

# Eye movements and the integration of visual memory and visual perception

JAMES R. BROCKMOLE

*Michigan State University, East Lansing, Michigan*

and

DAVID E. IRWIN

*University of Illinois at Urbana-Champaign, Champaign, Illinois*

Because visual perception has temporal extent, temporally discontinuous input must be linked in memory. Recent research has suggested that this may be accomplished by integrating the active contents of visual short-term memory (VSTM) with subsequently perceived information. In the present experiments, we explored the relationship between VSTM consolidation and maintenance and eye movements, in order to discover how attention selects the information that is to be integrated. Specifically, we addressed whether stimuli needed to be overtly attended in order to be included in the memory representation or whether covert attention was sufficient. Results demonstrated that in static displays in which the to-be-integrated information was presented in the same spatial location, VSTM consolidation proceeded independently of the eyes, since subjects made few eye movements. In dynamic displays, however, in which the to-be-integrated information was presented in different spatial locations, eye movements were directly related to task performance. We conclude that these differences are related to different encoding strategies. In the static display case, VSTM was maintained in the same spatial location as that in which it was generated. This could apparently be accomplished with covert deployments of attention. In the dynamic case, however, VSTM was generated in a location that did not overlap with one of the to-be-integrated percepts. In order to “move” the memory trace, overt shifts of attention were required.

The world contains more information than an observer can sample and process at any one moment. Physical properties of the eye limit the scope of perception by constricting the proportion of the environment from which information is received and the quality of the information extracted over different spatial regions. Foveal vision, corresponding to the center of gaze, resolves high spatial frequency and color components of a scene but covers only 2° of the visual world. In contrast, peripheral vision is tuned to lower spatial frequencies and derives degraded color information. In addition to these physical limitations that restrict sampling space, cognitive limitations on memory and attention set bounds to information-processing capabilities. Visual short-term memory (VSTM) is a limited capacity store estimated to hold three to five items (Irwin, 1992; Irwin & Andrews, 1996; Luck & Vogel, 1997). Although the scope of attention may vary (e.g., C. W. Eriksen & St. James, 1986), it has been estimated to have a maximum span of as little as 1° of visual

angle (B. A. Eriksen & C. W. Eriksen, 1974), with more diffuse states of attention leading to less refined processing (Shulman & Wilson, 1987). To view a scene in its entirety, then, observers shift their attention and their gaze from place to place.

Given the temporal extent of visual processing, information extracted from the world at one moment must be analyzed in conjunction with that which was obtained previously. What mechanism enables this analysis? A recent proposal is that visual percepts can be integrated with the contents of VSTM (Brockmole, Irwin, & Wang, 2003; Brockmole & Wang, 2003; Brockmole, Wang, & Irwin, 2002). That is, when an observer generates a representation of a visual stimulus in memory, a subsequently perceived stimulus can be directly incorporated into the existing representation. Thus, discontinuous, but nevertheless related, visual information can be represented as a single unit in memory, thereby cognitively linked once independent pieces of information.

Support for the *memory-percept integration hypothesis* has been obtained using a temporal integration paradigm (e.g., Di Lollo, 1980; C. W. Eriksen & Collins, 1967). Two arrays of dots were serially presented within a square grid. Together, the two arrays filled all but one space in the grid; subjects had to report the unfilled grid location. As one would expect, the time that separated the arrays was critical. If the delay was very short (<50 msec), sensory memory supported the perceptual

---

J.R.B. is partially supported by NSF Grant DGE-0114378. D.E.I. is supported by NSF Grant BCS 01-32292. We thank Aaron Benjamin, Arthur Kramer, Brian Ross, Frances Wang, Robert Logie, and two anonymous reviewers for their insightful comments on this research. We also thank Nicholas Cassavaugh for technical assistance and Shawn Bolin for help in collecting data. Address correspondence to J. R. Brockmole, Department of Psychology, Michigan State University, East Lansing, MI 48824 (e-mail: jim@eyelab.msu.edu).

