

The role of visual short-term memory in empty cell localization

ANDREW HOLLINGWORTH, JOO-SEOK HYUN, and WEIWEI ZHANG
University of Iowa, Iowa City, Iowa

Three experiments examined the visual memory representation supporting performance at long interstimulus intervals (ISIs) in an empty cell localization task. Two arrays of dots within a 4×4 grid were displayed briefly in succession. One grid cell did not contain a dot in either array, and the task was to localize the empty cell. In Experiment 1, we replicated previous findings of recovery to high levels of performance at long ISIs. In Experiment 2, we tested whether figural grouping in visual short-term memory (VSTM) supports long-ISI performance by manipulating the complexity of the array pattern. Pattern complexity had no effect on empty cell localization at 0-msec ISI, suggesting dependence on high-capacity visible persistence, but there was a large simple pattern advantage at long ISIs, suggesting dependence on figural grouping in VSTM. Experiment 3 demonstrated that participants typically remember the empty cells of the first array, and not the dots, for comparison with Array 2.

Current evidence suggests there are four memory systems for the retention of visual information: visible persistence, informational persistence, visual short-term memory (VSTM), and visual long-term memory (VLTM) (for reviews, see Hollingworth, in press; Irwin, 1992b; Palmer, 1999). Visible persistence and informational persistence preserve a precise, high-capacity, point-by-point, retinotopic sensory trace that decays very quickly after a stimulus event (Averbach & Coriell, 1961; Coltheart, 1980; Di Lollo, 1980; Irwin & Yeomans, 1986; Sperling, 1960). Together, visible persistence and informational persistence are often termed *iconic memory* or, preferably, *sensory persistence*. Visible persistence is a visible trace that decays within 80–100 msec after stimulus onset (Di Lollo, 1980). If the stimulus onset asynchrony (SOA) between two stimuli is less than 80–100 msec, the visible persistence of the first overlaps with the initial sensory processing of the second, causing the two stimuli to perceptually fuse into a composite image (Di Lollo, 1980; Eriksen & Collins, 1967). When the two stimuli are complementary, such as two sets of dots forming a letter string (Eriksen & Collins, 1967), participants are able to perceive the full, composite stimulus at short SOAs. When two stimuli are not complementary, sensory integration leads to backward masking at short SOAs: The contours of the trailing mask perceptually fuse with the leading target stimulus, making the target more difficult to perceive.

Informational persistence also maintains a high-capacity sensory trace but in a format that does not support visual awareness (i.e., unlike visible persistence, informational

persistence is not visible). Whereas the duration of visible persistence is locked to the onset of a stimulus, informational persistence lasts approximately 150–300 msec from stimulus offset (Irwin & Yeomans, 1986; Phillips, 1974; Yeomans & Irwin, 1985). Informational persistence does not integrate with a trailing stimulus, but it is nonetheless highly susceptible to backward masking (Irwin & Yeomans, 1986; Pashler, 1988; Phillips, 1974). It is likely that early studies examining high-capacity sensory memory (e.g., Averbach & Coriell, 1961; Sperling, 1960) primarily tapped informational persistence, since the advantage for partial report over full report extended well beyond the temporal range of visible persistence.

VSTM maintains visual representations abstracted away from precise sensory information. Unlike high-capacity forms of sensory persistence, VSTM has a limited capacity of three or four objects (Luck & Vogel, 1997; Pashler, 1988) and less spatial precision than point-by-point sensory persistence (Irwin, 1991; Phillips, 1974). However, VSTM is considerably more robust than sensory persistence. It is not significantly disrupted by backward masking (Pashler, 1988; Phillips, 1974) and can be maintained for longer durations—on the order of seconds (Phillips, 1974)—and across saccades (Henderson, 1997; Irwin, 1991, 1992a). Evidence for object-based representations in VSTM comes from the fact that VSTM capacity is driven primarily by the number of objects to remember and not by the number of visual features to remember (Irwin & Andrews, 1996; Luck & Vogel, 1997). In addition, VSTM is sensitive to higher order object structure (Howe & Jung, 1986; Pomerantz, 1977; Sebrechts & Garner, 1981). Using a sequential same-different task, Sebrechts and Garner found that patterns of dots were remembered more efficiently when the dots could be grouped into a simple global figure than when they could not. We use the term *higher level visual representation* to describe the type of

This research was supported by NIMH Grant R03 MH65456. Correspondence should be addressed to A. Hollingworth, Department of Psychology, University of Iowa, Iowa City, IA 52240 (e-mail: andrew-hollingworth@uiowa.edu).

abstracted, object-based visual information retained in VSTM. VSTM maintains visual representations of format similar to those maintained in VSTM but has the capability to accumulate hundreds of visual object representations (Hollingworth, 2004) over longer time scales and across multiple discrete inputs (Hollingworth, 2005; Standing, Conezio, & Haber, 1970).

As is evident from the above discussion, perhaps the most fundamental distinction between visual memory systems is between low-level, high-capacity storage (visible and informational persistence) and higher level, limited-capacity storage (VSTM). This distinction has been investigated most directly using two related paradigms: a change detection paradigm developed by Phillips (1974) and an empty cell localization paradigm developed by Di Lollo (1977; Hogben & Di Lollo, 1974). The Phillips paradigm has been used primarily to distinguish informational persistence from VSTM. Phillips tested memory for checkerboard objects consisting of a virtual grid with approximately half of the squares randomly filled. The initial stimulus (Array 1) was presented for 1,000 msec, followed by a variable interstimulus interval (ISI) and a test stimulus (Array 2). Array 2 was either identical to Array 1 or identical except for the addition or deletion of a single filled cell. The task was change detection. The Array 1 duration of 1,000 msec placed Array 2 well outside the SOA range of visible persistence (Di Lollo, 1980).¹ Phillips observed very high levels of performance at short ISIs and significantly lower levels of performance at longer ISIs. In addition, short- and long-ISI performance diverged on a number of variables. Short-ISI performance was strongly disrupted by an unpredictable shift in Array 2 location, whereas long-ISI performance was practically unaffected. In addition, long-ISI performance was significantly influenced by the size of the array (in terms of number of cells), whereas short-ISI performance was not. Phillips argued that short-ISI performance was supported by a high-capacity sensory memory that was maintained in low-level retinotopic coordinates (informational persistence) and that long-ISI performance was supported by a limited-capacity memory abstracted away from precise spatial organization (VSTM).

The second major paradigm, empty cell localization, has been used primarily to distinguish visible persistence from other forms of visual memory (Di Lollo, 1977, 1980; Di Lollo & Hogben, 1987; Dixon & Di Lollo, 1994; Enns & Visser, 2001; Hogben & Di Lollo, 1974). In this paradigm, two arrays of dots are displayed in sequence within a visible or virtual grid. The first array fills approximately half of the cells in the grid. Array 2 fills all but one of the cells not filled in Array 1. Between the two arrays, one cell is always empty, and the task is to specify the location of this "missing dot." Array 1 is typically presented for a very short duration (~30 msec), followed by a variable ISI and Array 2 (also presented very briefly). At short ISIs (and thus short SOAs, given the 30-msec Array 1 duration), the two arrays perceptually integrate, participants perceive a single dot array with one empty cell, and the task is very easy indeed. As ISI is increased, performance

drops very rapidly to near floor levels by 100-msec ISI.² The rapid drop from 0- to 100-msec ISI in the standard paradigm likely derives from multiple sources. First, it clearly reflects the decay of visible persistence. Informational persistence should still be available at 100-msec ISI, but informational persistence does not support phenomenological integration. The rapid reduction in performance as ISI increases to 100 msec is also likely to be caused by interruption masking, in which the consolidation of a stimulus into a stable, reportable form is interrupted by the processing demands of a trailing stimulus (see Enns & Di Lollo, 2000, for a review). Such masking is most evident at positive ISIs in the range of 50–100 msec.

