Journal of Experimental Psychology:
Human Perception and Performance
2007, Vol. 33, No. 5, 1003–1012

0096-1523/07/$12.00 DOI: 10.1037/0096-1523.33.5.1003

How Many Locations Can Be Selected at Once?

Steven L. Franconeri
Northwestern University

George A. Alvarez
Massachusetts Institute of Technology

James T. Enns
University of British Columbia

The visual system uses several tools to select only the most relevant visual information for further processing, including selection by location. In the present study, the authors explored how many locations can be selected at once. Although past evidence from several visual tasks suggests that the visual system can operate on a fixed number of 4 objects or locations at once, the authors found that this capacity varies widely in response to the precision of selection required by the task. When the authors required precise selection regions, only 2–3 locations could be selected. But when the selection regions could be coarser, up to 6–7 locations could be selected. The authors discuss potential mechanisms underlying the selection of multiple locations and review the evidence for fixed limits in visual attention.

Keywords: visual attention, selection, capacity, span, location

At any moment in time, the visual world presents a vast array of visual information to the eyes. The visual system has several tools available to help it select the information that is most relevant to current goals and behavior. One such tool is selection by features. People are able to select regions of the visual field that contain certain simple properties, such as a color (Egeth, Virzi, & Garbart, 1984; Friedman-Hill & Wolfe, 1995), direction of motion (Saenz, Buracas, & Boynton, 2003), or sudden luminance change (Belopolsky, Theeuwes, & Kramer, 2005; Donk & Theeuwes, 2001). Another tool, selection by surfaces, seems to operate more automatically. When individuals attend a part of an object or an element in a group, selection often seems to spread to the rest of the object or group (e.g., Nakayama, He, & Shimojo, 1995).

The present article explores a third major tool—selection by location. When told that a cue will appear at a future target's location on the majority of trials, participants are more accurate at detecting the presence of the target when the cue is valid relative to when it is invalid, suggesting that a cued location can be preferentially selected over other locations (Posner, 1980). And although past evidence suggests that only a single location can be selected at once (Eriksen & Yeh, 1985; McCormick & Klein, 1990; Posner, 1980), more recent work using improved behavioral methodologies has shown that selection can be split over at least two noncontiguous locations (Awh & Pashler, 2000; Hahn & Kramer, 1998; Kramer & Hahn, 1995).

For example, in one study, participants were briefly presented with a display containing two cued locations, followed by a grid containing letters and two numbers, and participants were asked to report the two numbers. These numbers appeared in the cued locations on the majority of trials, and in the rest of the trials, one number appeared in between the locations and one appeared at another location. Accuracy was high when the numbers appeared at the cued locations, but accuracy was low for numbers appearing in between those locations, suggesting that participants could split selection across two locations without selecting the area in between (Awh & Pashler, 2000). In another study, participants were asked to search for a target object among distractor objects, and when additional distractor objects were added to the displays, response times were longer (Burkell & Pylyshyn, 1997). However, when a set of object locations was cued before the search, performance suggested that participants could, to some degree, isolate their search to those locations. Participants could use one to five cues and possibly more because higher numbers of cues were not tested.

In the present study, we tested a wider range of cue set sizes and showed that location-based selection can sometimes be divided between up to seven locations. But more critically, our results indicate that the number of locations that can be selected is not fixed. Instead, this number depends on how precisely individuals must specify each location. That is, as a participant selects more locations, the locations must each be specified with less precision; conversely, if more precision is required, then fewer locations can be selected. We then discuss possible mechanisms underlying the selection of multiple locations.

Steven L. Franconeri, Department of Psychology, Northwestern University; George A. Alvarez, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology; James T. Enns, Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada.

Steven L. Franconeri was supported by an Isaac Walter Killam Postdoctoral Fellowship and a National Defense Science & Engineering Graduate Fellowship. George A. Alvarez was supported by a National Institutes of Health National Research Science Award. James T. Enns was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada. We thank Patrick Cavanagh, Joan Chiao, and Zenon Pylyshyn for helpful discussion, and Jennifer Charlton and Nigel Kam for their assistance in data collection.

Correspondence concerning this article should be addressed to Steven L. Franconeri, Northwestern University, 2029 Sheridan Road, Evanston, IL 60208. E-mail: franconeri@northwestern.edu

Can the Visual System Select Only a Fixed Number of Spatial Locations?

Evidence that visual selection can be divided over multiple locations is congruent with results from a variety of psychophysical tasks indicating that the visual system can restrict processing in any one moment to a small but fixed number of locations. When asked to enumerate a collection of dots, beans, or shapes, participants are accurate on sets of one, two, three, and even four with little increase in response time (this effect is often called *subitizing*), whereas sets greater than four lead to ploddingly slow response times and much greater error (e.g., Trick & Pylyshyn, 1993). In visual tracking tasks, participants can keep track of about four moving targets among identical distractors (Intriligator & Cavanagh, 2001; Pylyshyn & Storm, 1988; Yantis, 1992). When asked to search for a target letter among a set of distractor letters, participants can restrict their search to a group of letters that appear suddenly, but only if the group has fewer than about four letters (Yantis & Johnson, 1990; Yantis & Jones 1991). Together, these results suggest that something about the architecture of the visual system limits the number of objects that one can access concurrently to only a handful. The visual system might have a fixed number of functional or neural mechanisms each designed to deal with a single object or location (e.g., Cowan, 2001; Pylyshyn, 1989).

In contrast, there might not be a fixed number of objects that the visual system can handle concurrently. But if so, why would performance fall abruptly when participants are asked to deal with more than four objects at once? One possibility is that there is a trade-off between the number of objects that participants can select at once and the level of spatial precision that can be maintained for the location of each one. Perhaps, when a single item is selected, its position or its properties can be encoded very precisely. But as more objects are selected, each item's position or properties must be encoded more coarsely. Eventually, as more items are selected, each item's represented location becomes too coarse, and performance begins to suffer. In each of the tasks described above, perhaps the apparent limit of four items emerges simply because of the spatial precision that is demanded to perform the tasks in these experiments. If lower precision were required, perhaps more locations could be selected, and, conversely, if greater precision were required, perhaps fewer locations could be selected.

