), and nr is the height of a uniform distribution with range $-\pi$ to π .

Maximum likelihood estimation (MLE) was used to fit Eqs. 1 and 2 to each observer's distribution of response errors (separately for each set size and task—i.e., selection vs. WM). Initially we allowed all parameters in each model (μ and σ for the flexible-resource model; μ , σ , and nr for the discrete-resource model) to vary across participants and set sizes. In all of the experiments reported here, estimates of μ never deviated from 0. Thus, we fixed this parameter at 0 and excluded it from further analyses. Fitting was performed iteratively over a wide range of starting values in an effort to avoid local minima. Finally, we imposed lower boundaries on the range of possible parameter values (a minimum σ and nr of 0.01 and 0, respectively) in order to prohibit nonsensical estimates (e.g., negative values of σ).

To compare the models, we used Bayesian model comparison (BMC; MacKay, 2003; Wasserman, 2000). This method returns the likelihood of a model given the data, while correcting for model complexity (i.e., the number of free parameters). Unlike traditional model comparison methods (e.g., adjusted r^2 and likelihood ratio tests), BMC does not rely on single-point estimates of the model parameters. Instead, it integrates information over parameter space, and thus accounts for variations in a model's performance over a wide range of possible parameter values. We also report traditional goodness-of-fit measures (e.g., adjusted r^2 values, where the amount of variance explained by a model is weighted to account for the number of free parameters it contains) for the flexible and discrete models (after sorting response errors into 25 bins, each 14.4° wide). As is shown below, the adjusted r^2 values were larger for the discrete than for the flexible model in all experiments. However, we emphasize that these

values can be influenced by arbitrary choices about how to summarize the data, such as the number of bins to use when constructing a histogram of response errors (e.g., one can arbitrarily increase or decrease estimates of r^2 to a moderate extent by manipulating the number of bins). Thus, they should not be viewed as conclusive evidence suggesting that one model systematically outperforms another.

Results and discussion

Selection task Panels A and B of Fig. 2 depict the mean (± 1 SEM) distribution of report errors across participants for set size 4 (2A) and set size 8 (2B) trials of the selection task. Note that both distributions feature a prominent uniform component. At the individual-participant level, a mixture distribution (Eq. 2) capturing key predictions of a discrete-resource model provided a reasonable description of report errors for set sizes 4 (mean adjusted $r^2 = .89 \pm .01$) and 8 (mean adjusted $r^2 = .57 \pm .05$).² Conversely, a circular Gaussian distribution capturing the predictions of a flexible-resource model (Eq. 1) provided a poor description of the data (mean adjusted $r^2 = .39 \pm .02$ and $.23 \pm .02$ for set sizes 4 and 8, respectively). Bayesian model comparison revealed that the log likelihood of the mixture distribution was 35.61 ± 2.41 and 11.25 ± 1.57 units larger than the likelihood of the Gaussian distribution for the set size 4 and set size 8 trials, respectively. For exposition, a log likelihood difference of 11.25 units means that the data are $e^{11.25}$, or $\sim 76,880$ times, more likely to result from the mixture distribution relative to the Gaussian distribution.

The estimates of σ and nr returned by this model are listed in Table 1 (top row). Note that estimates of σ appear to increase as a function of set size (i.e., the precision of participants' report error appears to decrease with set size). This result is nominally consistent with the inverse relationship between number and granularity predicted by flexible-resource models. However, we would caution against this interpretation for three reasons. First, the increase in σ with set size was only marginally significant ($p = .07$). Second, note that both the discrete- and flexible-resource models performed relatively poorly at set size 8 (adjusted r^2 s = $.57$ and $.23$, respectively). Thus, the parameter estimates returned by these models should be treated as suspect. Finally—and most importantly—flexible-resource models predict no upper limit on the number of locations that can be selected. However, our data clearly indicate that a distribution containing a uniform component corresponding to an upper limit in the number of locations that can be selected provides a superior account of the observed data in all conditions.

² The smooth curves overlaid in Fig. 2 were fit to the mean distributions of report errors across participants. Fits at the individual-participant level were substantially noisier.

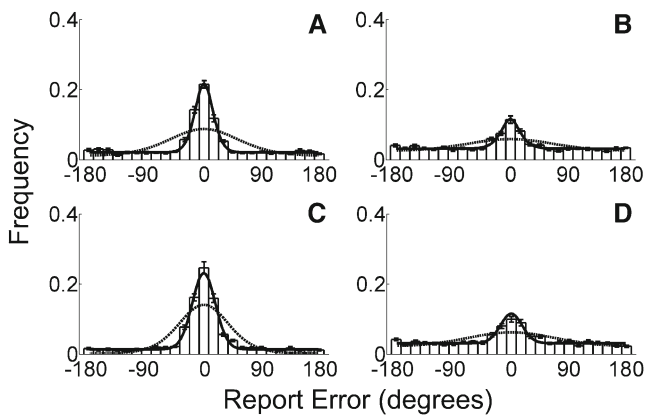


Fig. 2 Results of Experiment 1. Panels A and B depict the mean (± 1 SEM) distributions of response errors across participants for set size 4 (A) and set size 8 (B) trials of the selection task, while panels C and D depict the mean (± 1 SEM) distributions of response errors across participants for set size 4 (A) and set size 8 (B) trials of the WM task. In all panels, the best-fitting flexible- (dashed lines) and discrete- (solid lines) resource models are overlaid

WM task Panels C and D of Fig. 2 depict the mean distributions of response errors observed across participants during set size 4 (2C) and set size 8 (2D) trials of the WM task. As was the case in the selection task, these distributions were well-approximated by a mixture distribution (mean adjusted r^2 s = $.91 \pm .02$ and $.57 \pm .06$ for set size 4 and 8 trials, respectively) and poorly described by a Gaussian distribution (mean adjusted r^2 s = $.62 \pm .05$ and $.29 \pm .05$ for set size 4 and 8 trials, respectively), corroborating earlier findings (Zhang & Luck, 2008, 2009, 2011). Additionally, Bayesian model comparison revealed that the log likelihood of the mixture distribution was 23.28 ± 2.69 and 7.87 ± 1.36 units larger than the likelihood of the Gaussian distribution (for set size 4 and set size 8 trials, respectively). The estimates of σ and nr returned by this model are listed in Table 1 (bottom row). As in the selection task, estimates of σ increased as a function of set size. However, this increase was driven exclusively by one participant [a direct comparison of σ estimates revealed no reliable effect of set size; $t(15) = 1.31$, $p = .20$]. Critically, the discrete-resource model still provided a substantially better account of the empirically observed distributions of response errors, consistent with prior work (e.g., Zhang & Luck, 2008).