As a whole, research on early visual memory tells a consistent story. After a stimulus event, high-capacity, low-level sensory persistence is succeeded by higher level VSTM representations, with the capacity of VSTM limited to three or four objects. However, Brockmole, Wang, and Irwin (2002) produced evidence potentially difficult to reconcile with this view. Brockmole et al. used a version of the empty cell localization paradigm, testing longer ISIs than had been typically used in that paradigm. Array 1 was a 4×4 or 5×5 grid with either 7 or 12 dots, randomly placed. Array 1 was presented for 33 msec, followed by Array 2 after the ISI. Array 2 consisted of 8 or 12 dots, with one cell empty across the two arrays. Brockmole et al. found the standard high levels of empty cell localization at 0-msec ISI and rapid decline in performance by 100-msec ISI. Surprisingly, however, performance increased with even longer ISIs. By 1,500-msec ISI, localization performance had returned to levels close to those observed at 0-msec ISI. Brockmole et al. estimated the number of dots retained at long ISIs, concluding that for the 4×4 grid, participants had retained on average 6.2 of the 7 dots from Array 1 and 7.8 of the 8 dots from Array 2. Capacity estimates were even larger for the 5×5 grid, with an estimate of 9.6 of 12 dots retained from Array 1 and 11.7 of 12 dots from Array 2.

This accurate localization performance was observed at ISIs when memory for Array 1 should have been supported only by VSTM, yet the Array 1 capacity estimates quite significantly exceeded three or four individual dots. What accounts for this apparently high-capacity memory, and can these data be reconciled with current conceptions of limited-capacity VSTM? One possibility is that apparently high-capacity performance at long ISIs is supported by high-capacity sensory memory. If so, then the visual memory system framework reviewed above would need to be revised, allowing for much longer sensory retention than has previously been found. A second possibility is that long-ISI performance is supported by VSTM, and VSTM can maintain more than three or four individual objects, at least for this type of stimulus. If so, then the three- or four-object limitation on VSTM capacity would need to be reconsidered. A third possibility is that long-ISI performance is supported by VSTM, and VSTM groups individual dots into a larger scale object or objects. Under this *figural grouping hypothesis*, multiple dots are grouped into one or more higher order figures. VSTM

would still be limited to three or four objects, but those object representations would each contain information from multiple dots. Because the final hypothesis is most consistent with what is currently known about sensory persistence and VSTM, we considered it the most likely and sought to test it.

In Experiment 1, we replicated the Brockmole et al. (2002) paradigm, observing a similar recovery to relatively high levels of performance at longer ISIs. In Experiment 2, we tested the figural grouping hypothesis by manipulating the ease of figure extraction from Array 1, a manipulation of pattern complexity. Performance at 0-msec ISI was unaffected by pattern complexity, whereas performance at longer ISIs was quite dramatically higher for simple Array 1 patterns than for complex Array 1 patterns. These data demonstrate a VSTM locus for long-ISI performance and reconcile the apparently high-capacity performance observed by Brockmole et al. with the standard limited-capacity model of VSTM. Finally, in Experiment 3, we used a pattern complexity manipulation to test two competing models of how VSTM supports accurate localization performance at long ISIs: an integration hypothesis (Brockmole et al., 2002) and a negative-space comparison hypothesis (Jiang, Kumar, & Vickery, 2005). The results supported the latter model, indicating that performance at long ISIs is supported by VSTM for the Array 1 negative space (empty cells), which is then compared with Array 2 to isolate the empty cell.

EXPERIMENT 1

In Experiment 1, we replicated the basic method of Brockmole et al. (2002). An initial, random array of dots was presented within a 4×4 grid for 30 msec, followed by a variable ISI between 0 and 4,000 msec, followed by a second array of dots for 30 msec. One cell in the grid

did not contain a dot in either of the two arrays. After the appearance of the second array, the participants moved a mouse cursor to the cell that had not contained a dot and clicked to register their response.

Two dot array conditions were examined. In the 7–8 condition, Array 1 contained seven dots, and Array 2 contained eight dots (leaving one of the 16 cells unfilled). This condition was a direct replication of Brockmole et al. (2002, Experiment 1). In the 8–7 condition, Array 1 contained eight dots, and Array 2 contained seven dots. This condition was included to allow comparison with Experiment 2, which manipulated Array 1 complexity and used the 8–7 design given the availability of complexity ratings for eight-dot configurations. The two conditions were run as separate experiments with two sets of participants.

Method

Participants. Sixteen participants from the University of Iowa community completed the experiment, 8 in each experiment. All participants reported 20/20 uncorrected or corrected vision.

Stimuli. The stimuli were similar to those used in Brockmole et al. (2002, Experiment 1). Two arrays of dots were presented on each trial within a 4×4 grid. In the 7–8 condition, Array 1 consisted of seven dots, assigned to 7 of the 16 cells in the grid. Array 2 consisted of eight dots, assigned to eight of the nine cells not occupied in Array 1. In the 8–7 condition, the number of dots assigned to each array was reversed. In both conditions, one cell was not filled in either array. The positions of the dots were determined randomly. As a result, the position of the empty cell was also randomly determined. The stimuli were displayed against a light-gray background. The 4×4 grid was composed of light-blue lines superimposed over the background. Dots were presented in black. The entire grid subtended 13.7° of visual angle (both horizontally and vertically). Each cell in the grid subtended 3.4° . The diameter of each dot was 2.7° .

Apparatus. The stimuli were displayed at a resolution of 800 by 600 pixels on a 17-in. video monitor at a refresh rate of 100 Hz. The initiation of image presentation was synchronized to the monitor's vertical refresh. Responses were collected using a mouse. The presentation of stimuli and collection of responses were controlled by

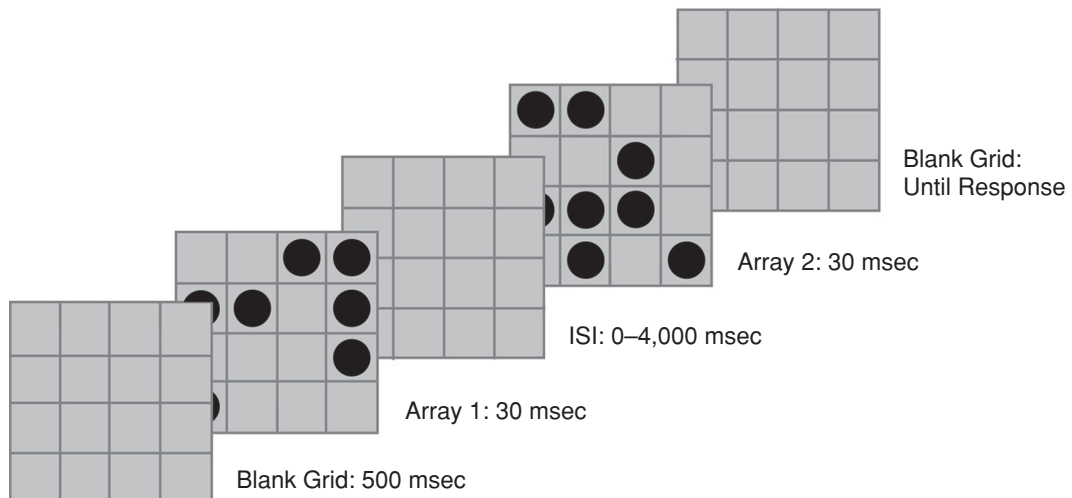


Figure 1. Sequence of events in a trial of Experiment 1. The participants pressed a pacing button to begin the trial, followed by the events illustrated in the figure. When the final blank grid appeared, the participants used a mouse cursor to click on the cell that had not contained a dot.