Consider for example the subitizing phenomenon in which participants can quickly report the number of items in a collection if there are four or fewer but report the number of larger sets much more slowly, suggesting that participants can simultaneously deal with about four items (e.g., Trick & Pylyshyn, 1993, 1994). According to an alternative account of this result, participants may be estimating the number of objects in a small collection in the same way they would for a large collection of dozens of objects, using a combination of simple visual cues (such as item density and area). This estimation process will become increasingly imprecise as the number of items increases (Whalen, Gallistel, & Gelman, 1999). The coarseness of this estimate might not be a problem when there are fewer than four items but could be a problem when there are more than four items, at which point participants must resort to serial counting (Gallistel & Gelman, 1992). Thus, the sudden change in performance when counting more than four objects could be a product of the inherent imprecision of visual number estimates, which just happens to support estimation of up to about four items for the density of the displays that is typically used.

A limit on spatial precision might also explain why many studies have found that participants can track only four moving objects. Participants might need to know each target object's position with some degree of precision to distinguish them from nearby distractors. When the speed of the objects is varied, which presumably varies the precision of the participant's representation of the target position, participant performance varies smoothly between tracking seven or eight slowly moving targets or only one rapidly moving target (Alvarez & Franconeri, 2007b). In the case of restricting search to a maximum of about four locations (Yantis & Johnson, 1990), further research showed that this limit is due to factors such as the difficulty of the search task—with difficult searches (requiring a more precise scrutiny of each item) diminishing the participant's ability to restrict search to more items (Yantis & Jones, 1991).

Is the Selection of Location Limited by Precision?

There are thus at least two explanations for the performance limitations in counting, object tracking, and search tasks. In each task, performance could suddenly drop at around four items because the visual system is somehow architecturally restricted to dealing with a fixed number of items at once, or it could occur because the precision of each item's representation becomes too coarse at that point. Deciding between these alternatives is fundamentally important for the understanding of the architecture of visual selection.

In the present study, we found first that participants can select about four locations at once when displays are relatively dense but that capacity increases up to seven locations when displays are more sparse (Experiment 1). This finding suggests that requiring a greater spatial precision for the selection decreases the number of locations that can be selected. It also implies that there is a flexible number of locations that can be selected and retained over a brief retention interval. In the General Discussion, we consider possible mechanisms for this trade-off between selection precision and selection capacity.

Experiment 1

We used a modified visual search task to measure how many locations a participant can select at one time. Figure 1 depicts a typical trial sequence. Participants were shown an initial array of small discs, which marked the locations of all potential search items. In addition, between two and six of the disc locations were occupied by plusses (superimposed vertical and horizontal bars). The plusses then disappeared for a brief period before search items consisting of either a single horizontal or vertical bar appeared at every location. The participant's task was to decide whether a vertical bar was present or absent in one of the locations cued by plus signs before the retention interval. To ensure that participants remain focused on only the locations occupied by plus signs, we presented false targets at the uncued set of locations. Thus, if participants lost track of even one cued item, accuracy would begin to drop. As the number of cued items increased, participants would be increasingly likely to lose track of cued items. We refer to the

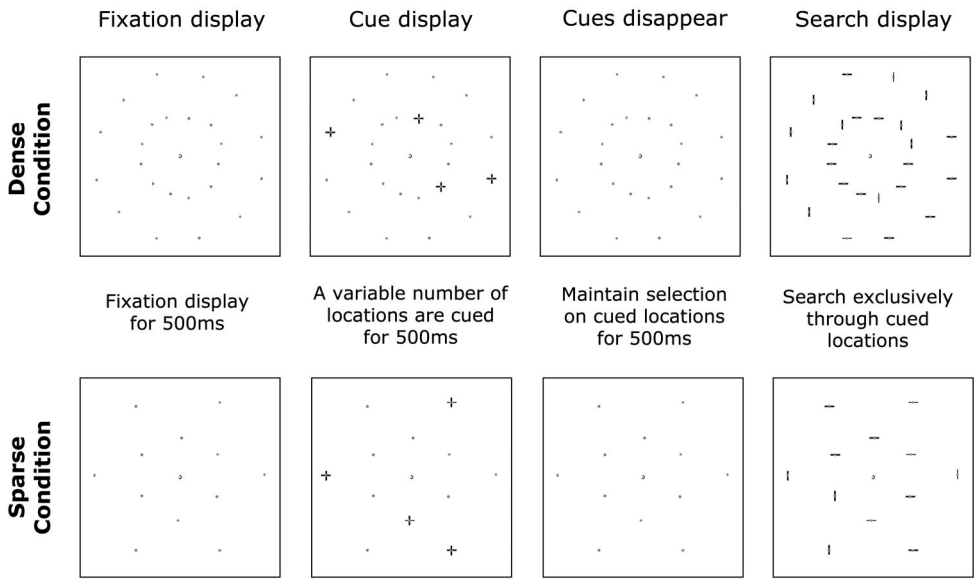

Figure 1. Example search displays in Experiment 1.

point at which accuracy drops below 90% as the participant's *capacity:* the number of locations a participant can remember before performance begins to suffer.

The displays were designed to minimize the distracting transient created by the disappearance of the cues and the appearance of the bars for two reasons. First, we did not want these transient events to disrupt the participant's ability to maintain his or her selection of the cued locations during the retention interval. Second, using cues that create a unique transient at the cued locations may allow participants to rely on a high-capacity but short-lasting representation of the transients that occur at the cued locations (Belopolsky et al., 2005; Donk & Theeuwes, 2001; Yantis & Jones, 1991). To illustrate, imagine using very bright cues, which would leave a brief afterimage at the cued locations, allowing the participant to see the cued locations even after they disappeared. To remove all luminance transients at the display transitions, we designed search items and cues that were perceptually equiluminant to the background (see *Method* section).

The displays were also designed to minimize a participant's ability to chunk multiple locations into a single unit. If all multiple cued locations happened to fall within a small area, without intervening uncued locations, the participant might choose to remember the center of the cluster, instead of each position individually. For example, in one study, participants were asked to detect a change to the contrast polarities in a set of simple objects. Surprisingly, their performance suggested that they were able to remember up to eight or nine locations. However, observers likely memorized the locations of the dark or light objects and chunked multiple locations together (Rensink, 2000). To minimize the usefulness of this strategy, we interleaved potential cue locations with locations that would never be cued. If participants selected these interleaved locations in addition to the cued locations, then their selection would add the locations of false targets to the remembered set locations.