Table 1 Mean (\pm SEM) parameter estimates returned by the discrete-resource model for the selection and WM tasks in Experiment 1

	Set Size 4		Set Size 8	
	σ	nr	σ	nr
Selection task	14.50 (0.63)	0.51 (0.22)	17.17 (1.29)	0.77 (0.02)
WM task	18.22 (0.53)	0.32 (0.05)	24.34 (4.48)	0.75 (0.03)

Values of σ are in degrees

Relationships between selection and WM capacity The aforementioned results are broadly consistent with a discrete-resource view of attentional selection and WM storage. Specifically, participants' distributions of orientation report errors were well-approximated by a mixture distribution designed to capture guesses, and poorly described by a function that assumed no guessing. Next, we wondered whether these apparent discrete limits in attentional selection and WM storage were determined by a common resource. To investigate this, we correlated capacity estimates³ obtained in the selection and WM tasks. Because the discrete-resource model was a poor predictor of response errors on set size 8 trials in both tasks, only set size 4 trials were used to compute this correlation. However, we found a marginally significant positive correlation between these two variables ($r = .42$, $p = .10$) when set size 8 trials were considered, and we observed a robust positive correlation between selection and WM capacity when we averaged capacity estimates over set size 4 and 8 trials: $r = .75$, $p < .01$. As is shown in Fig. 3, we observed a strong correlation between the capacity estimates in the two tasks (Fig. 3; $r = .77$, $p < .01$). This result is nominally consistent with the view that capacity limits in selective attention and WM storage are determined by a common resource.

Control experiments Before considering these findings in more detail, some alternative explanations merit consideration. For example, one possibility is that the discrete item limits that we observed in our selection task are an artifact of masking. To investigate this possibility, 18 new observers completed a version of Experiment 1 that omitted the masks (all other aspects of the design and analysis were identical to those described above). The results of this experiment are shown in Fig. 4. Specifically, panels A and B depict the mean (± 1 SEM) distributions of report errors across participants during set size 4 and 8 trials of the selection task (respectively). As in Experiment 1 (see Figs. 2A and B), both distributions feature prominent uniform components. At the individual-participant level, a mixture function (Eq. 2) accounted for $.90 (\pm .01)$ and $.91 (\pm .03)$ of the variance in participants' report errors during set size 4 and 8 trials, respectively. Conversely, a Gaussian distribution (Eq. 1) accounted for just $.68 (\pm .04)$ and $.53 (\pm .04)$ of the variance in report errors. Bayesian model comparison yielded similar findings: The likelihoods of the mixture distribution were 46.02 ± 4.73 and 49.54 ± 6.32 units larger than the likelihood of the Gaussian distribution (for set size 4 and 8 trials, respectively). The parameter estimates returned by the discrete-resource model are shown in Table 2 (top row). The estimates of nr

³ The nr variable corresponds to the proportion of trials on which the participant failed to select the cued location and was forced to guess. Thus, we defined "capacity" as $N(1 - nr_N)$, where N is the number of cued locations.

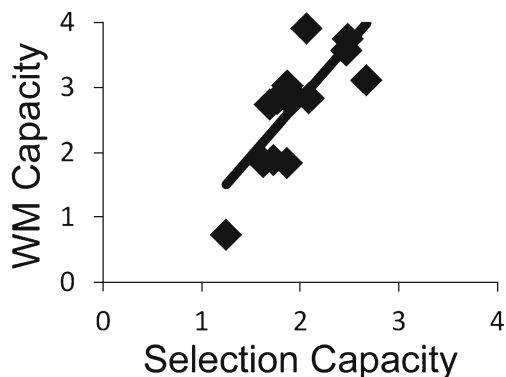


Fig. 3 Estimates of selection and WM capacity are strongly correlated ($r=.77$)

were substantially lower than those obtained in Experiment 1 (cf. Tables 1 and 2), suggesting that participants were more likely to successfully encode and store the target when masks were eliminated. However, the estimates of σ were comparable to those obtained in Experiment 1 and did not vary with set size.

Similar results were obtained in the WM task (which was identical in every respect to the memory task used in Exp. 1). Panels C and D of Fig. 4 depict the mean distributions of report errors observed across participants during set size 4 (C) and set size 8 (D) trials. These distributions were reasonably well-approximated by a mixture distribution (mean adjusted r^2 s=.84 \pm .05 and .45 \pm .07 for set size 4 and 8 trials, respectively), and poorly approximated by a Gaussian distribution (mean adjusted r^2 s=.60 \pm .04 and .28 \pm .04 for set size 4 and 8 trials). Furthermore, Bayesian model comparison revealed that the log likelihoods of the mixture distribution were 21.33 \pm 2.96 and 5.97 \pm 1.54 units larger than the likelihood

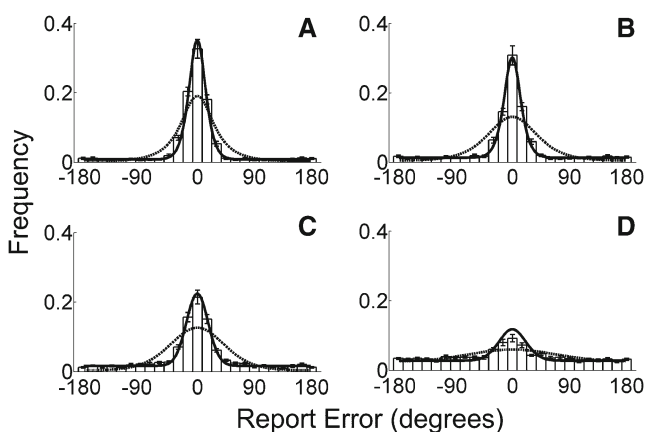


Fig. 4 Results of Control Experiment 1. As in Fig. 2, panels A and B depict the mean (± 1 SEM) distributions of response errors across participants for set size 4 (A) and set size 8 (B) trials of the selection task, while panels C and D depict the mean (± 1 SEM) distributions of response errors across participants for set size 4 (C) and set size 8 (D) trials of the WM task. In all panels, the best-fitting flexible- (dashed lines) and discrete- (solid lines) resource models are overlaid

Table 2 Mean (\pm SEM) parameter estimates returned by the discrete-resource model for the selection and WM tasks in Control Experiment 1, in which masks were removed from the target display

	Set Size 4		Set Size 8	
	σ	nr	σ	nr
Selection task	14.69 (0.87)	0.19 (0.03)	15.09 (0.95)	0.31 (0.04)
WM task	18.51 (0.98)	0.39 (0.05)	38.19 (9.78)	0.66 (0.06)

Values of σ are in degrees

of the Gaussian distribution (for set size 4 and set size 8 trials, respectively). The parameter estimates returned by the discrete-resource model are given in Table 2 (bottom row). Note that these values are strikingly similar to those obtained in Experiment 1. Finally, we also replicated the correlation between estimates of selection and WM capacity observed in Experiment 1 ($r=.48$, $p=.04$; set size 4 trials only). Thus, the results of this experiment were broadly consistent with those of Experiment 1, which suggests that the discrete limits that we observed cannot be attributed solely to masking.