integration of the arrays, and they appeared to be presented simultaneously. If the delay was extended to 100–300 msec, however, accuracy was poor, since the second array masked the first. In the past, this finding has been taken as evidence that visual integration is impossible once sensory memory has decayed (see Coltheart, 1980; Phillips, 1974). However, given that, in the absence of sensory memory, successful performance requires both the formation and the retention of a longer-lasting representation, integration might fail in these short delay situations, because of a failure to consolidate the initial representation into a stable and durable format. Indeed, Brockmole et al. (2002) demonstrated that when the delays are longer than 300 msec, performance begins to improve and reaches an asymptotic level between 1,000 and 1,500 msec. The improvement in accuracy over time was attributed to the generation of a mental representation of the lead array, because the increase in accuracy was accompanied by a reduction in errors associated with the lead array (i.e., erroneously selecting a position previously occupied by the first array as the position of the missing dot) over the same time period. In excess of 95% of the variance in accuracy improvement was accounted for by the decrease in Array 1 error (Brockmole et al., 2002).

Why postulate that the two arrays were integrated into a single representation in VSTM when, intuitively, there are several strategies that an observer could use to solve the missing dot task? For example, the first array might be used to *visually mark* (see Watson & Humphreys, 2000) grid positions as locations that cannot constitute the correct answer. As a result, these positions might be inhibited, and attention directed to the positions that Array 1 left empty, enhancing the detection of the space left unfilled by the second array (see Jiang & Kumar, 2004; Jiang, Kumar, & Vickery, 2005). Another possibility is that two separate memory representations of each stimulus array may be created in memory. If the two memory traces could then be compared “in the mind’s eye,” the combination or integration of two stimulus arrays need not be postulated. We investigated and rejected these possible explanations in two recent studies, described next.

To address the visual-marking hypothesis, a direct measure of the locus of spatial attention during the temporal integration task was used (Brockmole et al., 2003). During the interstimulus interval (ISI) separating the arrays, two-alternative forced choice probes appeared in positions either filled by or left empty by the first array. Probes that occupied positions filled by the first array enjoyed a processing benefit, in that they were responded to more quickly than probes that were in positions left empty by the first array. Thus, attention was selectively allocated to the positions filled by Array 1 during the time between arrays; no evidence for Array 1 inhibition or Array 2 priming was found. In a separate study (Brockmole et al., 2002), the two-representation hypothesis was assessed by examining the time required to consolidate

the second array into VSTM. Although 1,000–1,500 msec were required to consolidate Array 1 into VSTM, reaction time analyses demonstrated that the time needed to process the second array at long ISIs was about equal to that needed during perceptual integration (when the ISI was 0 msec), suggesting that Array 2 is not submitted to the same consolidation processes as Array 1. Because attention is allocated to the first array throughout the delay separating the arrays and because the second array does not undergo a slow consolidation process, it was concluded that over a period of 1,000–1,500 msec, observers can generate a VSTM representation of one stimulus and directly integrate subsequent perceptual information with that representation.

In other experiments, the flexibility, or abstractness, of the VSTM representations involved in integration was considered by examining whether integration can occur when the to-be-integrated stimuli do not match in terms of their spatial properties (Brockmole & Wang, 2003). Stimuli can be integrated even if they are presented in different size scales and orientations. This result indicates that during integration, attention is not dedicated to particular areas of space independently of the stimuli that occupied them: The memory representation can be modified to account for changes in stimulus structure and “projected” to different locations on the display. Expressed another way, the memory trace of the lead stimulus is not spatiotopically or retinotopically specific; rather, observers are able to retain their memory for the lead array abstractly, enabling it to be transformed to accommodate spatial changes, thereby providing spatial flexibility in the integration process.