This approach to assessing the number of locations that can be selected at one time is similar to one used by Burkell and Pylyshyn (1997), who found that participants had a capacity of at least five items. However, because Burkell and Pylyshyn did not test larger cue set sizes, it is unclear whether those participants had even higher capacities. In the present experiments, we tested a larger number of cue set sizes to find the upper limit on the number of locations that can be selected. However, we also tested whether that number varies with the degree of spatial precision that is required for each selection. The displays in Experiment 1 contained 24 or 12 possible locations. When the display contained 24 items, the displays were dense, and each item needed to be selected precisely to distinguish it from items close by. When the display contained only 12 items, the displays were sparse, and each item could be selected with less spatial precision. If the number of locations that can be maintained through the retention interval is fixed, then capacity should be the same for both display types. However, if capacity decreases as the precision required for each selection increases, then capacity should be lower when the display locations are more densely distributed.

Method

Participants. Twelve undergraduate students at the University of British Columbia (Vancouver, British Columbia, Canada) participated in exchange for course credit.

Stimuli. Stimuli were displayed on Apple eMac computers with 16-in. (40.6-cm) cathode ray tube monitors. Displays were created and presented using the VisionShell C libraries (http://www.kagi.com/visionshell). Although head position was unrestrained, viewing distance was approximately 50 cm. From this distance, the display subtended 35.3° in width × 26.5° in height and consisted of a blue background (10 cd/m^2), a gray outlined oval fixation point at the center of the display (39 cd/m^2, 0.3° in width and height), and either 24 (dense displays) or 12 (sparse displays) red rectangular bars (53 cd/m^2 on average, 1.03° in length and 0.14° in width), occluded at their centers by a set of gray (39 cd/m^2, 0.28° in diameter) discs. The dots and rectangular

bars were equally spaced along the diameters of two concentric imaginary circles (4.5° and 9° in diameter), with each position jittered randomly by 0.14°–0.28° in either the vertical and horizontal directions. See Figure 1 for sample displays.

The red bars served as both cues and search items, and they disappeared and reappeared during each trial. To prevent their appearance from causing disruptive transients, each participant set the red bars to a luminance that was perceptually equiluminant to the background, using a flicker-minimization procedure. Each participant was presented with an example search display in which the bars (randomly horizontal or vertical) flickered between red and blue (the same luminance as the background) at 15 Hz, and participants adjusted the red luminance with the mouse until the perception of flicker was minimized. Red values were set according to the median of 15 such observations. Because the gray discs were present throughout the trial, they did not need to be perceptually equiluminant to the background.

Procedure. Figure 1 depicts a typical trial sequence. Participants pressed a key to begin each trial. In the *cued* search condition, a blank blue background containing only a fixation point appeared for 400 ms, followed by a fixation display containing only the discs (which marked the potential cue positions) for 400 ms. Next, in the *cue display*, between two and six potential target locations were cued by presenting both horizontal and vertical bars behind the discs at those locations, for 500 ms. On each trial, cues were restricted to either the odd or even potential search locations to minimize the participant's ability to chunk multiple locations into single, larger locations. Then the cues disappeared for 500 ms, but the discs marking all possible search item locations remained on the screen. Results from experiments not reported here indicate that increasing this interval to 1,200 ms leads to only a tiny increase (1.7%) in accuracy. On the search display, vertical or horizontal bars appeared behind every disc, and the participants began their search for a vertical bar within one of the cued locations (either one or none were vertical). The remaining uncued bars were randomly horizontal or vertical, so that if participants could not maintain their selection of the cued locations, they would not be able to discriminate a real target (in the cued group) and a false target (in the uncued group).

Participants pressed the *P* key if the target was present and the *A* key if the target was absent. Participants were informed of the accuracy of each answer by a series of high (correct) or low (incorrect) tones. In the uncued condition, only the cued search items were shown, and all of the other items were hidden from view. Thus, the participant's task was simply to indicate whether a vertical bar was present or absent in the search display. This condition was included so that accuracy and response time in the cued condition, in which participants must select multiple noncontiguous locations, could be compared with a condition in which participants may select one large contiguous region. In both conditions, accuracy was stressed as the participant's first priority, and speed was stressed as the second priority.

The experiment consisted of one practice block and 12 test blocks of 40 trials each (8 trials each of set Sizes 2, 3, 4, 5, and 6). The uncued condition was tested in Blocks 3, 4, 9, and 10, whereas the cued condition was tested in the remaining blocks. Both conditions were tested with sparse and dense displays, with sparse in the odd blocks for odd-numbered participants and sparse in the even blocks for even-numbered participants. The experiment lasted about 45 min, and participants were encouraged to take breaks at any time by postponing the keypress that would start the next trial.

Results and Discussion

Accuracy in the uncued displays was high for all set sizes in both the sparse and dense displays (98% accuracy). Accuracy in the cued condition is shown in Figure 2a and 2b for sample participants with the highest and lowest total accuracy and also for a participant closest to average accuracy. Accuracy decreased with both cued set size and with denser displays. These data were submitted to a 2×5 repeated measures analysis of variance, with density (sparse, dense) and cued set size (2, 3, 4, 5, 6) as variables. Because of the often qualitatively different patterns of performance among participants, tests involving set size showed heterogeneity of variance, and we used a conservative Greenhouse–Geisser correction that lowered the degrees of freedom for these tests. Across both display types, accuracy dropped as the cued set size became larger (98% down to 82%), $F(2.6, 29) = 26.5$, $p < .001$, and participants were more accurate with sparse displays than with dense displays (95% vs. 89%, respectively), $F(1, 11) = 32$, $p < .001$. The effect of cued set size was greater for dense displays than for sparse displays, as evidenced by a significant interaction between display type and cued set size, $F(2.8, 31) = 13.0$, $p < .001$. More specifically, accuracy for both display types started at 98% at the smallest cued set size, but whereas accuracy in sparse displays dropped to only 91% at the largest cued set size, accuracy in dense displays dropped to 74%.

Because accuracy drops with higher set sizes, participants were not able to selectively remember six relevant locations. So how many locations could they retain? Because participants' subjective reports indicated a wide variance in capacities, we examined each participant's accuracy individually by finding the point at which accuracy began to drop below 90% for each participant. This capacity measurement reflects approximately the number of cued locations that can be selected to allow for near-ceiling search performance. The important result is not the capacity per se but whether capacity changes with display density. We also cannot know from the present data how participants store and use those cues (see General Discussion) or the strategies that participants use when there are too many cues to use (they might select only a subset or attempt to select each region but more coarsely, including some of the uncued locations).