Next, we considered whether the discrete item limits that we observed in our selection task reflect a difficulty in maintaining focus on individual items in the cue array (rather than the ability to *select* these items per se). For example, perhaps participants selected all of the relevant locations after the onset of the cue array (as is predicted by a flexible-resource model) but were incapable of “holding” attention over all of these locations for the entire cue-to-target interval (2,000 ms). To examine this possibility, 13 new participants completed a version of Experiment 1 in which the cue-to-target interval was drastically shortened. The design of this experiment was identical to that of Experiment 1, with the following exceptions. First, set size 8 trials were eliminated. Second, the temporal interval separating the cue and target arrays was shortened to a maximum of 250 ms, and varied unpredictably across trials (between 50, 100, 200, and 250 ms). Third, the WM task was eliminated (this ensured that we could collect sufficient data in each of the cue-to-target interval bins within a single 1.5-h session). Mean (\pm SEM) histograms of report errors are shown for each cue-to-target interval in Fig. 5 (panel A=50 ms, B=100 ms, C=200 ms, D=250 ms). As in Experiment 1, these distributions were well-described by a mixture function (mean adjusted r^2 s=.91, .94, .93, and .84 for the 50-, 100-, 200-, and 250-ms cue-to-target stimulus onset asynchronies, respectively) and poorly described by a Gaussian function (mean adjusted r^2 s=.43, .54, .60, and .65 for the 50-, 100-, 200-, and 250-ms cue-to-target stimulus onset asynchronies, respectively). Bayesian model comparison revealed that the likelihoods of the mixture distribution were 40.22 (± 3.18), 45.61 (± 4.25), 44.07 (± 4.34), and 44.94 (± 7.20) units greater than the likelihood of the Gaussian distribution (for the cue–target stimulus onset asynchronies

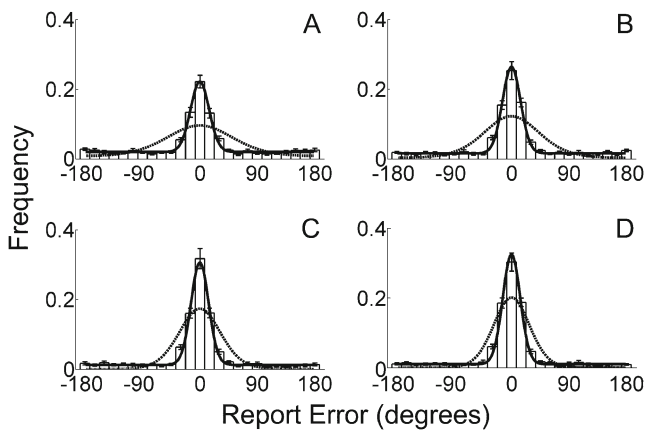


Fig. 5 Results of Control Experiment 2. Panels A–D depict the mean (± 1 SEM) distributions of response errors for the 50-, 100-, 200-, and 250-ms cue-to-target stimulus onset asynchronies, respectively. In all cases, a discrete-resource model (solid lines) outperformed a flexible-resource model (dashed lines)

[SOAs] of 50, 100, 200, and 250 ms, respectively). The parameter estimates returned by the discrete-resource model are shown in Table 3. Briefly, increasing the cue-to-target SOA increased the likelihood that participants would encode the target (i.e., estimates of nr fell as the SOA increased). However, the same manipulation had no discernible effect on estimates of σ . Moreover, the estimates of σ obtained for each SOA were comparable to those observed for set size 4 trials in Experiment 1 (cf. Tables 1 and 3). Thus, the results of this experiment are broadly consistent with those of Experiment 1, which suggests that the item limits that we observed in the selection task cannot be attributed to difficulties in holding attention at cued locations for sustained periods.

Experiment 2

Given that our selection task required observers to briefly hold a single orientation in memory (e.g., during the 750-ms mask interval), one could argue that the apparent item limits we observed reflect well-known item limits in WM storage. For example, perhaps participants performed our selection task by encoding and storing information from each cued location, then executing a search through memory to find the appropriate representation when the probe was presented. This strategy

Table 3 Mean (\pm SEM) parameter estimates returned by the discrete-resource model in Control Experiment 2

	50 ms	100 ms	200 ms	250 ms
σ	14.93 (0.67)	14.42 (0.61)	14.19 (0.70)	14.48 (0.89)
nr	0.49 (0.05)	0.39 (0.05)	0.30 (0.06)	0.25 (0.05)

Values of σ are in degrees

seems farfetched for at least two reasons: First, relative to the WM task (in which observers were required to encode and remember four or eight unique orientations), the memory load imposed by the selection task was relatively low—participants were only required to remember the orientation of a single stimulus. Second, the substitution mask that followed the target array also served as a 100 %-valid postcue regarding the target’s location, obviating any need for participants to store information from irrelevant locations. Nevertheless, if participants performed the “selection” task using a mnemonic strategy, it could very well explain the strong correlation between estimates of selection and WM capacity shown in Fig. 3. Thus, we felt it prudent to conduct an additional experiment to examine this possibility. In Experiment 2, we compared observers’ performance on the “single-target” selection task used in Experiment 1 (i.e., in which participants were cued to attend a variable number of locations and required to discriminate the orientation of a single target) with performance on a “multiple-target” task that required participants to encode and store information from a variable number of cued locations. Our reasoning was that if participants performed the single-target task by encoding and storing information from each cued location, then requiring participants to use this strategy in a multiple-target task should have a negligible effect on their performance.

Method

Participants A group of 18 undergraduate students were tested in a single 1.5-h session in exchange for course credit. All of the participants gave written and oral informed consent, and all reported normal or corrected-to-normal visual acuity and color vision.

Stimuli and procedure Each participant completed a set-size-4-only version of the selection and WM tasks described in Experiment 1. Additionally, participants completed a “multiple” target variant of the selection task. Like the “single-target” task, each trial began with a cue array that instructed observers to attend four locations. Next, we presented radial targets in each cued location. Participants knew that they would be asked to report one of these items, but the location of the probed item was unknown in advance. The target display was followed by a 750-ms delay, during which each cued location was surrounded by a four-dot substitution mask. At the end of the trial, participants were probed to report a single target item (as in Exp. 1).

Results and discussion

Figure 6 depicts the mean (± 1 SEM) distributions of report errors across participants for the single-target (A), multiple-

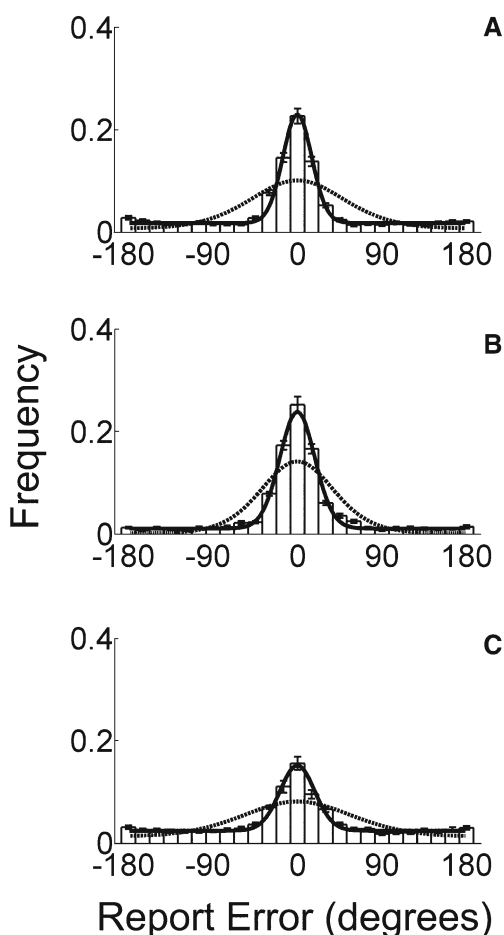


Fig. 6 Results of Experiment 2. Panels A, B, and C depict the mean (± 1 SEM) distributions of report errors during the single-target, multiple-target, and working memory tasks, respectively. The best-fitting flexible- (dashed lines) and discrete- (solid lines) resource models are overlaid

target (B), and WM (C) tasks. For all three tasks, the report error distributions were well-approximated by a mixture distribution (mean adjusted r^2 s = .90, .77, and .92, for the single-target, multiple-target, and WM tasks, respectively) and poorly described by a Gaussian function (adjusted r^2 s = .48, .46, and .68, respectively). Bayesian model comparison revealed that the discrete-resource model outperformed the flexible-resource model by 45.00 (± 4.39), 20.50 (± 3.32), and 31.39 (± 3.28) units for the single-target, multiple-target, and WM tasks (respectively). The parameter estimates returned by this model for each task are shown in Table 4. Critically, estimates