E-Prime software running on a Pentium IV-based computer. Viewing distance was maintained at 80 cm by a forehead rest. The room was dimly illuminated by a low-intensity light source.

Procedure. Each trial consisted of the following events (see Figure 1). First, the empty grid (superimposed over the gray background) was displayed. When ready to begin, the participants pressed a mouse button to start the trial. There was a 500-msec delay before presentation of Array 1. Array 1 was then presented within the grid for 30 msec (three refresh cycles at 100 Hz). The ISI between the end of Array 1 presentation and the beginning of Array 2 presentation was 0, 20, 50, 100, 250, 500, 1,000, 1,500, 2,000, 2,500, 3,000, or 4,000 msec. During the ISI, the blank grid was displayed. In the 0-msec ISI condition, Array 1 was presented for three refresh cycles (as in all conditions), and Array 2 was written to the screen at the onset of the fourth refresh cycle. Following the ISI, Array 2 was presented within the grid for 30 msec. Array 2 was followed by the empty grid with a mouse cursor presented in the center. The participants moved the mouse cursor to the cell they thought had been empty and clicked to respond. A mouse click outside of the grid or on a grid line was not accepted, and the cursor was returned to the center of the screen so the participant could provide a valid response. A mouse click within one of the cells terminated the trial. There was a 1,000-msec blank (gray screen) delay between the end of a trial and the beginning of the next trial.

The participants were given the following instructions, similar to those used in Brockmole et al. (2002):

When the blue grid appears, press the space bar to begin the trial. Two sets of dots will be displayed sequentially within the grid. The delay between the two sets will vary from almost no delay to a delay of about 4 seconds. There will always be one grid cell that did not contain a dot in either of the two sets. Your task is to decide the location of this "missing dot." After the two sets of dots have been displayed, a mouse cursor will appear on the screen. Move the cursor to the cell corresponding to the missing dot. Then click. Please strive to do this as accurately as possible. A good strategy may be to imagine that the first set of dots is still present after they have disappeared.

The participants first completed a practice session of 24 trials, two in each of the 12 ISI conditions, randomly intermixed. Feedback was provided in the practice session; the 1,000-msec gray screen between trials contained either the word *correct* or the word *incorrect*. The practice session was followed by the experimental session, consisting of four blocks of trials. Each block contained 96 trials, 8

in each of the 12 ISI conditions. Thus, the participants completed a total of 384 trials, 32 in each of the ISI conditions. Within each block, trial order was determined randomly, producing a random intermixing of trials from the different ISI conditions. Feedback was not provided in the experimental session. The entire experiment lasted approximately 45 min.

Results and Discussion

7–8 condition. The results of the 7–8 condition are displayed in Figure 2. In this and in all subsequent analyses, participant was treated as a random effect. Percentage correct localization data were examined in a one-way repeated measures ANOVA, with ISI as the independent variable. Localization performance varied significantly with ISI [$F(11,77) = 22.64, p < .001$]. The pattern of performance replicated the main characteristics of the Brockmole et al. (2002, Experiment 1) data. Performance at 0-msec ISI was quite accurate (79.3%), characteristic of sensory integration. There was an initial rapid decline with increasing ISI, which was most pronounced at 100-msec ISI (21.9%). This initial drop in localization performance was followed by a subsequent increase in performance as ISI was increased further, the critical result obtained by Brockmole et al. Maximum recovery was observed at 3,000-msec ISI (58.6%). Consistent with this clearly nonmonotonic pattern of performance, there was a reliable cubic trend in the percentage correct data [$F(1,7) = 71.74, p < .001$]. Examining short-ISI performance, localization was reliably more accurate at 0-msec ISI than at 100-msec ISI [$F(1,7) = 175.28, p < .001$]. Examining the subsequent increase in performance, localization approached asymptote by 1,000-msec ISI, and performance was reliably higher at 1,000-msec ISI than at 100-msec ISI [$F(1,7) = 12.26, p < .01$]. Performance at longer ISIs did not completely rebound to the 0-msec ISI level. Performance at 0-msec ISI was reliably higher

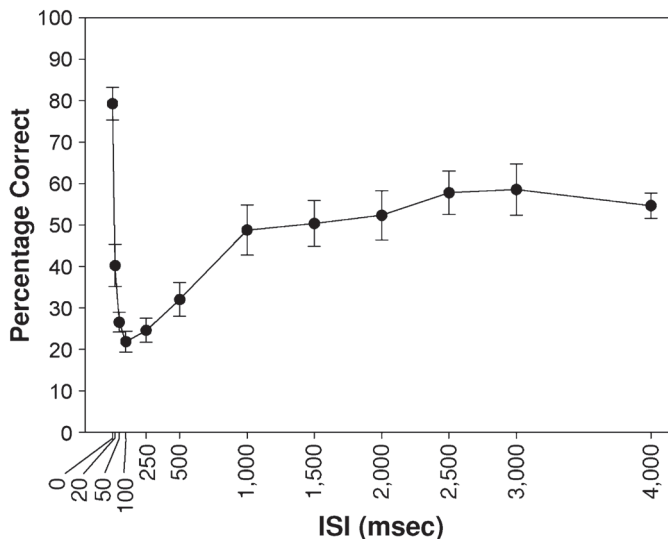


Figure 2. Localization accuracy as a function of ISI for the 7–8 condition in Experiment 1. Error bars represent standard errors of the means.

than performance at 3,000-msec ISI [$F(1,7) = 10.25, p < .05$].

Absolute levels of localization accuracy were very similar to those in Brockmole et al. (2002, Experiment 1), who observed 79% correct at 0-msec ISI and 21% correct at 100-msec ISI (compared with 79.3% and 21.9% in the present experiment). The only major difference between the two experiments was that long-ISI performance in the present experiment did not quite recover to the same levels as that in Brockmole et al. Long-ISI performance reached asymptote at approximately 68% correct in Brockmole et al., compared with approximately 54% correct in the present experiment (the average of performance at 1,000 to 4,000-msec ISI). This difference may derive from the one significant methodological difference between the two experiments. Trials in Brockmole et al. were blocked by ISI condition, whereas trials from different ISI conditions were randomly intermixed in the present experiment. If long-ISI performance is supported by strategic encoding of Array 1, such as figural grouping, then such strategic processing may have been facilitated by ISI blocking in Brockmole et al., because the strategy could be consistently applied across all trials in a block.

8–7 condition. The results of the 8–7 condition are displayed in Figure 3. One participant was replaced in this condition because of very low levels of performance (at 0-msec ISI, which typically yields accurate localization, the replaced participant's mean was more than 3 standard deviations lower than that of the remaining 7 participants).

The 8–7 condition data produced a pattern very similar to that in the 7–8 condition. Localization performance varied significantly with ISI [$F(11,77) = 23.03, p < .001$]. As in the 7–8 condition, performance at 0-msec ISI was quite accurate (77.0%). There was an initial rapid decline with increasing ISI, which was most pronounced at 50-msec ISI (20.7%). This initial drop in localization per-

formance was followed by a subsequent increase in performance as ISI was increased further. Maximum recovery was observed at 3,000-msec ISI (63.3%). Consistent with this clearly nonmonotonic pattern of performance, there was a reliable cubic trend in the percentage correct data [$F(1,7) = 313.57, p < .001$]. Examining short-ISI performance, localization was reliably more accurate at 0-msec ISI than at 50-msec ISI [$F(1,7) = 477.47, p < .001$]. Examining the subsequent increase in performance, localization approached asymptote by 1,500-msec ISI, and performance was reliably higher at 1,500-msec ISI than at 50-msec ISI [$F(1,7) = 79.50, p < .001$]. Performance at longer ISIs did not completely rebound to the 0-msec ISI level. Performance at 0-msec ISI was marginally higher than performance at 3,000-msec ISI [$F(1,7) = 5.53, p = .05$].