We used this capacity measurement to estimate that participants in the dense condition could remember an average of 4.2 items, which is consistent with past results using a similar paradigm (Burkell & Pylyshyn, 1997). However, the present results provide an upper limit for this capacity by showing that performance quickly degrades with larger cued set sizes. Capacities in the dense condition were also variable among participants, with a low of 2.4 items and a high of at least 6 items. This range of individual variation is also reported in studies of rapid visual enumeration (Akin & Chase, 1978), multiple object tracking (Oksama & Hyona, 2004), and search through objects that suddenly appear in a display (Yantis & Jones, 1991).

Although both of these capacity estimates and the capacity estimates of the dense condition vary among participants, they still could be fixed for any individual participant. In contrast, capacity

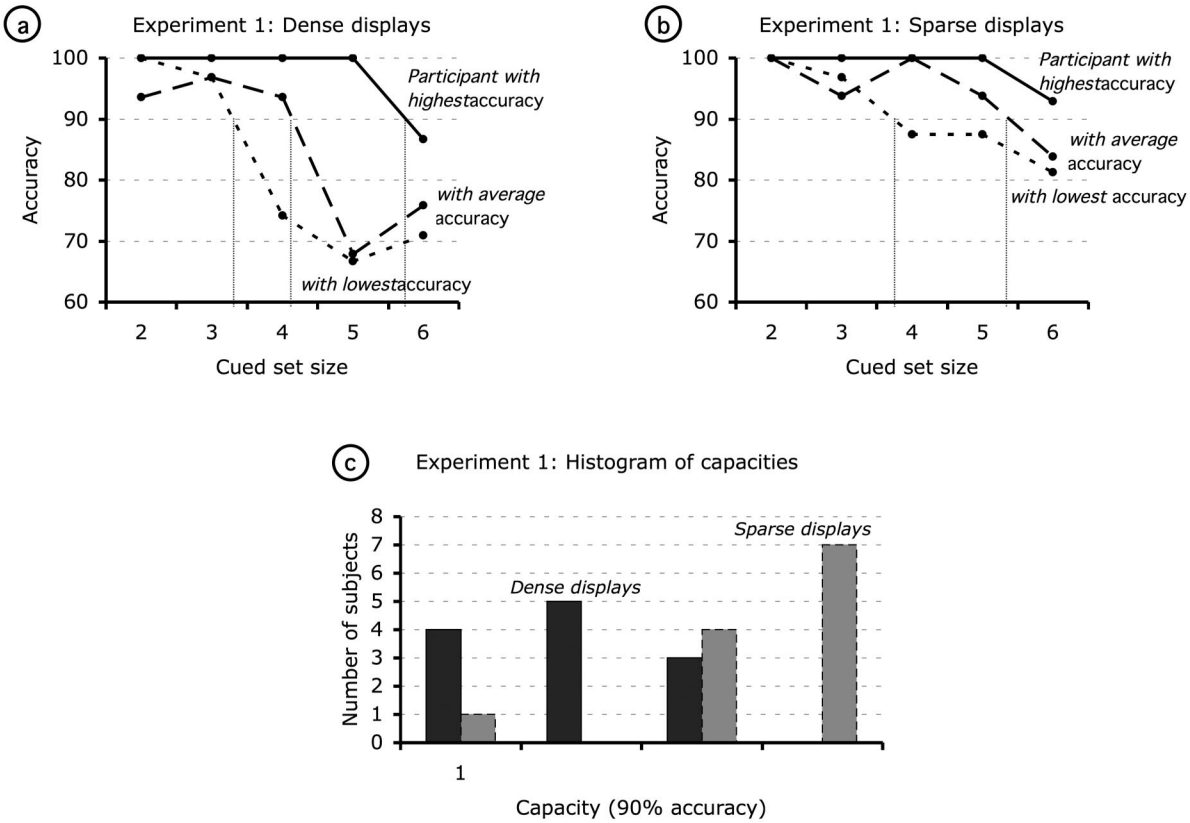

Figure 2. Average accuracy in Experiment 1 for each of 3 sample participants with highest, average, and lowest accuracy in dense displays (a) and in sparse displays (b); histogram of capacities in the sparse and dense conditions (c).

estimates in the sparse condition show that the number of locations that can be selected is not fixed for an individual participant. Average capacity in the sparse displays (5.6 items) was higher than average capacity in the dense displays (4.2 items), $t(11) = 5.7$, $p < .001$. Figure 2c shows a histogram of participant capacities in the two display types. For sparse displays, accuracy never dropped below 90% at any cued set size for 7 of the 12 participants, suggesting that most participants could search through at least 6 cued locations. With the exception of a single outlier participant with a capacity of 3.8 locations, the participant with the lowest capacity could still search through about 5.2 items on average. In contrast, for dense displays, no participants showed performance above 90% at the largest cued set size. The lowest capacity was 3.1 locations, and the highest was 5.8.

Although accuracy was stressed more strongly than speed, we also analyzed response times. Response times for sparse and dense displays in both the uncued and cued conditions are shown in Figure 3. Response times from both cue conditions were submitted to a 2 × 2 × 5 repeated measures analysis of variance, with cue type (search cued items, uncued items hidden), density (sparse, dense) and cued set size (2, 3, 4, 5, 6) as variables. All tests involving set size as a variable again showed heterogeneity of variance and were again Greenhouse–Geisser corrected. Responses were faster in uncued-hidden trials (674 ms) than in cued item search trials (959 ms), $F(1, 11) = 85$, $p < .001$, and response times increased with larger set sizes, $F(1.7, 19) = 44$, $p < .001$. Response times rose more in the cued condition than in the uncued condition, as shown by the interaction between cue condition and set size, $F(1.9, 21) = 27$, $p < .001$. Overall, response times on