Table 4 Mean (\pm SEM) parameter estimates returned by the discrete-resource model for each task in Experiment 2

	σ	nr
Single target	16.18 (0.72)	0.43 (0.04)
Multiple target	20.19 (1.28)	0.58 (0.04)
WM	19.43 (0.92)	0.26 (0.04)

Values of σ are in degrees

of selection capacity were substantially lower in the multiple-target task ($M=2.27$) than in the single-target task ($M=1.65$), $t(17)=4.38$, $p<.001$, whereas estimates of σ were larger in the multiple-target ($M=20.19^\circ$) than in the single-target ($M=16.18^\circ$) task, $t(17)=4.08$, $p<.001$. Thus, explicitly requiring participants to encode and store information from each cued location had a deleterious effect on their performance.

Experiment 3

To provide additional evidence that performance in our selection task was limited by visual selection per se rather than by memory demands, we conducted another experiment in which we measured an electrophysiological marker of selective attention: the N2pc ERP component. Briefly, the N2pc is a transient contralateral negative wave appearing over posterior electrode sites approximately 200 ms after stimulus onset, which has been localized to generators in extrastriate cortex, including V4 and posterior regions of inferior temporal cortex (Hopf et al., 2000). Moreover, recent evidence has suggested that the N2pc amplitude scales with attentional demands before reaching an asymptotic limit with behavioral estimates of subitizing (e.g., Ester et al., 2012), visual search (Anderson et al., 2013), multiple-object tracking (Drew & Vogel, 2008), and WM (Anderson et al., 2011) capacity. In one recent example, Anderson et al. (2013) examined set-size-dependent changes in N2pc amplitudes while participants performed a speeded search task (see Ester et al., 2012, for similar findings in the context of a subitizing task). On each trial, participants were simply asked to report the orientation of a target “L” presented among a variable number of tilted “T”s as quickly and as accurately as possible. On each trial, the presentation of the search display evoked a large N2pc response. Critically, the amplitudes of these evoked responses increased monotonically with set size (i.e., the total number of to-be-searched items) for displays containing one, two, or three items, but reached an asymptotic limit for displays containing four or more targets. This response profile was well-approximated by a bilinear function, and individual differences in the inflection point of these functions were strongly correlated with search slopes. This result is naturally accommodated by a discrete selection model in which participants sequentially select and search sets of N items (where N is typically ≤ 4). Under this scheme, participants who can select and search more items at any instant will, on average, take less time to find the target. Moreover, the fact that each participant’s N2pc-by-set-size response profile correlated with search slopes suggests that this component reflects—at least in part—the N items that observers are capable of selecting at an instant.

In Experiment 3, we examined set-size-dependent changes in N2pc amplitudes evoked by the presentation of a cue

display (appearing 500–700 ms prior to the onset of any to-be-encoded or remembered information) using a design similar to that of Experiment 1. A reliable relationship between cue-evoked N2pc responses and participants' behavioral performance would further support the conclusion that participants' performance is determined by the selection of cued locations rather than the subsequent storage of target and distractor information at these locations.

Method

Participants A group of 26 adults (aged 18–34) participated in a single 2.5-h session in exchange for monetary compensation (\$25). All of the participants gave both written and oral informed consent and reported normal or corrected-to-normal visual acuity and color vision. The data from two participants were discarded due to excessive eye movement artifacts; the data reported here reflect the remaining 24 participants.

Design and procedure The design and procedure were similar to those of Experiment 1. A representative trial is depicted in Fig. 7. Prior to beginning the experiment, each participant was randomly assigned to one of two cue color conditions: red or green (luminance matched at 23 cd/m²). On each trial, one, two, or four cue circles appeared along the perimeter of an imaginary circle (radius 5°) in one hemifield of the display. An equivalent number of distractor items were presented in corresponding locations in the opposite visual field. No items were presented in uncued locations. The cue array was presented for a randomly chosen interval between 500 and 700 ms, and was followed by the presentation of a target array for 100 ms. The target was followed by a 400-ms substitution mask and the presentation of a randomly oriented probe at the target's location. As in Experiment 1, participants adjusted the orientation of this probe until it matched their percept of the target's orientation.

Change detection task To examine whether the correlations between estimates of selection and WM capacity would generalize to this new experiment, each participant completed a short (approximately 15-min) change detection task similar to one described by Luck and Vogel (1997). Each trial began with the presentation of a "sample" display containing four or eight colored squares (each subtending 1.1°) for 100 ms. The color of each square was randomly chosen with replacement from a set including red, green, yellow, blue, black, and white, with the constraint that no color appeared more than twice. The stimuli were presented within a 12°×9° window centered at fixation, with the following constraints: (1) Stimuli were evenly spaced across the four quadrants of the display (i.e., during set size 4 trials, a single square appeared in each quadrant, and during set size 8 trials, two squares appeared

in each quadrant), and (2) stimuli were separated by a minimum of 2.1°.

Following a 1,000-ms blank interval, a single "probe" square replaced one of the sample items. On 50 % of the trials, the probe had the same color as the item that it replaced; on the remaining 50 % of trials, the probe was assigned a different color (randomly chosen from the aforementioned set of six colors). Participants indicated via keyboard press whether the probe was the same color as the corresponding sample item (z ="yes," ?="no"). Participants were encouraged to prioritize accuracy, and no response deadline was imposed. Each participant completed a total of three blocks of 48 trials. Short rest periods were provided between blocks.

To estimate WM capacity, we relied on an analytical solution developed by Pashler (1988) and refined by Cowan (2001). Here, capacity (K) is defined as

$$K = N * [HR_N - (1 - FA_N)], \quad (3)$$

where N is the number of items in the sample display, HR is the participant's hit rate (or accuracy on "different" trials) for displays containing N items, and FA is the participant's false alarm rate.

Electrophysiological recording and analysis ERPs were recorded using standard recording and analysis procedures that have been described in detail elsewhere (McCullough, Machizawa, & Vogel, 2007). Participants were asked to hold fixation and to refrain from blinking. Trials contaminated by blocking, blinks, or large (>1°) eye movements were excluded from all analyses. We recorded from 22 electrodes spanning the scalp, including the International 10–20 Standard sites F3, F4, C3, C4, P3, P4, O1, O2, PO3, PO4, T5, and T6, as well as nonstandard sites OL and OR (midway between O1/2 and T5/6). The horizontal electrooculogram (EOG) was recorded from electrodes placed 1 cm to the left and right of the external canthi, and the vertical EOG was recorded from an electrode beneath the right eye referenced to the left mastoid. EEG and EOG were amplified by an SA Instrumentation amplifier with a bandpass of 0.01–80 Hz (half-power cutoff, Butterworth filters) and were digitized at 250 Hz by a PC-compatible microcomputer.