EXPERIMENT 2

Experiment 1 replicated the major effects of Brockmole et al. (2002, Experiment 1) and, in particular, replicated the recovery to relatively high levels of performance at long ISIs. This result sets the stage for the central question in the present study: What accounts for the apparently high-capacity performance at long ISIs? To test the hypothesis that such performance is supported by figural grouping in VSTM, we used a method developed by Irwin (1991, Experiment 6). Irwin used a modified Phillips paradigm along with a manipulation of pattern complexity. The basic paradigm displayed a dot array (Array 1) consisting of eight dots in 8 of 16 possible locations in an invisible 4×4 grid. Array 1 was followed by a second array after a variable ISI. Array 2 was either identical to Array 1 or identical except for the position of a single dot. The task was change detection. To examine the nature of Array 1 coding, Irwin manipulated the pattern complexity

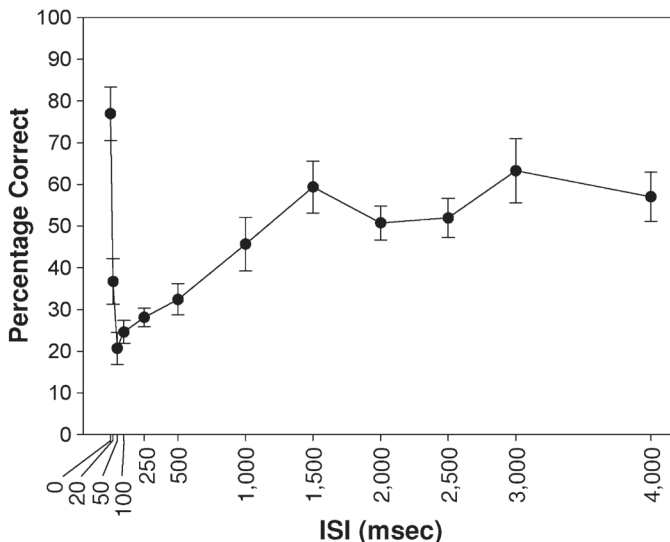


Figure 3. Localization accuracy as a function of ISI for the 8–7 condition in Experiment 1. Error bars represent standard errors of the means.

of Array 1: Array 1 formed either a simple pattern or a complex pattern. Simple patterns tended to group individual dots into one or more larger scale shapes. In the complex pattern set, larger scale shapes were not as evident, with dots haphazardly distributed within the array. Array 1 was presented for 250 msec, which placed Array 2 well outside the SOA window for visible persistence, even at the shortest ISI of 1 msec, thus eliminating the possibility of sensory integration. Short-ISI performance was therefore likely to have been dominated by high-capacity informational persistence. Long-ISI performance (5,000 msec) was likely to have depended on VSTM. Irwin found no advantage for simple versus complex patterns at the shortest ISIs—suggesting dependence on a high-capacity memory that could represent the full array—but a reliable advantage for simple patterns at 5,000-msec ISI—suggesting dependence on a limited-capacity memory that could benefit from coherent, higher order image structure.

Given the Irwin (1991) result, it is possible that figural grouping plays a significant role in the relatively high levels of performance found at long ISIs in the empty cell localization paradigm. However, there are reasons to be cautious in generalizing the Irwin result to the empty cell localization task. The Phillips paradigm requires comparison and change detection, whereas the standard interpretation of the empty cell localization task is that it requires integration of two visual representations to form a composite image (Brockmole et al., 2002). These two tasks have historically depended on different forms of visual representation. Visible persistence supports integration in the empty cell task but impairs change detection in the Phillips task (Irwin, 1992b), whereas informational persistence supports accurate comparison and change detection in the Phillips task but does not support sensory integration in the empty cell task (Di Lollo, 1980). Given the traditional dependence of the two tasks on different forms of visual memory, evidence of a role for VSTM at long ISIs in the Phillips paradigm (Irwin, 1991) does not necessarily imply a role for VSTM at long ISIs in the empty cell localization paradigm.

To test whether figural grouping in VSTM supports apparently high-capacity performance at long ISIs in the empty cell localization task, we replicated the 8–7 condition of Experiment 1 with a manipulation of Array 1 pattern complexity.

Method

Participants. Eight participants from the University of Iowa community completed the experiment. All participants reported 20/20 uncorrected or corrected vision. None participated in previous experiments.

Stimuli. As in Experiment 1, two dot arrays were displayed on each trial. The Array 1 stimuli were drawn from two sets: a simple pattern set and a complex pattern set. Simple and complex patterns were generated on the basis of complexity ratings compiled by Ichikawa (1985). These same ratings were used to group stimuli in Irwin (1991). Ichikawa collected complexity ratings for stimuli highly similar to those used in Experiment 1. Each stimulus was a 4×4 grid with eight of the cells filled by dots. A total of 140 different patterns were rated by participants, and the stimuli were ordered from least complex (Item 1) to most complex (Item 140). The set of 40 simple pattern stimuli for Experiment 2 were Items 11–50 from Ichikawa (1985). The set of 40 complex pattern stimuli were Items 101–140 from Ichikawa. Figure 4 shows a sample stimulus from the simple pattern set and from the complex pattern set. The Array 2 stimuli were constructed by randomly filling seven of the eight cells not filled in Array 1.

Apparatus. The apparatus was the same as that in Experiment 1.

Procedure. The procedure was the same as that in Experiment 1, with the following exceptions. The ISI conditions were limited to 0, 100, 750, and 2,500 msec. These were chosen to mark the major performance landmarks in Experiment 1: sensory integration (0-msec ISI), decay of visible persistence and maximum interruption masking (100-msec ISI), midrecovery (750-msec ISI), and full recovery (2,500-msec ISI). The participants first completed a practice session of 16 trials, two in each of the eight conditions created by the 2 (Array 1 pattern complexity) \times 4 (ISI) design. Practice items were not used in the experimental session. The participants then completed four blocks of trials in the experimental session. Each block contained 80 trials, one for each of the 80 Array 1 stimuli. In each block, 10 trials were presented in each of the eight conditions, randomly intermixed. Across the four blocks, each stimulus was presented four times. The participants completed a total of 320 trials, one for each of the 80 stimuli in each of the four ISI conditions. Between participants, stimulus–ISI assignments were rotated be-

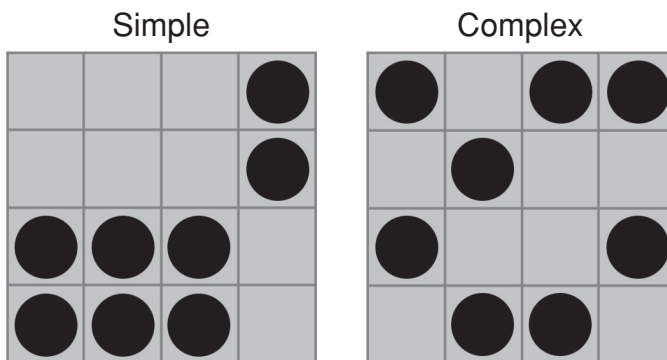


Figure 4. Sample Array 1 stimulus from the simple pattern set (left) and from the complex pattern set (right) in Experiment 2.

tween blocks by Latin square, so that each stimulus–ISI assignment appeared equally often in each block.