An important question remaining concerns the time course of integration. Specifically, why does optimal integration require up to 1,500 msec of separation between visual stimuli? We assume that this time reflects the speed with which the lead stimulus is consolidated into VSTM. In this article, we focus on uncovering the mechanisms that underlie this consolidation process by examining observers’ eye movements during integration. Specifically, we ask whether overt shifts of attention to the to-be-remembered elements are required for integration. That is, in order to be included in the VSTM representation, must a stimulus element—or in the case of a brief presentation, its location—be overtly attended, or fixated? In essence, the issue is one of whether or not the initial stimulus is maintained by eye movements that rehearse its spatial structure. On the basis of temporal characteristics of eye movements and assuming an average fixation duration of 200 msec and an average saccade duration of 25 msec, it would take 1,575 msec to individually fixate seven elements, a time course matching that in integration experiments requiring the maintenance of seven items. If the ISI is less than this value, fewer elements can be overtly attended or rehearsed, and integration will suffer.

The idea that eye movements are important to integration stems from research in which the use of eye move-

ments during perception has been compared with the maintenance or recall of information from VSTM. In his theoretical work on cell assemblies, Hebb (1968) suggested that visual memory (i.e. imagery) may constitute a top-down reinstatement of perceptual processes. Thus, just as eye movements are necessary to sequentially view different parts of a visual stimulus, so too may they be necessary to sequentially assemble image parts to construct complete VSTM representations of those same stimuli. Specifically, he argued that eye movements may serve an “organizing function” in both visual perception and visual memory (see Neisser, 1967, for similar arguments). The recent interest in eye movements and visual memory has led to a variety of empirical papers in which Hebb’s proposal has been tested over the past few years.

Brandt and Stark (1997) had observers study random patterns of black-and-white squares arranged in a matrix and to later maintain the patterns in visual memory after their disappearance. Systematic scanpaths were found during both the study and the memory maintenance phases of the experiment, in that the eyes were directed to the dark, or “filled,” regions of the matrix. For each stimulus, the eye movements during the memory phase were more closely related to the eye movements made during the study of that stimulus than to those made during the study of other stimuli. Spivey and Geng (2001) noted that while observers answered questions about objects from a since-removed visual display, the eyes were spontaneously directed to the region of space once housing the object. Finally, Laeng and Teodorescu (2002) manipulated an observer’s ability to make eye movements during the study and recall of visual stimuli. Subjects who maintained a single fixation during the perception phase also maintained a single fixation during the recall phase. Observers who were allowed to move their eyes freely during the perceptual phase moved their eyes during the recall phase in systematic ways. The percentage of time fixating any one location during perception was correlated with the percentage of time that region of space was fixated during the recall phase, as was the order in which locations were fixated during the perceptual and recall phases. The strength of these relationships predicted performance accuracy. Together, these studies suggest that eye movements during the maintenance, recall, and processing of previously viewed visual stimuli are not random but reflect the content of the visualized scene in the same way that eye movements during scene viewing are related to the scene’s content (Buswell, 1935; Yarus, 1967). That is, eye scanpaths engaged by visual memory recapitulate those during perception—the interrogation of VSTM is accompanied by an oculomotor inspection of space. The functional role of these eye movements may be to spatially link mental and retinal images.

The research above suggests that processes that call on memory to recreate the visual world operate similarly to processes that call on direct perception of the world. As such, the success of memory–percept integration may

be closely linked to eye movement behavior. Therefore, determining the relationship between eye movements and the integration of visual memories and percepts is important for a complete understanding of integration as a mechanism that links temporally distinct but conceptually related information in VSTM. If fixation on a stimulus element is required for integration, integration can take place only between stimuli that enjoy overt selection and processing. This selection process would restrict the possible aspects of a scene that would be linked in VSTM and elucidate a circumstance necessary for information to be linked in VSTM. On the other hand, independence between eye movements and integration performance would suggest that covert attention alone is sufficient for integration to occur. A specific link between covert attention shifts and the consolidation of VSTM representations has not been established, because the vast majority of experiments on memory consolidation have not monitored eye movements, rendering it impossible to dissociate covert and overt attention shifts. Indeed, Neisser (1967) has pointed out that the role of covert attention in memory processes cannot be dismissed, because even during object perception, shifts of attention can occur without ocular motion. A primarily covert selection process would give greater cognitive control over integration to the observer, since he or she would be able to integrate any attended information, not just that overtly scanned, enabling greater flexibility in the integration process.