Contralateral waveforms were computed by averaging the activity recorded over the right hemisphere when participants were required to select items in the left visual hemifield, and vice versa for trials on which the to-be-selected items appeared in the right visual hemifield. This was done separately for each set size (1, 2, or 4). The N2pc was defined as the difference in mean amplitudes between contralateral and ipsilateral waveforms averaged across electrode sites O1/2, OL/R, and PO3/4 during a period extending from 175 to 250 ms following the onset of the cue array.

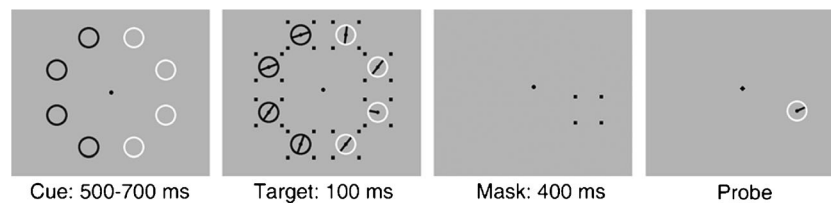


Fig. 7 Selection task used in Experiment 3. Prior to beginning the task, participants were instructed to attend red (shown here in white) or green (shown in black) circles. Each trial began with the presentation of one, two, or four cue circles in one hemifield of the display; an equivalent

number of distractors were presented in the opposite field. The target array was then rendered for 100 ms; as in Experiment 1, participants were instructed to discriminate the orientation of a single radial target that was presented in one of the cued locations

Results and discussion

As in Experiments 1 and 2, performance on the behavioral task was described better by a discrete-resource model (Fig. 8A; mean adjusted $r^2 = .94 \pm .02$ for set size 4; set sizes 1 and 2 were not modeled, due to near-ceiling performance) than by a flexible-resource model (adjusted $r^2 = .50 \pm .03$). Moreover, estimates of selection capacity were strongly predictive of estimates of WM capacity obtained from the color change detection task (Fig. 8B; $r = .57, p < .01$).

As is shown in Fig. 9 (top), the presentation of the cue array evoked a robust N2pc whose amplitude increased monotonically with set size. To examine the putative relationships

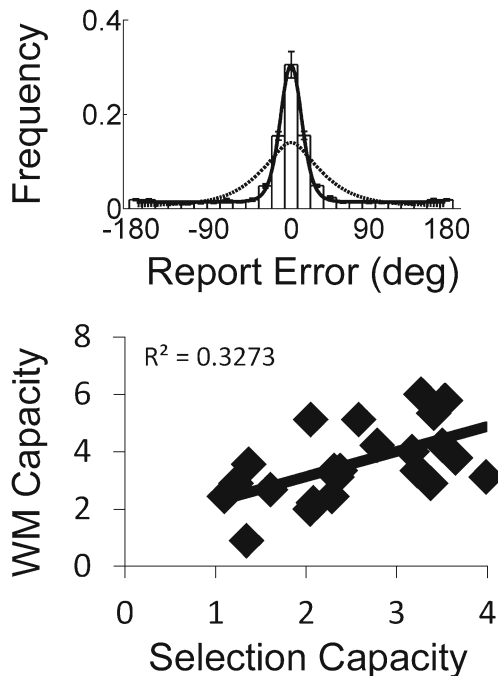


Fig. 8 Behavioral performance in Experiment 3. Panel A depicts the mean distribution of response errors (± 1 SEM) observed during set size 4 trials of a lateralized variant of a selection task. These data were well-described by a discrete-resource model (solid line) and poorly described by a flexible-resource model (dashed line). Panel B depicts the correlation between estimates of selection (using data from set size 4 trials) and working memory (WM) capacity. Here, WM capacity was estimated using a color change detection procedure similar to the one described by Luck and Vogel (1997)

between behavioral performance and N2pc amplitudes, participants were first divided into high- and low-capacity groups on the basis of a median split (in which the median capacity was 2.48 items) of selection capacity at set size 4. The mean N2pc amplitudes are plotted as a function of group (high vs. low capacity; $M = 3.33$ and 1.85 , respectively) and set size in Fig. 9, middle panel. A 2 (capacity group) \times 3 (set size) analysis of variance (ANOVA) with Capacity Group as the sole between-subjects factor on mean N2pc amplitudes revealed a main effect of set size, $F(2, 20) = 11.86, p < .001$, no main effect of group, $F(1, 10) = 3.12, p = .10$, and a marginal interaction between these factors, $F(2, 20) = 3.27, p = .059$. Planned comparisons revealed that the low- and high-capacity groups did not differ when only one or two items had to be selected, $t(11) = 1.87, p = .08$, and $t(11) = 0.40, p = .70$, for set sizes 1 and 2, respectively. However, the mean N2pc amplitudes were larger in the high- ($M = -0.59$) than in the low- ($M = -0.23$) capacity groups on set size 4 trials, $t(11) = 3.09, p = .01$. Moreover, interparticipant variability in selection capacity was strongly correlated with N2pc amplitudes on set size 4 trials (Fig. 9, bottom), such that individuals with a larger selection capacity tended to have higher N2pc amplitudes ($r = .51, p = .01$). Thus, a cue-evoked electrophysiological response occurring approximately 300–500 ms before the onset of the target array—well before there was any need to encode or store any information—was a reliable predictor of participants' behavioral performance. This strongly suggests that performance in this task was limited by each participant's ability to select and individuate information from these locations rather than by their ability to encode or store this information.

Experiment 4

One could argue that the N2pc set size effects documented in Experiment 3 reflect factors unrelated to the selection of cued locations (e.g., perhaps they reflect a general preparatory response; though see Anderson et al., 2011, 2013; Drew & Vogel, 2008; and Ester et al., 2012, for additional evidence suggesting that N2pc amplitudes are sensitive to attentional