The participants were instructed in the same manner as in Experiment 1. They were not informed of the complexity manipulation. The entire session lasted approximately 40 min.

Results and Discussion

Localization performance was examined as a function of Array 1 complexity and ISI. These data are presented in Figure 5. First, there was a reliable main effect of ISI [$F(3,21) = 30.11, p < .001$], and there was a reliable main effect of complexity [$F(1,7) = 66.77, p < .001$]. As is evident from Figure 4, complexity had little effect on performance at 0-msec ISI but had a very large effect on performance at longer ISIs. This pattern was confirmed by a reliable interaction between complexity and ISI [$F(3,21) = 12.68, p < .001$]. Comparing 0-msec ISI to each of the other individual ISI conditions, the interaction between complexity and ISI was reliable for each comparison [0 vs. 100, $F(1,7) = 8.75, p < .05$; 0 vs. 750, $F(1,7) = 19.78, p < .005$; 0 vs. 2,500, $F(1,7) = 40.76, p < .001$]. Examining the complexity effect, there was no complexity difference at 0-msec ISI [$F(1,7) = 2.21, p = .18$], but there was a reliable advantage for the simple pattern condition at 100-msec ISI [$F(1,7) = 12.11, p < .05$], 750-msec ISI [$F(1,7) = 32.82, p < .001$], and 2,500-msec ISI [$F(1,7) = 73.14, p < .001$].

In the simple pattern condition, complete recovery was observed, with localization performance at 2,500-msec ISI numerically higher (83.8% correct) than localization performance at 0-msec ISI (80.0% correct); the difference was not statistically reliable ($F < 1$). In the complex pattern condition, recovery was much more limited, with 43.5% correct at 2,500-msec ISI. Figure 5 also plots

the data from the 8–7 condition of Experiment 1. Experiment 1 arrays were randomly generated. The average complexity of Array 1 stimuli from Experiment 1 was intermediate between the complexity of the complex and simple sets of Experiment 2, since both of these sets were chosen from the extremes of the complexity distribution. Consistent with the figural grouping hypothesis, mean localization accuracy at long ISIs for random arrays in the 8–7 condition fell between performance for the complex and simple arrays of Experiment 2.

In summary, stimulus complexity did not have a significant effect on empty cell localization at 0-msec ISI but had a large effect on localization at longer ISIs, consistent with the finding of Irwin (1991) using the Phillips change detection paradigm. Thus, the Experiment 2 results support the hypothesis that apparently high-capacity visual memory at long ISIs results, at least in part, from figural grouping in VSTM. One potentially puzzling result from Experiment 2 is the complexity effect at 100-msec ISI, which is earlier than one would expect to find on the basis of the time course of recovery in the standard empty cell task (Experiment 1; Brockmole et al., 2002). A possible explanation for this effect is that at short ISIs, the participants formed a figural representation of Array 1 in the simple array condition that was relatively resistant to backward masking by Array 2. This possibility is consistent with evidence that metacontrast masking is reduced when discrete stimuli can be grouped into a higher level figure (Ramachandran & Cobb, 1995). Note, however, that long ISIs were of principal concern in this experiment, and differential backward masking cannot account for the complexity effect at long ISIs. Masking interference should have been essentially eliminated by 750-msec ISI and would have been nonexistent at 2,500-msec ISI.

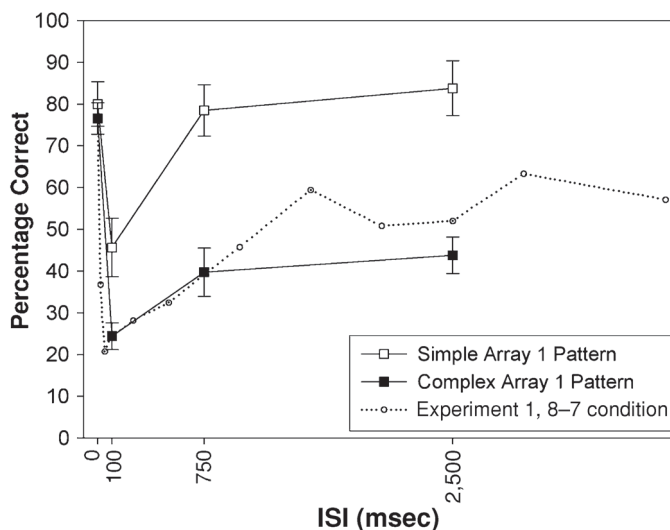


Figure 5. Localization accuracy as a function of ISI and pattern complexity in Experiment 2. For comparison, the dotted line plots performance for the 8–7 condition in Experiment 1. Error bars represent standard errors of the means.

EXPERIMENT 3

In Experiment 2, we examined the properties of the visual memory representation supporting long-ISI performance. Two related questions remain. First, what is the process by which Array 1 and Array 2 are related to each other so as to localize the empty cell? Second, how do representation and process interact to produce the pattern of performance as a function of ISI?

At very short ISIs, the mechanism supporting localization performance is relatively well understood: Sensory processing of Array 2 integrates with (cannot be temporally segregated from; Dixon & Di Lollo, 1994) the persisting sensory processing of Array 1 (visible persistence). At 100-msec ISI, however, visible persistence of Array 1 has decayed before the arrival of Array 2, eliminating the possibility of sensory integration. In addition, Array 2 interferes with the consolidation of Array 1 into a stable form in VSTM through interruption masking (Enns & Di Lollo, 2000). Thus, empty cell localization performance at 100-msec ISI is very poor indeed.

What then accounts for improvement in performance at longer ISIs? Brockmole et al. (2002) argued that at longer ISIs, a VSTM representation of Array 1 is formed. This VSTM representation is then projected back onto visual areas of the brain responsible for early sensory processing (such as V1), a process of *image formation* that they describe as akin to mental imagery. In essence, Brockmole et al. claim that participants re-create the sensory-level representation of Array 1 on the basis of an Array 1 representation maintained in VSTM. When Array 2 appears, it directly integrates with the projected sensory representation of Array 1. Thus, the mechanism of integration is identical at very short ISIs and long ISIs, except that in the case of very short ISIs, the Array 1 representation is the brief sensory persistence of Array 1, whereas in the case of long ISIs, the Array 1 representation is a sensory-level projection of the contents of VSTM. Brockmole et al. claimed that the time course of recovery in the empty cell task suggests that this process of deriving a VSTM representation and then projecting it back to early visual areas takes approximately 1,300 msec to complete.

The Brockmole et al. (2002) account has three stages: the construction of a stable VSTM representation, the projection of that representation back onto sensory cortex, and the integration of that projected representation with sensory activation of Array 2. Recently, Jiang and colleagues (Jiang, 2004; Jiang & Kumar, 2004; Jiang et al., 2005) have questioned whether the second two stages—projection back to sensory cortex and sensory integration—are necessary. Jiang et al. argued that localization performance at long ISIs is supported by comparison within VSTM itself. In this view, participants use a strategy of attending to the empty locations of Array 1, forming a VSTM representation of the Array 1 negative space, which they then compare with the Array 2 dot pattern. The difference between these two representations (i.e., the one Array 1 negative space cell that does not have an Array 2

dot) is the empty cell. In essence, Jiang et al. claim that by remembering the negative space pattern, participants turn the empty cell localization task into a change detection task, much like the Phillips paradigm. This account is attractive in its parsimony, since it does not depend on additional mechanisms of sensory projection and integration. In addition, the negative space comparison hypothesis relies on well-established properties of VSTM: the ability to support comparison and change detection.