## EXPERIMENT 1

There are two general means by which attention can be deployed about a scene. First, attention can be assigned to locations or objects covertly—that is, without the aid of eye movements. This kind of attention allocation can be likened to attending to something “out of the corner of your eye” and enables an observer to deploy attention to many locations or objects in parallel. Alternatively, attention can be assigned to objects or locations overtly—that is, by executing eye movements to bring the object of attention to the fovea. Overtly guided attention allocation brings objects or locations into the scope of attention serially, as eye movements serially move from one locus of fixation to another. Overt attention may also provide a rehearsal strategy: Subjects may systematically scan the grid, focusing their fixations on locations previously occupied by a dot from Array 1.

The purpose of Experiment 1 was to determine which mode of attention deployment is used to commit the to-be-remembered grid locations to memory in preparation for integration with a subsequent stimulus. In order to assess whether attention is deployed covertly or overtly during the interval separating the to-be-integrated arrays, eye movement behavior was recorded while subjects engaged in a temporal integration task. If attention is allocated to the to-be-remembered grid locations covertly, performance should not share a strong rela-

tionship with eye behavior. Alternatively, if attention is allocated overtly, specific patterns of eye behavior should be observed. First, as the ISI between arrays increases, so too should the number of eye movements made during the delay, since the subjects would begin to serially scan the display. Second, the increasing number of eye movements should strongly correlate with an increase in the number of grid locations that are fixated. That is, if the subjects attempt to fixate as many Array 1 positions as possible, each eye movement should be directed to a new grid location, with few intragrid position fixations or intergrid position revisitations. Third, on the basis of Brockmole et al. (2003), the eye movements that are made should also be more likely to bring the eyes to fixation on a grid location that was occupied by a dot from the first array. Finally, some researchers have argued that fixation durations reflect the amount of processing that is needed at that location (e.g., Henderson, Weeks, & Hollingworth, 1999; Just & Carpenter, 1980; Rayner, 1998). Assuming that each eye movement is used to assign attention to a new location, the fixation durations associated with each fixation should either remain constant, because equal processing time should be needed at each fixation, or potentially increase, because each fixation would add more information to the maintained memory trace, thereby increasing processing demands.

### Method

#### Subjects

Eight members of the University of Illinois community participated after providing informed consent. All the subjects were naive with respect to the experimental hypotheses and were paid \$6 for participating.

#### Stimuli

The stimuli were very similar to those used in Brockmole et al. (2002). Two unique dot arrays were presented within an enclosed  $4 \times 4$  square grid. The first array contained seven dots, and the second array contained eight dots. Together, the arrays filled all but one square in the grid. The grid was composed of interconnected lines, drawn over the background in such a way that the color within the grid spaces and the area surrounding the grid was the same. The subjects viewed the stimuli at a viewing distance of 57 cm. The total display subtended  $41^\circ$  horizontally and  $31^\circ$  vertically. The square grid subtended  $20^\circ$ . Each square within the grid subtended  $5^\circ$ . Each dot presented in the array subtended  $4.5^\circ$ . The display background was light gray, the grid lines were light blue, and the dots were black.

#### Apparatus

The stimuli were presented at a refresh rate of 60 Hz on a Hitachi Superscan Elite 21 monitor. A Gateway P5-150 microcomputer controlled stimulus presentation and recorded responses. Eye position was sampled at a rate of 250 Hz (every 4 msec) with an Eye-link eyetracker (Senso-Motoric Instruments, Inc.) and was recorded on an IBM microcomputer. This system used video-based infrared oculography to measure eye and head positions. A chinrest was used to help stabilize the head and to ensure constant distance from the display.

#### Design and Procedure

In general, the design and procedure were very similar to those used in Brockmole et al. (2002). On each trial, two dot arrays (seven and eight dots, respectively) were presented sequentially within an enclosed square grid separated by a variable ISI. On any given trial, one position within the grid was never filled, and the subjects were instructed that they were to identify the position of the missing dot.