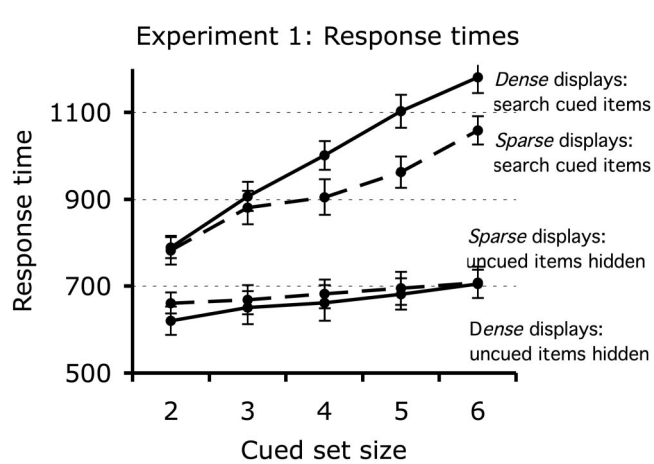

Figure 3. Average response times across all participants in the sparse and dense displays, in both the *search cued items* and *uncued items hidden* conditions. Error bars are standard errors.

sparse displays (792 ms) were not significantly faster than those on dense displays (809 ms), $F(1, 11) = 2.4$, $p = .15$, because of an interaction between density and cue condition $F(1, 11) = 10$, $p < .01$. Response times for sparse displays (920 ms) were faster than those for dense displays (998 ms) in the cued search item condition, $t(11) = 4.1$, $p = .002$, but there was an opposite, but nonsignificant, trend for the uncued hidden condition (sparse displays = 684 ms, dense displays = 664 ms, $t < 1$). There was also a nonsignificant trend for response times to increase more with set size for dense displays (56 ms/item) relative to sparse displays (35 ms/item), $F(2.5, 28) = 2.9$, $p = .06$. In the General Discussion, we explore some potential reasons for the higher response times in the cued condition.

In summary, participants could maintain selection for fewer items in the dense displays (an average of 4.2 items) relative to the sparse displays (5.6 items). This finding suggests that the number of locations that can be selected at once is not fixed but instead depends on the precision of the spatial selection that is required. However, the capacity estimate in the sparse displays is limited by the ceiling performance for most of the participants. Capacity would likely be even higher if cue set sizes over 6 had been tested. In Experiment 2, we tested 3 participants (using multiple sessions each) on a wider range of display densities and cued set sizes.

Experiment 2

In Experiment 2, we sought to generalize the relationship between the capacity of location selection and the spatial precision required to select those locations to a wider range of display densities and cued set sizes. We tested a small group of participants, who each participated in multiple testing sessions.

Method

S. L. Franconeri and G. A. Alvarez and one naive participant (J. C.) participated in the experiment. Stimuli were identical to those in Experiment 1, except that displays either contained 8, 16, or 20 total search items (see Figure 4), and 1–8 items were cued. The uncued condition was not tested.

Results and Discussion

Average accuracy for each participant is shown in Figure 5. As in Experiment 1, accuracy decreased as the number of relevant search locations increased. Also as in Experiment 1, accuracy began to drop sooner with dense than with sparse displays. We estimated the location capacity of participants S. L. Franconeri, G. A. Alvarez, and J. C. to be 6.3, 8.0, and 5.7, respectively, in the sparse displays; 3.8, 3.8, and 3.3 in the medium displays; and 3.0, 3.5, and 2.4 in the dense displays. When we varied display density more widely in this experiment than in previous ones, location capacity varied from a low of 2–3 items to a high of 6–7 items.

Experiment 3

In Experiments 1 and 2, participants' location capacity was lower in dense displays relative to sparse displays. We have attributed this lower capacity to a trade-off between the number of locations that can be selected and the precision of selection at each location. However, there are two potential alternative interpretations. One possibility is that chance performance is simply worse in denser displays, because denser displays contain more items. Perhaps a participant's number of selections is fixed, but when the number of cued locations exceeds this fixed limit, guessing strategies result in higher performance in sparse displays relative to dense displays. Another possibility is that participants show lower capacities in denser displays not because of the coarseness of their selection regions but because the cues in the initial cue display take longer to segment or interpret. That is, there could be a fixed limit to a participant's number of selected locations, but performance is worse on dense displays because participants need more time to interpret the initial display.

Experiment 3 rules out both of these possibilities by manipulating density without changing the number of items on the screen and by testing two durations for the initial cue display. If increasing item density still impairs accuracy when the number of items on the screen is the same, then guessing strategies cannot explain the results of Experiments 1 and 2. And if the duration of the cue display does not affect accuracy, then difficulties in interpreting the cue display cannot explain the lower accuracy in the dense conditions.

Method

Displays were similar to the displays in the sparse and medium conditions of Experiment 2, except that both always contained 16 total items. In sparse displays, these items were distributed over

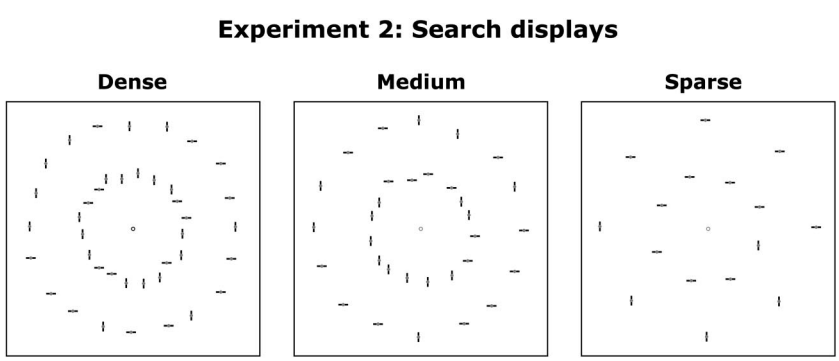

Figure 4. Sample search displays in Experiment 2.