In Experiment 3, a pattern complexity manipulation was used to discriminate between the integration and negative comparison accounts of long-ISI performance. Under the Brockmole et al. (2002) integration view, at long ISIs, participants form a VSTM representation of Array 1 for sensory projection and integration with Array 2. Under the Jiang et al. (2005) negative comparison view, participants form a VSTM representation of the Array 1 negative space for comparison with Array 2. The integration view predicts that manipulations facilitating the representation of the Array 1 dot pattern will improve long-ISI performance. The negative comparison view predicts that manipulations facilitating the representation of Array 1 negative space will improve long-ISI performance. Experiment 2 clearly demonstrated that simple patterns facilitate array representation at long ISIs, but simple Array 1 dot patterns tended to produce simple negative space patterns, and complex Array 1 patterns tended to produce complex negative space patterns (see Figure 4), so Experiment 2 could not discriminate between the two hypotheses.

In Experiment 3, we created asymmetric stimuli in which the Array 1 dot pattern was simpler than the negative space pattern, and vice versa. Figure 6 shows a sample pair of Array 1 stimuli. For the left-hand stimulus, the dot pattern is simpler than the negative space pattern (array simple condition). The right hand stimulus is the complement of the left-hand stimulus, and complexity is therefore reversed: The negative space pattern is simpler than the dot pattern (negative simple condition). Thirty pairs of complementary Array 1 stimuli were used in Experiment 3, ensuring that precisely the same patterns were formed for the dot array and negative space in the array simple and negative space simple conditions. In all conditions, Array 2 was formed by randomly filling seven of the eight cells not filled by Array 1. Under the Brockmole et al. (2002) integration view, long-ISI performance should be higher when the Array 1 dot pattern is relatively simple. Under the Jiang et al. (2005) negative comparison view, long-ISI performance should be higher when the Array 1 negative space pattern is relatively simple.

Method

Participants. Twenty-four participants from the University of Iowa community completed the experiment. All participants reported 20/20 uncorrected or corrected vision. None participated in previous experiments.

Stimuli. Fifty-five pairs of complementary eight-dot arrays were created, chosen so that one pattern appeared simpler than its complement (see Figure 6). Four raters viewed each pair, with the two complementary arrays displayed side by side, as in Figure 6

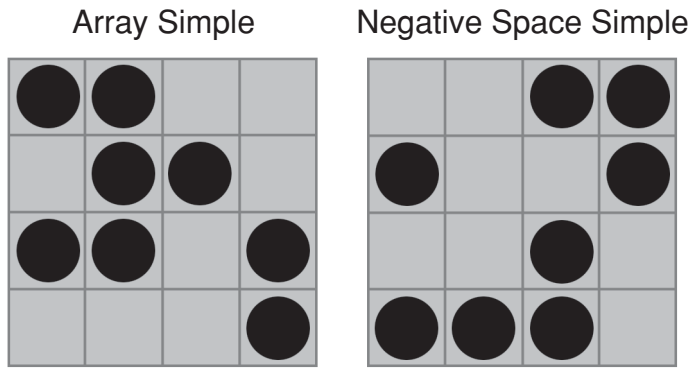


Figure 6. Sample pair of complementary Array 1 stimuli from Experiment 3. In the left Array 1 stimulus, the dot array forms a simpler pattern than the negative space (array simple condition). In the complementary Array 1 stimulus on the right, the negative space forms a simpler pattern than the array (negative space simple condition).

(without the labels, of course). The assignment of arrays to the left or right side of the screen was determined randomly and counter-balanced across raters. For each pair, the raters examined the pair as long as necessary and chose the pattern that appeared to be simpler, rating the ease of decision on a 9-point scale. To ensure that simplicity decisions were based on ease of figural processing, the raters were told that simple patterns tended to form coherent, connected shapes. For 40 of the 55 pairs, all 4 raters agreed on the simpler pattern. In each of these cases, the simpler pattern was the one chosen as simpler by the 2 research assistants who created the stimulus set, so the complexity decision was consistent across six sets of judgments. Of these 40, the 30 pairs with the highest ease of decision rating were selected for use in Experiment 3. Thus, for half of the 60 Array 1 stimuli, the dot array pattern was simpler than the negative space pattern, and for the other half (their complements) complexity was reversed. In each case, Array 2 was formed by randomly filling seven of the eight cells not filled in Array 1. In all other respects, the stimuli were the same as those in Experiments 1 and 2.

Apparatus. The apparatus was the same as that in Experiments 1 and 2.

Procedure. The procedure was the same as that in Experiments 1 and 2, with the following exceptions. The ISI conditions were limited to 500, 1,500, and 2,500 msec. Each of these ISIs was in the range of ISIs showing a simple pattern advantage in Experiment 2. The participants first completed a practice session of 30 trials. All practice arrays were randomly generated. The participants then completed three blocks of trials in the experimental session. Each block contained 108 trials. Forty-eight of these were filler trials with randomly generated arrays (16 in each ISI condition). For the 60 critical trials in a block, all 60 Array 1 experimental stimuli were presented. The 60 trials were evenly divided across the ISI conditions. Across the three blocks, each Array 1 stimulus appeared in each ISI condition. Between participants, stimulus–ISI assignments were rotated between blocks by Latin square, so that each stimulus–ISI assignment appeared equally often in each block.

The participants were instructed in the same manner as in Experiment 1. They were not informed of the complexity manipulation. The entire session lasted approximately 40 min.

Results and Discussion

Localization performance was examined as a function of Array 1 complexity (array simple, negative space simple) and ISI (500, 1,500, and 2,500 msec). Trials with random arrays were filler trials and were not directly relevant

to the question of localization mechanism. For the sake of completeness, however, they are reported in Figure 7. First, there was a reliable main effect of ISI [$F(2,46) = 4.96, p < .05$], with higher performance at longer ISIs. The critical effect of complexity was reliable [$F(1,23) = 5.42, p < .05$], with higher performance when the Array 1 negative space was simple (54.9%) than when the Array 1 dot pattern was simple (51.1%). These two factors produced a reliable interaction [$F(2,46) = 3.21, p < .05$], driven by the absence of an advantage for the negative simple condition at 2,500-msec ISI.

There was a surprising number of low-performing participants in this experiment. Six participants performed below 40% correct across conditions. The purpose of the experiment was to examine the strategies and mechanisms used to perform the localization task accurately at long ISIs. Thus, it is important to ensure that the effects were driven by participants who were able to perform the task accurately and not by participants who were unable to perform the task accurately. Figure 8 displays the results after elimination of the 6 participants who performed below 40% correct. The pattern of data obtained in this high-performing subset of participants was the same as that obtained in the full data set. The advantage for negative simple over array simple was still reliable [$F(1,17) = 6.70, p < .05$], as was the effect of ISI [$F(2,34) = 3.73, p < .05$]. However, the interaction between complexity and ISI was not reliable [$F(2,34) = 1.31, p = .28$].