The trials were divided into blocks on the basis of ISI. The order of blocks was randomized, and the subjects were informed of the ISI duration prior to the start of each block. A block of trials began with a calibration routine in which a fixation point was presented serially in eight locations around the perimeter of the display, as well as one location in the center of the display. The subjects' eye positions were monitored during this procedure, which served to calibrate the output of the eyetracker against spatial position. Following successful calibration, the subjects were informed of the ISI duration that would be used for the trials in that block. The subjects indicated their readiness to begin the trials with a keypress.

The procedure is illustrated in Figure 1. At the start of each trial, an empty  $4 \times 4$  square grid was presented in the center of the display. When ready to see the first array, the subjects pressed the space bar. Immediately, the first array of seven dots was presented within the grid for 33 msec. The ISI between the offset of the first dot array and the onset of the second array was 0, 100, 500, 1,000, 1,500, 2,000, 2,500, or 3,000 msec, during which the empty grid remained on the screen. Following the delay, the second array, containing eight dots, was presented for 33 msec. After both arrays had been presented, the subjects used a number keypad to indicate the row and column coordinates corresponding to the grid position they thought was not filled on that trial. The subjects were told to respond as accurately as possible and that they were under no speed stress.

Each ISI (eight) occurred equally often. With a repetition factor of 32, each subject completed a total of 256 trials, divided into blocks of 8 trials. The order of blocks was randomized. Prior to beginning the experimental trials, the subjects completed 32 practice trials, which consisted of 8 trials each at ISIs of 0, 100, 750, and 1,500 msec. During practice, the subjects were given feedback concerning the accuracy of their responses. However, during the experimental trials, no feedback was given.

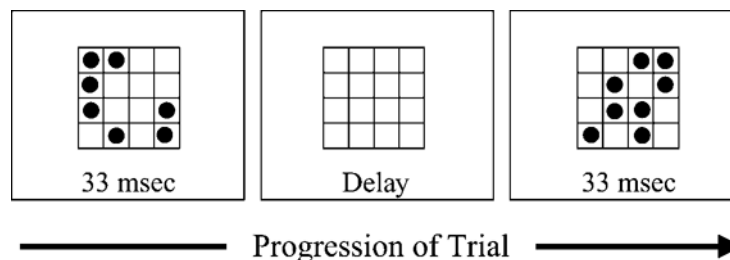


Figure 1. Schematic illustration of the procedure in Experiment 1. The subjects were instructed to locate the one grid space that was not filled with a dot.

The eyetracker monitored eye position during the presentation of Array 1, the ISI, and the presentation of Array 2. The subjects began each trial by looking at the center of the grid but were free to move their eyes without constraint during the trial. Eye position was scored in terms of the grid space that was fixated. Saccades were operationally defined as a change in fixation that exceeded  $0.2^\circ$  of visual angle accompanied by a velocity exceeding  $30^\circ/\text{sec}$  or an acceleration exceeding  $8,000^\circ/\text{sec}^2$  that was maintained for a minimum of 4 msec. The saccade was considered to be terminated when these criteria were no longer met.

### Results

The results will be reported in two sections. First, general accuracy and error rates for the integration task were assessed to establish that the present experiment replicated Brockmole et al.'s (2002) result that, as the ISI was increased beyond 100 msec, accuracy improved. This replication would establish the appropriateness of interpreting the eye movement record in relationship to prior results. Second, various aspects of eye movement behavior were examined with the goal of linking particular patterns of fixation to performance. Variables of interest included number of fixations, location of fixations, and fixation durations.