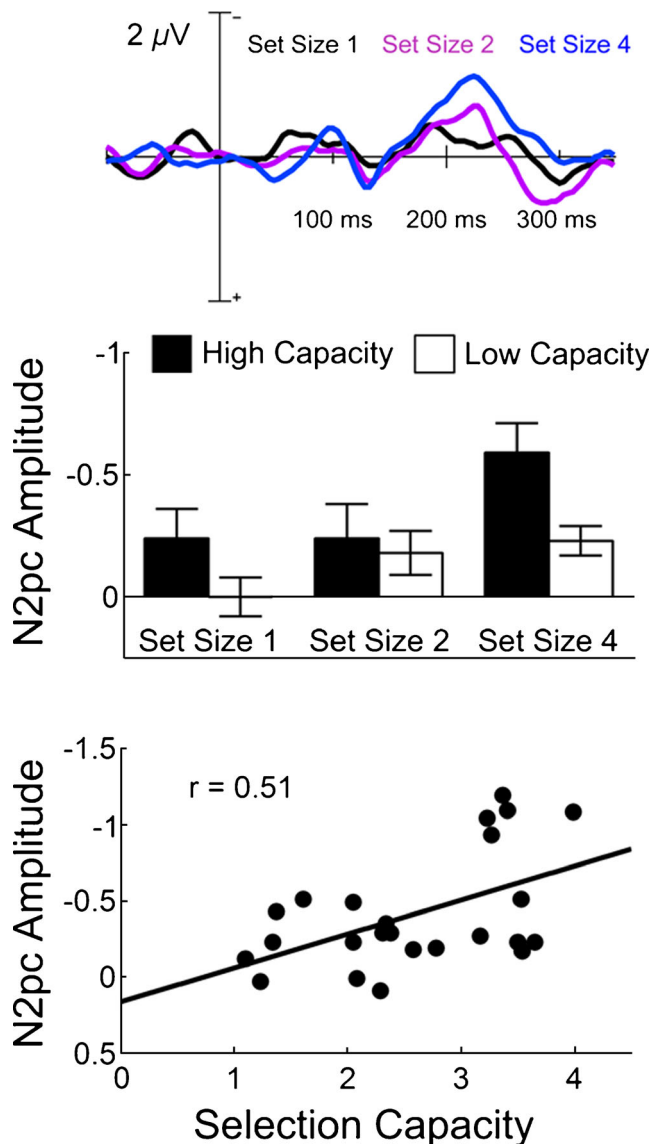


Fig. 9 Cue-evoked N2pc responses observed in Experiment 3. Top panel: Grand-averaged electroencephalographic waveforms time-locked to the onset of the cue array. An N2pc (a transient negative response) emerges approximately 175 ms following the onset of the array. Note that the N2pc amplitudes increase monotonically with set size. Middle panel: To examine whether individual differences in N2pc amplitudes predicted participants' performance on this task, we first divided participants into low- and high-capacity groups on the basis of a median split of capacity estimates. We then plotted the mean N2pc amplitudes (using a window from 175 to 250 ms following the onset of the target array) as a function of capacity group. No group differences in amplitudes were observed at set sizes 1 and 2, presumably because putative capacity limits in selection were not exceeded. However, a reliable difference in N2pc amplitudes across groups was observed at set size 4. Error bars represent ± 1 SEM. Bottom panel: Estimates of selection capacity (abscissa) were strongly correlated with the N2pc amplitudes observed during set size 4 trials (ordinate). Thus, the amplitude of an electrophysiological response that occurred 300–500 ms prior to the onset of the target array—well before participants were required to encode or store any information from the display—was strongly predictive of the participants' performance on this task

demands). In Experiment 4, we tested this possibility by eliminating any need for the observers to select items in the cue array. The task used in Experiment 4 was identical to that used in Experiment 3, with the exception that all nontarget items were removed from the target array. This obviated any need for the observer to select items in the cue array: The lone target would simply “pop out.” Our reasoning was that if the cue-locked N2pc set size effect observed in Experiment 3 reflected a preparatory response, then this effect should also manifest in the present experiment when selection demands were minimized. Alternately, if this effect reflected the number of objects or locations in the cue array that an observer had selected, then it should be abolished.

Method

Participants A group of 14 adults (ages 18–33) completed a single 2.5-h testing session in exchange for monetary compensation. None of the observers who participated in this experiment had participated in Experiment 3. All of the observers gave both written and oral consent, and all reported normal or corrected-to-normal visual acuity and color vision.

Design and procedure The design of Experiment 4 was identical to that of Experiment 3, with the exception that distractors were removed from the target array (i.e., the target array contained only the target). The acquisition and analysis of EEG data were identical to the procedures described for Experiment 3.

Results and discussion

Observers' behavioral performance was near ceiling, with very few high-magnitude errors. Because of this, observers' response profiles were well-approximated by both the flexible (adjusted $r^2 = .95$) and discrete (adjusted $r^2 = .98$) models, and the mean selection capacity (at set size 4) was approximately 3.95 items. Thus, deleting nontarget items from the target array greatly improved observers' task performance relative to Experiment 3. Grand-averaged ERP waveforms time-locked to the onset of the cue array as a function of set size are depicted in Fig. 10, top panel. To compare the results of this experiment with those of Experiment 3, N2pc amplitudes (defined as the mean response during a period 175–250 ms following the onset of the cue array, or roughly 125 ms following the onset of the initial sensory response) were initially submitted to a 2 (Exp. 3 vs. Exp. 4) \times 3 (set size) mixed-model ANOVA with Experiment as the sole between-group factor. This analysis revealed a marginal main effect of experiment, $F(1, 36) = 3.67, p = .06$, a main effect of set size, $F(2, 72) = 6.28, p = .003$, and an interaction between these two

factors, $F(2, 72)=5.23$, $p=.007$. Mean N2pc amplitudes are plotted as a function of set size and Experiment (3 vs. 4) in Fig. 10, bottom. Note that in contrast to the monotonic set size effects observed in Experiment 3 (black bars), no N2pc set size effect was observed in the present experiment. Because the only difference between the tasks used in Experiments 3 and 4 concerned whether observers would need to select cued locations in advance, this strongly suggests that the N2pc set size effect observed in Experiment 3 reflected the selection of items in the cue array.

General discussion

Our findings support the view that observers can select only a handful of items at an instant. Specifically, we found that performance in an attention-demanding selection task that required participants to select and monitor multiple locations was well-described by a discrete-resource model that assumes a fixed limit in the number of locations that can be selected. Conversely, a flexible-resource model that allowed the entire target array to be selected provided a relatively poor

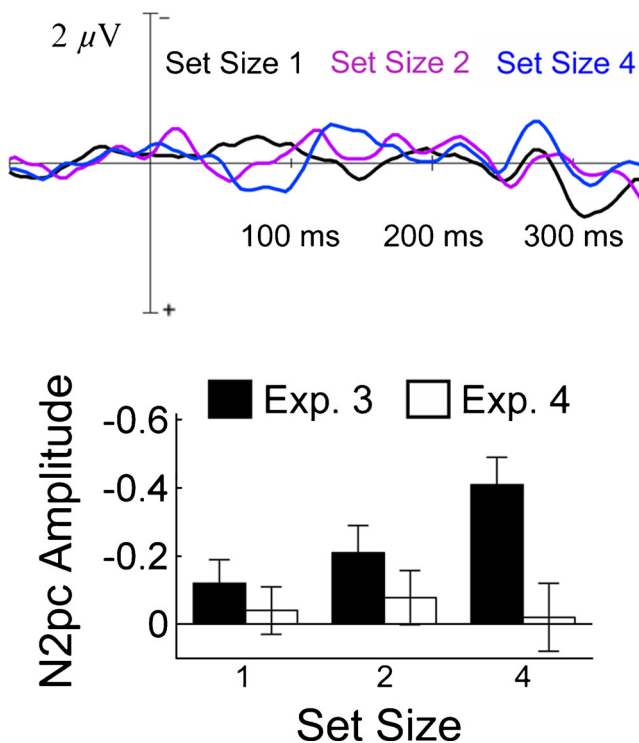


Fig. 10 Results of Experiment 4. Top panel: Grand-averaged waveforms time-locked to the onset of the cue array, in a variant of the task used in Experiment 2 in which we omitted distractors from the target array. Under these conditions, the N2pc observed in Experiment 3 was virtually eliminated. Bottom panel: Comparison of mean N2pc amplitudes observed across set sizes in Experiments 3 (black bars) and 4 (white bars). Error bars represent ± 1 SEM

description of these data. Of course, one explanation for the discrete resource limit observed in our selection task is that the number of selected locations actually increased with set size (consistent with a flexible-resource model), but a subsequent limit in WM storage (in which capacity is thought to be limited by a discrete resource—see, e.g., Zhang & Luck, 2008; though see Van den Berg, Shin, Chou, George, & Ma, 2012, for an alternative view) created the appearance of a discrete resource limit. However, Experiment 3 revealed a robust correlation between individual differences in selection capacity and the amplitude of a cue-evoked N2pc appearing 300–500 ms prior to the onset of the to-be-encoded target. This strongly suggests that performance in our selection task was limited by participants' ability to select and individuate multiple locations rather than by their ability to encode and store information from these locations.