If accurate performance at long ISIs is supported by a strategy of attending to and remembering the negative space pattern of Array 1, then the negative simple advantage should have been larger in later blocks, after the participants had an opportunity to develop a consistent strategy. Figure 9 shows the data for the high-performing group as a function of block. The advantage for the negative simple condition was most pronounced in the third block of trials. In addition, the absence of an overall effect of complexity condition at 2,500-msec ISI appears to have been generated by poor Block 1 performance in the

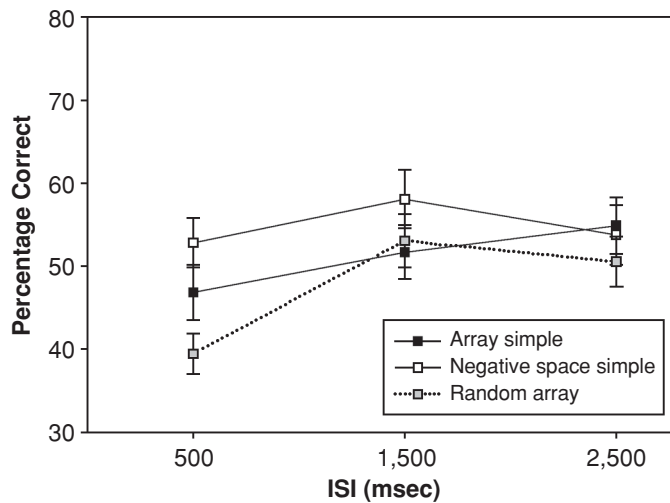


Figure 7. Localization accuracy as a function of ISI in the array simple, negative space simple, and random array conditions in Experiment 3. Error bars represent standard errors of the means.

negative simple condition at 2,500-msec ISI. By Block 3, all ISIs demonstrated an advantage for the negative simple condition.

Although there was a reliable advantage for the negative simple condition, the difference between complexity conditions was smaller than the complexity effect observed in Experiment 2. This is not surprising. The complexity manipulation in Experiment 3 was constrained by the use of complementary Array 1 stimulus pairs. The simple patterns were significantly less simple than the simple patterns of Experiment 2, and the complex patterns were significantly simpler than the complex patterns in Experiment 2 (compare Figures 4 and 6). Accordingly, performance in the negative simple condition in Experiment 3 was lower than in the simple condition of Experiment 2,

and performance in the array simple (negative complex) condition in Experiment 3 was higher than in the complex condition of Experiment 2. Given the more subtle manipulation of complexity in Experiment 3, the advantage for the negative simple condition provides strong support for the negative space comparison hypothesis proposed by Jiang et al. (2005). Thus, we must revise our interpretation of Experiment 2. Instead of grouping multiple dots into higher order figures, participants appear to group multiple empty cells into higher order figures.

GENERAL DISCUSSION

In the present study, we examined the memory systems and mechanisms supporting performance in an empty cell

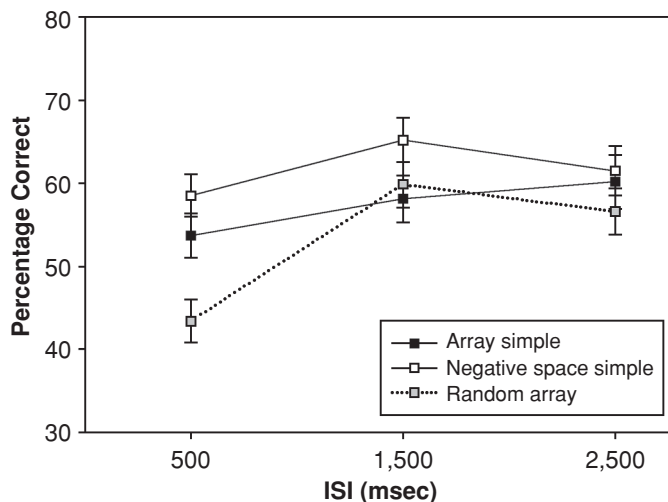


Figure 8. Localization accuracy for the 18 participants performing above 40% correct in Experiment 3. Error bars represent standard errors of the means.

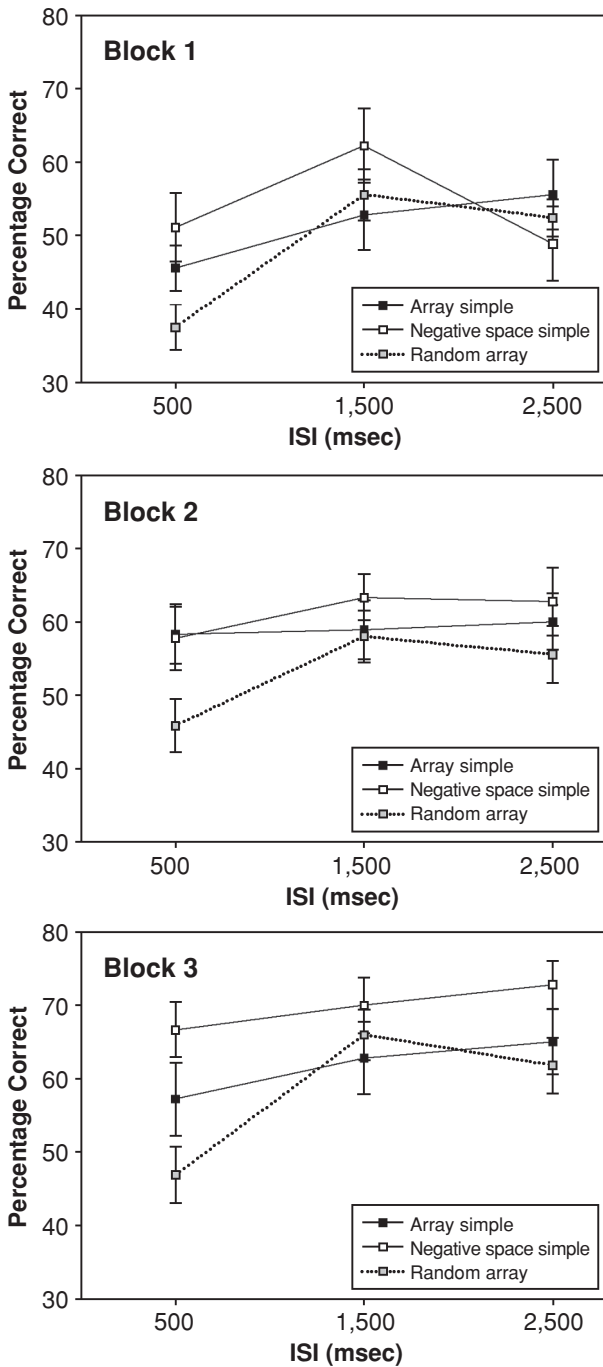


Figure 9. Localization accuracy for the 18 participants performing above 40% correct as a function of block in Experiment 3. Error bars represent standard errors of the means.

localization paradigm. In Experiment 1, we replicated the Brockmole et al. (2002) effect of relatively accurate empty cell localization at long ISIs. In Experiment 2, we tested whether this apparently high-capacity performance is supported by figural grouping in VSTM using a manipulation of pattern complexity. Array 1 formed either a simple pattern or a complex pattern. If performance at long ISIs is

supported by figural grouping in VSTM, then long-ISI performance should have been higher when figure extraction was relatively easy (simple pattern condition) relative to when figure extraction was difficult (complex pattern condition). Consistent with the figural grouping hypothesis, a simple pattern advantage was observed at long ISIs. This result suggests that performance at long ISIs is supported by a limited-capacity, object-based visual memory that is sensitive to higher order image structure: VSTM. The effect of pattern complexity provides a means to reconcile apparently high-capacity memory at long ISIs with evidence that VSTM has a limited capacity of three or four objects. At long ISIs, Array 1 is likely represented in VSTM as one or more higher order objects, each containing information from more than one array element.

The complexity effect at long ISIs contrasts with the absence of a complexity effect at 0-msec ISI. This latter result is consistent with the idea that brief forms of sensory memory are maintained in early, retinotopic stages of visual processing, before the computation of higher order image structure. From a slightly different perspective, if visible persistence maintains a complete representation with sufficient capacity to represent the positions of essentially all of the dots, then performance will be accurate, regardless of how the dots are configured. The reduced performance at long ISIs for complex patterns demonstrates that, unlike visible persistence, VSTM does not have the capacity to represent all the array elements individually (Irwin, 1991; Phillips, 1974).