#### Integration Task Accuracy and Error

A response was classified as correct, an Array 1 error (erroneously selecting a position occupied by the first array), or an Array 2 error (erroneously selecting a position occupied by the second array) and were measured in terms of the percentage of trials on which they occurred. The results are illustrated in Figure 2. A one-way repeated measures analysis of variance with ISI as the independent factor demonstrated that accuracy varied as a function of ISI [ $F(7,49) = 24.2$ ,  $MS_e = 95.5$ ]. Considering those ISIs greater than or equal to 100 msec, single degree of freedom polynomial tests indicated a reliable positive slope linear trend [ $F(1,7) = 20.6$ ,  $MS_e = 233$ ] and a quadratic trend [ $F(1,7) = 14.4$ ,  $MS_e = 114$ ] in the accuracy data. These trends indicate that accuracy increased with ISI in a curvilinear fashion. However, post hoc contrasts showed that accuracy never truly reached an asymptotic level, attaining a maximum of 61% in the data space measured. This level of performance is comparable to the asymptotic level observed by Brockmole et al. (2002). Array 1 errors occurred, on average, on 40% of the trials and accounted for 82% of all the errors. The Pearson correlation between Array 1 errors and correct responses across ISIs was  $-.97$ , revealing almost perfect dependence between accuracy and Array 1 error. Array 2 errors, however, as a percentage of total trials occurred, on average, on 9% of the trials and accounted for 18% of all the errors. The Pearson correlation between Array 2 errors and correct responses across ISIs was  $-.26$ . These overall patterns of results closely parallel those observed by Brockmole et al. (2002).

#### Eye Movement Behavior

The locus of overt attention is indicated by the location in space that is fixated. Thus, an analysis of patterns

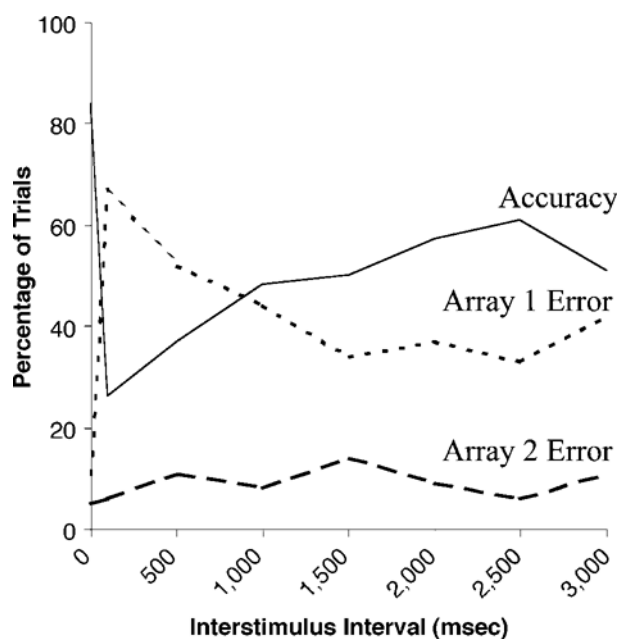


Figure 2. Accuracy and error results from Experiment 1.

in the distribution of fixations during the ISI separating the arrays can give insight into the manner in which overt attention is used during VSTM consolidation. In this section, fixations are analyzed according to their number, location, and duration as a function of ISI.

**Number.** If the locations of the dots in Array 1 are maintained by overtly attending to each of the locations that once housed an Array 1 dot, the number of fixations observed on a particular trial should be positively correlated with ISI, since longer ISIs afford more time to make more eye movements. Indeed, as Figure 3 illustrates, that is exactly the pattern that was observed. The Pearson correlation between ISI and number of fixations was  $.75$ . A linear regression analysis indicated that 1.5 additional fixations were made for every 1,000-msec increase in ISI [ $F(1,62) = 80.7$ ,  $MS_e = 1.90$ ].

**Location.** Three separate analyses were performed to consider the location of fixations within the grid in order to determine the locus of overt attention during integration. In these analyses, the absolute number of grid positions fixated, the distribution of fixations to Array 1 and Array 2 grid positions, and fluctuations in the distribution of fixations on the 16 grid positions over time were considered.