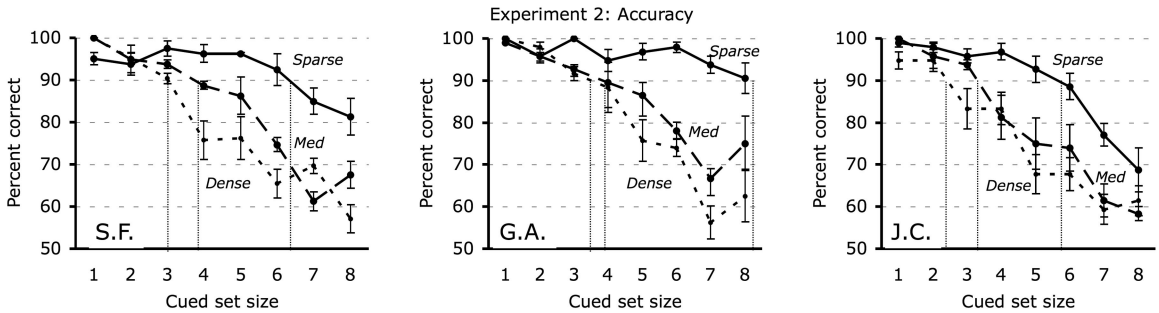

Figure 5. Accuracy data for all 3 participants in Experiment 2. Accuracy is shown separately for the sparse, medium (Med), and dense displays. S. F., G. A., and J. C. refer to participants. Error bars are standard errors.

the whole display, as in Experiment 2. But in dense displays, these 16 items were squeezed into the bottom half of the display, which is equivalent to a search display from the medium condition of Experiment 2 but with all of the items in the top half removed. Because some items fell directly on the horizontal midline (and were not in either the top or bottom of the screen) in Experiment 2, search displays in Experiment 3 were slightly rotated around fixation by half of the angular spacing between items, so that item positions were symmetrical on the left and right halves of the screen.

Initial cue displays were presented for either 500 or 1,200 ms, and this manipulation was crossed with the density manipulation. Each of these 4 block types contained 32 trials, and were presented 4 times, for a total of 512 trials. Block types were counterbalanced across participants, and each experiment began with the same color calibration procedure described in previous experiments, followed by 12 practice trials. Eight people participated in the experiment, including S. L. Franconeri, a research assistant familiar with the hypothesis, and 6 naive observers.

Results and Discussion

Two observers had accuracies below 60% and were removed from the analysis. Average accuracy for each participant for sparse and dense displays is shown in Figure 6. As in Experiments 1 and 2, accuracy was better in the sparse condition (M = 85.0%) relative to the dense condition (M = 68.8%), $t(5) = 5.2$, $p < .004$. Even though dense displays had the same number of items as sparse displays, participants still could select more items in sparse displays than in dense displays. Furthermore, performance was equal when cue displays were shown for a shorter 500-ms period (M = 76.9%) and for a longer 1,200-ms period (M = 76.9%). Because the duration of the cue display did not affect performance, lower accuracy in the dense conditions cannot be due to a lack of time for extracting the cued locations. There was also no hint of an interaction between the two factors.

General Discussion

In Experiment 1, when spatial locations in a search display were densely arranged, participants were able to search efficiently through a precued subset of up to about four spatial locations. However, when the spacing among those locations was sparsely arranged, participants were able to extend their search to more locations. Across Experiments 1 and 2, individual capacity varied widely according to the density of the search displays, from two to three locations up to six to seven locations. These results indicate that the number of locations that can be selected at once is not fixed but instead varies with the precision of the spatial selection that is required.

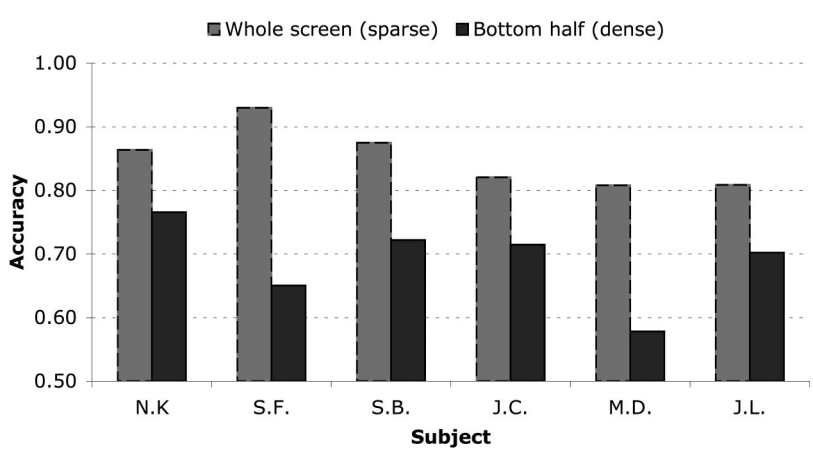

Figure 6. Accuracy data for all 6 participants in Experiment 3. Accuracy is shown separately for the sparse and dense displays.

Note that across our experiments, displays were always sparse enough to allow participants to select at least one location. Past research shows that when displays become too dense, elements in the display can be still discriminated from one another, but participants can no longer count their number or maintain selection on a single element. This effect of this interelement crowding seems to reflect a lower limit on the precision of selection when a single item is selected. When the distance among elements is denser than this level of precision, elements can no longer be selected independently (Intriligator & Cavanagh, 2001). Although this lower limit of resolution appears fixed, the present results show that the number of selected locations affects spatial precision well before this limit is reached.

How Do Individuals Select Multiple Locations?

Why does the number of object locations that can be selected vary with the precision of the spatial selection that is required? We believe the answer depends on how the participants selected the cued items. Participants were required to select specific locations, maintain this selection across a brief interval, and then perform a simple visual search using information from only these selected locations. But these results in themselves do not specify the nature of the process or representation that degrades as more locations are selected. What is meant by *selection* in the first place? There are at least two possibilities (see Figure 7 for illustration), which differ on whether the mechanism used to store the locations of the cued items is the same as the mechanism used to encode information from the cued locations.

According to the *single mechanism* account, the mechanism that stores the cued locations is the same one that encodes information from those locations. That is, additional locations cannot be stored without a relative decrease in the amount of visual information encoded from those locations. This is the strongest sense in which participants could select multiple locations at once. Previous results support the idea that memory for spatial locations and selection of spatial locations each rely on a shared pool of cognitive resources. Several studies have shown that maintaining a location in memory speeds processing at that location, whereas memory for a spatial location is impaired when a participant is forced to encode information from another location (Awh, Jonides, & Reuter-Lorenz, 1998; Smyth, 1996; Smyth & Scholey, 1994). In addition, performance in a visual search task (which requires encoding information from locations in space) was impaired when participants were required to concurrently remember a set of spatial locations (Oh & Kim, 2004; Woodman & Luck, 2004).