Our interpretation of these data rely upon the assumption that the uniform component of the mixture model reflects trials in which the participant fails to select the target location and is forced to guess randomly. However, a previous study (Bays, Catalao, & Husain, 2009) suggested that a large proportion of these random responses can be attributed to binding errors, in which the participant erroneously reports a feature value from a nontarget location. To explore this possibility, we generated a distribution of response errors relative to each distractor orientation in the cued locations during the set size 4 trials of Experiment 3. We found no evidence of a central tendency in these distributions (Fig. 11), suggesting that binding errors were not a major determinant of task performance.

Our findings are reminiscent of multiple studies also suggesting that WM capacity is determined by a discrete resource (Anderson et al., 2011; Anderson & Awh, 2012; Awh, Barton, & Vogel, 2007; Barton, Ester, & Awh, 2009; Luck & Vogel, 1997; Rouder et al., 2008; Vogel & Machizawa, 2004; Zhang & Luck, 2008). In addition, a substantial amount of evidence suggests that both visual selection and WM storage are mediated by similar neural mechanisms. For example, similar cortical regions are recruited during both sustained attention and WM maintenance (Awh & Jonides, 2001). Moreover,

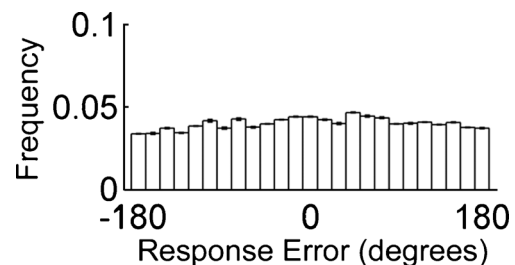


Fig. 11 Distribution of response errors relative to the distractor orientations observed in Experiment 3. The data are from set size 4 trials and have been pooled across the three distractors presented in the target hemifield. Error bars represent ± 1 SEM

activity in posterior parietal regions such as intraparietal sulcus has been shown to scale with the number of items that are being attended (Cusack, Mitchell, & Duncan, 2010; Mitchell & Cusack, 2008) or held in WM (Xu & Chun, 2005). Together, these findings raise the possibility that visual selection and WM storage are dependent on a common, discrete resource (Anderson et al., 2013; Ester et al., 2012). Here, we provide evidence supporting this view. Specifically, individual differences in selection capacity were strongly correlated with corresponding differences in WM capacity. Critically, this relationship cannot be explained by WM contributions to our selection task: The amplitude of a cue-evoked N2pc component—appearing approximately 300–500 ms before any to-be-encoded stimulus—was a robust predictor of behavioral performance.

The flexible-resource model scrutinized here assumes that encoding and memory precision are constant (i.e., resources are divided equitably between each of the cued stimuli) across stimuli and trials. However, some recent work (Mazyar, van den Berg, & Ma, 2012; Van den Berg et al., 2012) has reported that precision varies considerably across stimuli and trials. Thus, one could characterize our flexible-resource model as a “straw man” and argue that a suitably complex model (e.g., one in which precision varies across stimuli and trials, or one that incorporates recent history—i.e., which location contained the target on the previous trial) could easily match or better a discrete-resource model. However, note that we reported strong links between capacity estimates obtained via a discrete-resource model and a neural measure of attentional processing (the N2pc ERP component). Prior work has demonstrated that this component reaches an asymptotic limit once putative item limits in multiple-object tracking (e.g., Drew & Vogel, 2008), subitizing (Ester et al., 2012), visual search (Anderson et al., 2013), and WM storage (e.g., Anderson et al., 2011) have been reached. It is unclear how a flexible-resource model—however complex—could account for any these findings, since a fundamental prediction of this class of models is that all items within a display are selected (or remembered), albeit with dwindling levels of precision. Thus, the extant neural data provide strong converging evidence for discrete-resource models of perception and visual short term memory.

Previous theoretical (Lisman & Idiart, 1995; Raffone & Wolters, 2001) and experimental (Liebe, Hoerzer, Logothetis, & Rainer, 2012; Sauseng et al., 2009; Siegel, Warden, & Miller, 2009) work has suggested that the storage of multiple items in WM is mediated by a phase-coding scheme. Here, each item held in WM is represented through a unique pattern of high-frequency, synchronous firing across large populations of neurons. When multiple items must be held in memory, the high-frequency activity related to each remembered item may be multiplexed within distinct phases of slower oscillatory activity. One attractive aspect of this phase-

coding scheme is that it provides a relatively straightforward explanation of the discrete capacity limits that have been reported in numerous studies of WM. For example, if information from each selected location must be segregated from the others in a different range of phase orientations, there should be a maximum number of locations that could be distinctly represented at once. Our suspicion is that a similar phase-dependent coding scheme mediates the selection of multiple locations, though further research will be needed to evaluate this possibility. One attractive aspect of this account is that it implies that the discrete resource limits observed in visual selection and WM storage are due to basic biophysical limitations in how information can be represented in the brain.

In summary, although multiple past studies have documented clear capacity limits in the online selection of visual information, conclusive evidence regarding the nature of the limiting resource has been elusive. The findings reported here are consistent with a fixed item limit in visual attention. Moreover, the item limits that were evident in this visual selection task were strongly correlated with the item limits that were observed in a parallel test of storage capacity in visual WM, suggesting that visual selective attention and working memory may be constrained by a common discrete resource.

Author note Supported by Grant No. NIH R01-MH087214 to E.A. and E.K.V. E.F.E. and E.A. conceived and designed all experiments. E.F.E. and L.M.M. collected the data and analyzed them using software developed by K.F. Author E.K.V. provided guidance and assistance with electrophysiological recordings. E.F.E. and E.A. wrote the manuscript.

References

- Alvarez, G. A., & Cavanagh, P. (2005). Independent resources for attentional tracking in the left and right visual hemifields. *Psychological Science*, *16*, 637–643.
- Alvarez, G. A., Gill, J., & Cavanagh, P. (2012). Anatomical constraints on attention: Hemifield independence is a signature of multifocal spatial selection. *Journal of Vision*, *12*(5), 9. doi:10.1167/12.5.9
- Anderson, D. E., & Awh, E. (2012). The plateau in mnemonic resolution across large set sizes indicates discrete resource limits in visual working memory. *Attention, Perception, & Psychophysics*, *74*, 891–910. doi:10.3758/s13414-012-0292-1
- Anderson, D. E., Vogel, E. K., & Awh, E. (2011). Precision in visual working memory reaches a stable plateau when individual item limits are exceeded. *Journal of Neuroscience*, *31*, 1128–1138. doi:10.1523/JNEUROSCI.4125-10.2011
- Anderson, D. E., Vogel, E. K., & Awh, E. (2013). A common discrete resource for visual working memory and visual search. *Psychological Science*, *24*, 929–938.
- Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, *18*, 622–628. doi:10.1111/j.1467-9280.2007.01949.x
- Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, *5*, 119–126. doi:10.1016/S1364-6613(00)01593-X