First, if the subjects attempted to fixate as many grid positions previously occupied by a dot in Array 1 as possible, given the available ISI, the number of grid locations that were fixated should also increase with increases in ISI. More specifically, if each eye movement carried fixation from one grid position to another, the rate of this increase should exactly match the rate at which the number of fixations increased as a function of ISI. The extent to which multiple fixations occur within a single grid position—whether caused by intraposition

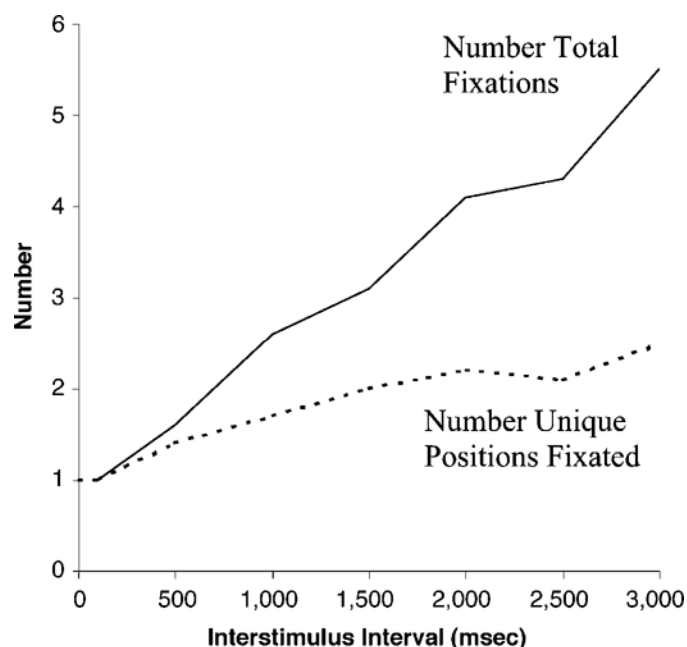


Figure 3. Number of total fixations versus the number of fixations made in unique grid locations in Experiment 1.

refixation or interposition revisitation—is indicated by the discrepancy between the rates at which the number of fixations and the number of grid positions fixated increase with increases in ISI. Figure 3 illustrates that although the number of grid positions fixated did increase with increases in ISI, the slope of this increasing function was 0.5 items for every 1,000-msec increase in ISI [ $F(1,62) = 51.6$ ,  $MS_e = 0.314$ ]. This rate is one third that of the increase in the number of fixations observed, indicating that one out of every three fixations was on a unique grid position. This disparity is striking. For example, assuming an average fixation duration of 300–500 msec (Rayner, 1998), between 7 and 11 grid positions could have been fixated at the longest ISI (3,000 msec), enough to fixate the location of each dot individually. However, on average, only 2.5 grid positions were actually fixated.

Second, because the number of grid positions fixated did increase with increases in ISI, another analysis was performed to examine where those fixations occurred. The results of Brockmole et al. (2003) showed that attention is deployed to Array 1 positions. If the unique positions that were fixated in the present experiment were made to overtly guide attention systematically to these positions, a plurality of fixations should rest on grid locations that were occupied by a dot from the first array. Because Array 1 contained seven dots, the probability of fixating an Array 1 position randomly was 44%. Figure 4 illustrates the percentage of fixations on a grid position that was occupied by an Array 1 dot as a function of ISI.

Although a reliable but nonsystematic effect of ISI was observed [ $F(5,35) = 4.11$ ,  $MS_e = 43.2$ ], 95% con-

fidence intervals constructed around each data point (illustrated by the error bars in Figure 4) revealed that the percentage of fixations on Array 1 positions (the black line in Figure 4) did not reliably differ from chance (the dashed line in Figure 4) at any ISI, with the exception of 1,500 msec. A subsequent analysis also demonstrated that the percentage of fixations on Array 1 positions did not vary as a function of fixation number. In sum, the

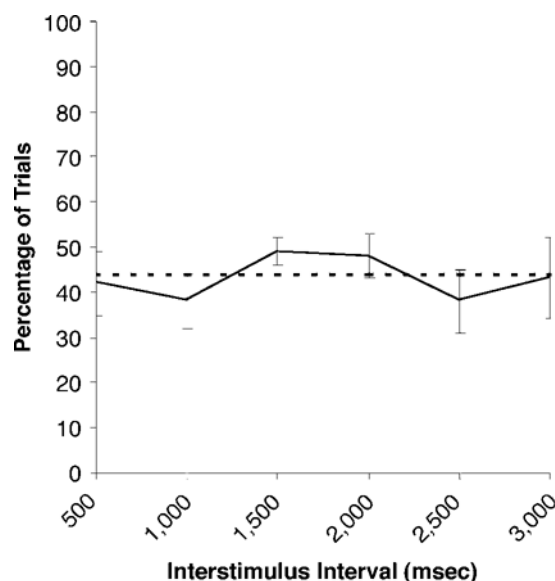


Figure 4. Percentage of fixations on positions originally occupied by a dot from Array 1 in Experiment 1.