This single mechanism account thus attributes the trade-off between the precision of selection and the number of selected locations to the overloading of a central selection mechanism. This possibility is supported by results suggesting that it is easiest to attend to more than one target at a time when multiple targets are relatively far away from one another (Bahcall & Kowler, 1999; Cutzu & Tsotsos, 2003; Fecteau & Enns, 2005; Kristjansson & Nakayama, 2002). The most complete computational account of these findings to date is that selective attention to an item is implemented by the creation of an inhibitory spatial surround around the attended item (Cutzu & Tsotsos, 2003; Tsotsos, Culhane, Wai, Davis, & Nuflo, 1995). Such an inhibitory surround for each attended location would result in mutual interference when two nearby locations are both the focus of attention.

This possibility is also consistent with the results from tasks showing a direct link between visual attention and spatial precision. For example, focused attention on a spatial location increases the spatial resolution of object representations in that location (Yeshurun & Carrasco, 1998), implying decreased spatial resolution for items of divided attention. Participants also localize the position of a briefly presented stimulus at a cued location more precisely than at an uncued location (Tsal & Bareket, 1999). Finally, in a study similar to the present one, participants were required to perform a difficult visual search task but were cued to the potential target locations either within or outside of an attentional blink triggered by a preceding rapid serial visual presentation task. Participants could store and use fewer cued locations when the cues were presented during the blink. Similarly, when asked to point to the location of a single cue, participants were less accurate when that cue had been presented during an attentional blink (Olivers, 2004).

According to a second *dual-mechanism* account of selection, participants remember the cued positions using a different offline form of storage, while only encoding information from a single cued position or a small subset of positions at once. This memory could be based on the shape created by the constellation of cued locations and might be similar to the representations that allow people to recognize and remember the shapes of objects. There is some evidence that this type of representation aids performance in multiple object tracking tasks, in which participants might store the locations of the target items as the vertices of a complex polygon

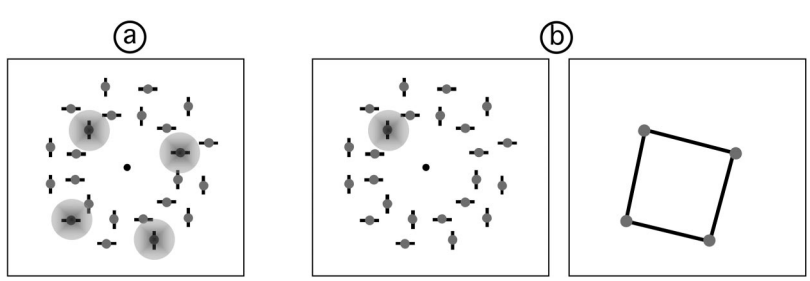

Figure 7. According to the *single mechanism* account, the mechanism that stores the cued locations is the same one that encodes information from those locations (a); according to the *dual-mechanism* account, participants remember the cued positions using a different offline form of storage, while only encoding information from a single cued position or a small subset of positions at once (b).

(Yantis, 1992; but see also Pylyshyn, 2004). It is also possible that this offline form of storage is not based on shape memory. According to the finger of instantiation model (Pylyshyn, 1989; Pylyshyn & Storm, 1988), a small number of objects (e.g., four) are indexed, which allows quick access to their properties and easy computation of spatial properties such as spatial relations or colinearity between objects. These indexes could allow participants to remember a set of cued locations and search them even after the cues have disappeared (Burkell & Pylyshyn, 1997). Note that under both proposals, the memory for the cued locations still represents a form of selection, because some locations are still privileged over others. But the way that the locations are selected and the temporal pattern of the subsequent encoding of information are quite different.

From a dual-mechanism perspective, the drop in precision with more cued locations must be related to this offline storage mechanism. If the cued items were stored as a shape constellation, it would be reasonable to expect the resolution of this constellation to decrease as the shape becomes more complex, based on evidence from visual memory tasks (Alvarez & Cavanagh, 2004; Phillips, 1974). The finger of instantiation account, on the other hand, cannot directly explain the present results, because the number of indexes is fixed and should not vary with the precision of selection required. To explain these data, the indexes would have to be supplemented by a second memory system, such as a shape memory system, whose capacity is affected by the precision of selection required by the task.

At first glance, the dual-mechanism account seems supported by the high search slopes found in the cued conditions of Experiment 1 (high search slopes were also found by Burkell & Pylyshyn, 1997). In contrast to the uncued conditions, in which search slopes were low (but not flat), search slopes were high in the cued conditions (M = 65 ms/item for the sparse condition, M = 99 ms/item for the dense condition). When we included only set sizes within each participant's capacity limit, slopes decreased only slightly in the sparse condition (M = 58 ms/item), $t(11) = 1.96$, $p = .08$, and stayed the same in the dense condition (M = 110 ms/item), $t(11) < 1$. This result could be taken to suggest that participants were forced to search only one item at a time in the cued conditions, when participants were required to search multiple noncontiguous locations. However, these high search slopes can also be explained by the single-mechanism account. Participants may have selected and encoded information from all objects at the same time, but when more objects were selected, the selection regions became more coarse, allowing more information from uncued objects (noise) to interfere with the information being encoded from the cued location (signal). Thus, both accounts can explain the finding of high search slopes for cued locations equally well.

Future research should explore which of these two accounts best describes the trade-off between the number of selections and the precision of those selections. Is the reduction in spatial precision due to increasingly coarse areas of active encoding of information? Or is it due to the inherent coarseness of the offline spatial memory for complex shapes? For example, if selecting multiple locations requires a memory for the shape of the constellation of cued locations, then a secondary task requiring memorization of a complex shape should interfere with a participant's ability to select multiple locations. One could also examine whether information is concurrently encoded from the memorized locations. If it is then making the task more difficult at each location (e.g., by increasing search difficulty) should require that the selection mechanism focus on fewer locations at a time, thus reducing a participant's estimated capacity. But if participants encode information from only one location at a time and rely on an offline shape memory system to remember the other locations, then changing the difficulty of the task should not affect a participant's capacity (see Yantis & Jones, 1991, for a similar manipulation).