- Awh, E., & Pashler, H. (2000). Evidence for split attentional foci. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 834–846.
- Barton, B., Ester, E. F., & Awh, E. (2009). Discrete resource allocation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 1359–1367. doi:10.1037/a0015792
- Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, 9(10):7, 1–11. doi: 10.1167/9.10.7
- Bergen, J. R., & Julesz, B. (1983). Parallel versus serial processing in rapid pattern discrimination. *Nature*, 303, 696–698.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–436. doi:10.1163/156856897X00357
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87–114. doi:10.1017/S0140525X01003922. disc. 114–185.
- Cusack, R., Mitchell, D. J., & Duncan, J. (2010). Discrete object representation, attention switching, and task difficulty in the parietal lobe. *Journal of Cognitive Neuroscience*, 22, 32–47.
- Drew, T., & Vogel, E. K. (2008). Neural measures of individual differences in selection and tracking multiple moving objects. *Journal of Neuroscience*, 28, 4183–4191. doi:10.1523/JNEUROSCI.0556-08.2008
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96, 433–458. doi:10.1037/0033-295X.96.3.433
- Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science*, 8, 135–139. doi:10.1111/j.1467-9280.1997.tb00696.x
- Ester, E. F., Drew, T., Klee, D., Vogel, E. K., & Awh, E. (2012). Neural measures reveal a fixed item limit in subitizing. *Journal of Neuroscience*, 32, 7169–7177.
- Franconeri, S. L., Alvarez, G. A., & Enns, J. T. (2007). How many locations can be selected at once? *Journal of Experimental Psychology: Human Perception and Performance*, 33, 1003–1012. doi:10.1037/0096-1523.33.5.1003
- Hopf, J., Luck, S. J., Girelli, M., Hagner, T., Mangun, G. R., Scheich, H., & Heinze, H. J. (2000). Neural sources of focused attention in visual search. *Cerebral Cortex*, 10, 1233–1241.
- Huang, L., & Pashler, H. (2005). Attention capacity and task difficulty in visual search. *Cognition*, 94, B101–B111. doi:10.1016/j.cognition.2004.06.006
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24, 175–219. doi:10.1016/0010-0285(92)90007-O
- Kramer, A. F., & Hahn, S. (1995). Splitting the beam: Distribution of attention over noncontiguous regions of the visual field. *Psychological Science*, 6, 381–386.
- Liebe, S., Hoerzer, G. M., Logothetis, N. K., & Rainer, G. (2012). Theta coupling between V4 and prefrontal cortex predicts visual short-term memory performance. *Nature Neuroscience*, 15, 456–462.
- Lisman, J. E., & Idiart, M. A. P. (1995). Storage of 7±2 short-term memories in oscillatory subcycles. *Science*, 267, 1512–1515.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279–281. doi: 10.1038/36846
- MacKay, D. J. (2003). *Information theory, inference, and learning algorithms*. Cambridge: Cambridge University Press.
- Malinowski, P., Fuchs, S., & Müller, M. M. (2007). Sustained division of spatial attention to multiple locations within one hemifield. *Neuroscience Letters*, 414, 65–70.
- Mazyar, H., van den Berg, R., & Ma, W. J. (2012). Does precision decrease with set size? *Journal of Vision*, 12(6), 10. doi:10.1167/12.6.10
- McCullough, A. W., Machizawa, M. G., & Vogel, E. K. (2007). Electrophysiological measures of maintaining representations in visual working memory. *Cortex*, 43, 77–94.
- McMains, S. A., & Somers, D. C. (2004). Multiple spotlights of attentional selection in human visual cortex. *Neuron*, 42, 677–686.
- McMains, S. A., & Somers, D. (2005). Processing efficiency of divided spatial attention mechanisms in human visual cortex. *Journal of Neuroscience*, 25, 9444–9448.
- Mitchell, D. J., & Cusack, R. (2008). Flexible, capacity-limited activity of posterior parietal cortex in perceptual as well as visual short-term memory tasks. *Cerebral Cortex*, 18, 1788–1798.
- Müller, M. M., Malinowski, P., Gruber, T., & Hillyard, S. A. (2003). Sustained division of the attentional spotlight. *Nature*, 424, 309–312.
- Pashler, H. (1988). Familiarity and visual change detection. *Perception and Psychophysics*, 44(4), 369–378.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442. doi:10.1163/156856897X00366
- Raffone, A., & Wolters, G. (2001). A cortical mechanism for binding in visual working memory. *Journal of Cognitive Neuroscience*, 13, 766–785.
- Rensink, R., & Enns, J. T. (1995). Preemption effects in visual search: Evidence for low-level grouping. *Psychological Review*, 102, 101–130. doi:10.1037/0033-295X.102.1.101
- Rouder, J. N., Morey, R. D., Cowan, N., Zwilling, C. E., Morey, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences*, 105, 5975–5979.
- Sagi, D., & Julesz, B. (1985). “Where” and “what” in vision. *Science*, 228, 1217–1219.
- Sauseng, P., Klimesch, W., Heise, K. F., Gruber, W. R., Holz, E., & Hummel, F. C. (2009). Brain oscillatory substrates of visual short-term memory capacity. *Current Biology*, 19, 1846–1852. doi:10.1016/j.cub.2009.08.062
- Scharff, A., Palmer, J., & Moore, C. M. (2011a). Evidence of fixed capacity in visual object categorization. *Psychonomic Bulletin & Review*, 18, 713–721. doi:10.3758/s13423-011-0101-1
- Scharff, A., Palmer, J., & Moore, C. M. (2011b). Extending the simultaneous–sequential paradigm to measure perceptual capacity for features and words. *Journal of Experimental Psychology: Human Perception and Performance*, 37, 813–833. doi:10.1037/a0021440
- Siegel, M., Warden, M. R., & Miller, E. K. (2009). Phase-dependent neuronal coding of objects in short-term memory. *Proceedings of the National Academy of Sciences*, 106, 21341–21346.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136. doi:10.1016/0010-0285(80)90005-5
- Trick, L. M., & Pylyshyn, Z. W. (1993). What enumeration studies can show us about spatial attention: Evidence for limited-capacity preattentive processing. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 331–351. doi:10.1037/0096-1523.19.2.331
- Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review*, 101, 80–102. doi:10.1037/0033-295X.101.1.80
- Van den Berg, R., Shin, H., Chou, W. C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences*, 109, 8780–8785.
- Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, 428, 748–751. doi:10.1038/nature02447
- Wasserman, L. (2000). Bayesian model selection and model averaging. *Journal of Mathematical Psychology*, 44, 92–107.
- Xu, Y., & Chun, M. M. (2005). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, 440, 91–95. doi:10.1038/nature04262

- Xu, Y., & Chun, M. M. (2009). Selecting and perceiving multiple visual objects. *Trends in Cognitive Science*, *13*, 167–174. doi:[10.1016/j.tics.2009.01.008](https://doi.org/10.1016/j.tics.2009.01.008)
- Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*, 233–235. doi:[10.1038/nature06860](https://doi.org/10.1038/nature06860)
- Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychological Science*, *20*, 423–428. doi:[10.1111/j.1467-9280.2009.02322.x](https://doi.org/10.1111/j.1467-9280.2009.02322.x)
- Zhang, W., & Luck, S. J. (2011). The number and quality of representations in working memory. *Psychological Science*, *22*, 1434–1441. doi:[10.1177/0956797611417006](https://doi.org/10.1177/0956797611417006)