In Experiment 3, we examined the means by which the empty cell is localized at long ISIs. We pitted the Brockmole et al. (2002) hypothesis that localization is supported by sensory projection of Array 1 and direct integration with Array 2 against the Jiang et al. (2005) hypothesis that localization is supported by memory for the Array 1 negative space followed by comparison with Array 2 and change detection. The relative complexity of the Array 1 dot pattern and negative space pattern was manipulated. Localization performance was reliably higher when the Array 1 negative space pattern was simple relative to when the Array 1 dot pattern was simple, supporting the Jiang et al. negative comparison view. These results suggest that the same set of mechanisms support long-ISI performance both in the Phillips change detection paradigm (Irwin, 1991) and in the empty cell localization paradigm: Both depend on VSTM comparison and change detection. In the case of the Phillips task, participants remember the dot array pattern. In the empty cell task, however, participants remember the negative space pattern.

These results converge with the results reported in Jiang et al. (2005). In that study, participants first viewed an array of seven dots. After a variable ISI, participants performed one of two tasks: the empty cell localization task or a modified Phillips change detection task. In the empty cell task, a second array of eight dots was presented, and participants attempted to determine the empty cell. In the change detection task, Array 1 was re-presented along with one new dot, and participants responded to indicate which dot was new. Under the Brockmole et al.

(2002) view, at long ISIs both tasks should depend on a VSTM representation of Array 1 dot pattern. Under the Jiang et al. view, the two tasks should depend on different representations: a representation of the Array 1 negative space for the empty cell task and a representation of the Array 1 dot pattern for the change detection task. In different blocks, the proportion of trials in each task was manipulated. Consistent with the negative space comparison explanation of empty cell localization at long ISIs, successful performance on one task was achieved at the expense of the other, suggesting they were supported by different representations in VSTM.

In conclusion, Brockmole et al. (2002) found a striking recovery in empty cell localization performance at longer ISIs. In that study, long ISIs produced levels of performance that rivaled performance at very short ISIs, when localization is known to be based on high-capacity visible persistence. The present study demonstrated that long-ISI performance depends on VSTM and is supported by the grouping of individual array elements into one or more higher order figures. These results reconcile the apparently high-capacity performance at long ISIs with the standard, limited-capacity model of VSTM. In addition, at long ISIs, participants appear to perform the task not by remembering the Array 1 dot pattern but rather by remembering the Array 1 pattern of empty cells, which can then be compared with Array 2 to isolate the missing dot.

REFERENCES

- AVERBACH, E., & CORIELL, A. S. (1961). Short-term memory in vision. *Bell System Technical Journal*, **40**, 309-328.
- BROCKMOLE, J. R., WANG, R. F., & IRWIN, D. E. (2002). Temporal integration of visual images and visual percepts. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 315-334.
- COLTHEART, M. (1980). The persistences of vision. *Philosophical Transactions of the Royal Society of London: Series B*, **290**, 269-294.
- DI LOLLO, V. (1977). Temporal characteristics of iconic memory. *Nature*, **267**, 241-243.
- DI LOLLO, V. (1980). Temporal integration in visual memory. *Journal of Experimental Psychology: General*, **109**, 75-97.
- DI LOLLO, V., & HOGBEN, J. H. (1987). Suppression of visible persistence as a function of spatial separation between inducing stimuli. *Perception & Psychophysics*, **41**, 345-354.
- DIXON, P., & DI LOLLO, V. (1994). Beyond visible persistence: An alternative account of temporal integration and segregation in visual processing. *Cognitive Psychology*, **26**, 33-63.
- ENNS, J. T., & DI LOLLO, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, **4**, 345-352.
- ENNS, J. T., & VISSER, T. A. W. (2001). The role of attention in temporal integration. *Perception*, **30**, 135-145.
- ERIKSEN, C. W., & COLLINS, J. F. (1967). Some temporal characteristics of visual pattern perception. *Journal of Experimental Psychology*, **74**, 476-484.
- HENDERSON, J. M. (1997). Transsaccadic memory and integration during real-world object perception. *Psychological Science*, **8**, 51-55.
- HOGBEN, J. H., & DI LOLLO, V. (1974). Perceptual integration and perceptual segregation of brief visual stimuli. *Vision Research*, **14**, 1059-1069.
- HOLLINGWORTH, A. (2004). Constructing visual representations of natural scenes: The roles of short- and long-term visual memory. *Journal of Experimental Psychology: Human Perception & Performance*, **30**, 519-537.
- HOLLINGWORTH, A. (2005). The relationship between online visual representation of a scene and long-term scene memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 396-411.
- HOLLINGWORTH, A. (in press). Visual memory for natural scenes: Evidence from change detection and visual search. *Visual Cognition*.
- HOWE, E., & JUNG, K. (1986). Immediate memory span for two-dimensional spatial arrays: Effects of pattern symmetry and goodness. *Acta Psychologica*, **61**, 37-51.
- ICHIKAWA, S. (1985). Quantitative and structural factors in the judgment of pattern complexity. *Perception & Psychophysics*, **38**, 101-109.
- IRWIN, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology*, **23**, 420-456.
- IRWIN, D. E. (1992a). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 307-317.
- IRWIN, D. E. (1992b). Visual memory within and across fixations. In K. Rayner (Ed.), *Eye movements & visual cognition: Scene perception & reading* (pp. 146-165). New York: Springer.
- IRWIN, D. E., & ANDREWS, R. (1996). Integration and accumulation of information across saccadic eye movements. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 125-155). Cambridge, MA: MIT Press.
- IRWIN, D. E., & YEOMANS, J. M. (1986). Sensory registration and informational persistence. *Journal of Experimental Psychology: Human Perception & Performance*, **12**, 343-360.
- JIANG, Y. (2004). Time window from visual images to visual short-term memory: Consolidation or integration? *Experimental Psychology*, **51**, 45-51.
- JIANG, Y., & KUMAR, A. (2004). Visual short-term memory for two sequential arrays: One integrated representation or two separate representations? *Psychonomic Bulletin & Review*, **11**, 495-500.
- JIANG, Y., KUMAR, A., & VICKERY, T. J. (2005). Integrating sequential arrays in visual short-term memory. *Experimental Psychology*, **52**, 39-46.
- LUCK, S. J., & VOGEL, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, **390**, 279-281.
- PALMER, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.
- PASHLER, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, **44**, 369-378.
- PHILLIPS, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, **16**, 283-290.
- POMERANTZ, J. R. (1977). Pattern goodness and speed of encoding. *Memory & Cognition*, **5**, 235-241.
- RAMACHANDRAN, V. S., & COBB, S. (1995). Visual attention modulates metacontrast masking. *Nature*, **373**, 66-68.
- SEBRECHTS, M. M., & GARNER, W. R. (1981). Stimulus-specific processing consequences of pattern goodness. *Memory & Cognition*, **9**, 41-49.
- SPELTING, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, **74** (11, Whole No. 498).
- STANDING, L., CONEZIO, J., & HABER, R. N. (1970). Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. *Psychonomic Science*, **19**, 73-74.
- YEOMANS, J. M., & IRWIN, D. E. (1985). Stimulus duration and partial report performance. *Perception & Psychophysics*, **37**, 163-169.

NOTES

1. If in the Phillips (1974) paradigm Array 1 is presented for a very brief duration, then short-ISI performance does fall within the SOA range of visible persistence, and performance is very poor: The two arrays fuse, and it becomes difficult to determine whether or not Array 1 has changed (Irwin, 1992b).
2. If the duration of Array 1 is increased to over 130 msec, performance is very poor, even at 0-msec ISI, demonstrating that visible persistence is locked to stimulus onset, not offset (Di Lollo, 1980).