Conclusions

Is the visual system designed to deal with a fixed number of objects or locations at once? Our results suggest not. Participants in Experiments 1–3 had selection limits that varied from two to eight items, suggesting that the number of locations that can be simultaneously selected is flexible. One might still argue that this number is fixed and that the flexibility shown in the present experiments is due to the flexibility of an auxiliary memory system. But this account lacks any direct support at present, leaving it only as a theoretically interesting way to salvage the concept of a fixed number of selection mechanisms for this task. In contrast, if one takes into account that performance limits in other tasks such as rapid enumeration, visual search, and multiple object tracking (and now location selection) can all be explained with flexible mechanisms that trade capacity for precision, there is little reason left to posit fixed capacity mechanism in the first place.

In conclusion, the present results show that one's capacity to select multiple locations is not fixed but is instead limited by the spatial precision required to uniquely specify each location. These results are consistent with other evidence showing that the number of objects that can be dealt with at once is not fixed but instead is limited by the spatial precision required to represent each item, in tasks such as multiple object tracking (Alvarez & Franconeri, 2007b), remembering the locations of newly introduced objects (Yantis & Jones, 1991), rapid visual enumeration (Whalen et al., 1999), and visual short-term memory (Alvarez & Franconeri, 2007a).

References

Akin, O., & Chase, W. (1978). Quantification of three-dimensional structures. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 397–410.

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by total information load and by number of objects. *Psychological Science, 15,* 106–111.

Alvarez, G. A., & Franconeri, S. L. (2007a). *The allocation of visual short-term memory capacity: Evidence for a flexible storage mechanism.* Manuscript submitted for publication.

Alvarez, G. A., & Franconeri, S. L. (2007b). *How many objects can you attentively track? Evidence for a resource-limited tracking mechanism.* Manuscript submitted for publication.

Awh, E., Jonides, J., & Reuter-Lorenz, P. A. (1998). Rehearsal in spatial working memory. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 780–790.

Awh, E., & Pashler, H. (2000). Evidence for split attentional foci. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 834–846.

Bahcall, D. O., & Kowler, E. (1999). Attentional interference at small spatial separations. *Vision Research, 39,* 71–86.

Belopolsky, A. V., Theeuwes, J., & Kramer, A. F. (2005). Prioritization by transients in visual search. *Psychonomic Bulletin & Review, 12,* 93–99.

Burkell, J. A., & Pylyshyn, Z. W. (1997). Searching through subsets: A test of the visual indexing hypothesis. *Spatial Vision, 11,* 225–258.

Cowan, N. (2001). The magical number four in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87–185.

Cutzu, F., & Tsotsos, J. K. (2003). The selective tuning model of visual attention: Testing the predictions arising from the inhibitory surround mechanism. *Vision Research, 43,* 205–219.

Donk, M., & Theeuwes, J. (2001). Visual marking beside the mark: Prioritizing selection by abrupt onsets. *Perception & Psychophysics, 63,* 891–900.

Egeth, H. E., Virzi, R. A., & Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 32–39.

Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 583–597.

Fecteau, J. H., & Enns, J. T. (2005). Visual letter matching: Hemispheric functioning or scanning biases? *Neuropsychologia, 43,* 1412–1428.

Friedman-Hill, S., & Wolfe, J. M. (1995). Second-order parallel processing: Visual search for the odd item in a subset. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 531–551.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44,* 43–74.

Hahn, S., & Kramer, A. E. (1998). Further evidence for the division of attention between noncontiguous locations. *Visual Cognition, 5,* 217–256.

Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology, 43,* 171–216.

Kramer, A. E., & Hahn, S. (1995). Splitting the beam: Distribution of attention over noncontiguous regions of the visual field. *Psychological Science, 6,* 381–386.

Kristjansson, A., & Nakayama, K. (2002). The attentional blink in space and time. *Vision Research, 42,* 2039–2050.

McCormick, P. A., & Klein, R. (1990). The spatial distribution of attention during covert visual orienting. *Acta Psychologica, 75,* 225–242.

Nakayama, K., He, Z. J., & Shimojo, S. (1995). Visual surface representation: A critical link between lower-level and higher-level vision. In S. M. Kosslyn & D. N. Osherson (Eds.), *Vision: An invitation to cognitive science* (pp. 1–70). Cambridge, MA: MIT Press.

Oh, S., & Kim, M. (2004). The role of spatial working memory in visual search efficiency. *Psychonomic Bulletin & Review, 11,* 275–281.

Oksama, L., & Hyona, J. (2004). Is multiple object tracking carried out automatically by an early vision mechanism independent of higher-order cognition? An individual difference approach. *Visual Cognition, 11,* 631–671.

Olivers, C. (2004). Blink and shrink: The effect of the attentional blink on spatial processing. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 613–631.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics, 16,* 283–290.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3–25.

Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition, 32,* 65–97.

Pylyshyn, Z. W. (2004). Some puzzling findings in multiple object tracking: I. Tracking without keeping track of object identities. *Visual Cognition, 11,* 801–822.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3,* 1–19.

Rensink, R. (2000). Visual search for change. *Visual Cognition, 7*(1–3), 345–376.

Saenz, M., Buracas, G. T., & Boynton G. M. (2003). Global feature-based attention for motion and color. *Vision Research, 43,* 629–637.

Smyth, M. M. (1996). Interference with rehearsal in spatial working memory in the absence of eye movements. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 49*(A), 940–949.

Smyth, M. M., & Scholey, K. A. (1994). Interference in immediate spatial memory. *Memory & Cognition, 22,* 1–13.

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,* 331–335.

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,* 1–23.

Tsal, Y., & Bareket, T. (1999). Effects of attention on localization of stimuli in the visual field. *Psychological Bulletin & Review, 6,* 292–296.

Tsotsos, J. K., Culhane, S. W., Wai, Y. L., Davis, N., & Nuflo, F. (1995). Modeling visual attention via selective tuning. *Artificial Intelligence, 78,* 507–547.

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science, 10,* 130–137.

Woodman, G. F., & Luck, S. J. (2004). Visual search is slowed when visuospatial working memory is occupied. *Psychonomic Bulletin & Review, 11,* 269–274.

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology, 24,* 295–340.

Yantis, S., & Johnson, D. N. (1990). Mechanisms of attentional priority. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 812–825.

Yantis, S., & Jones, E. (1991). Mechanisms of attentional selection: Temporally modulated priority tags. *Perception & Psychophysics, 50,* 166–178.

Yeshurun, Y., & Carrasco, M. (1998, November 5). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature, 396,* 72–75.

Received October 22, 2005
Revision received December 22, 2006
Accepted December 29, 2006 ■