percentage of fixations on Array 1 positions was at chance and did not vary by ISI or fixation number.

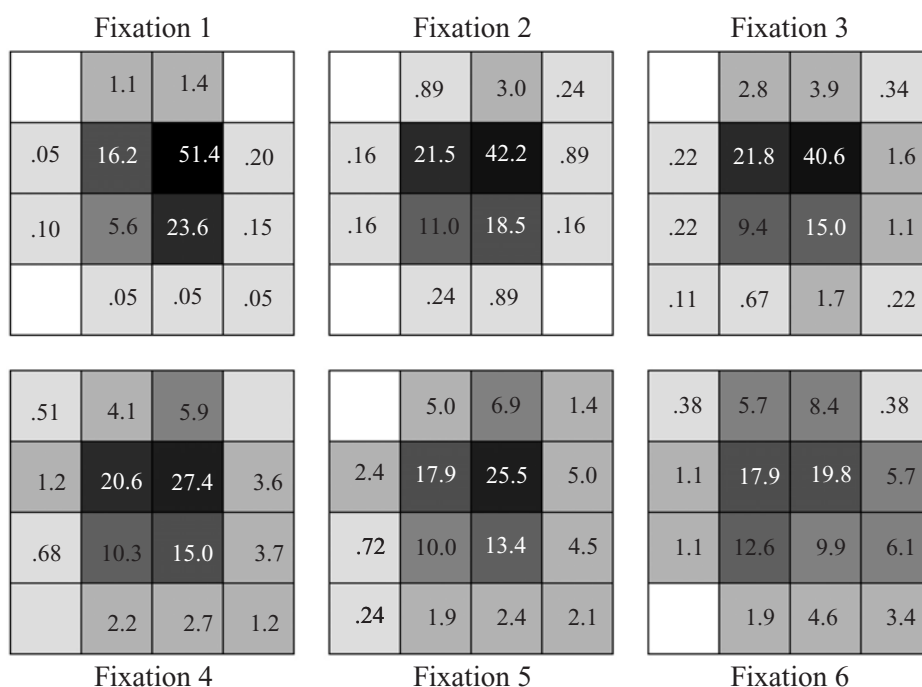
Third, the location of fixations was analyzed according to the frequency with which each grid position was fixated conditionalized on fixation number. The subjects were instructed to fixate the center of the grid at the start of the trial but were permitted to move their eyes freely once the trial began. As a result, the initial fixation in the grid would be in one of the four center boxes. If the subjects scanned the grid, each subsequent fixation should be more likely to occupy a position along the periphery of the grid. The percentage of total fixations across subjects that coincided with each grid position is illustrated in Figure 5 by fixation number. Note that relatively few fixations fell on peripheral grid locations, even after six fixations on the grid.

In an effort to characterize the preferences that the subjects displayed for fixating various grid positions, the percentage of fixations that occurred in the center four grid positions, where the first fixation occurred, was calculated for each fixation. The solid line in Figure 6 illustrates the discrete percentage of trials on which the center four grid positions were fixated as a function of fixation number. The dashed line in Figure 6 illustrates the cumulative percentage of fixations that occurred in the center four grid positions. This figure depicts a strong bias for the subjects to maintain fixation in the center region of the grid. Through the first six fixations, 84% of all the fixations occurred in the center four grid positions. Considering the sixth fixation alone, this fixation fell in the center four grid positions on 60% of the trials.

**Duration.** Finally, fixation durations were analyzed as a function of fixation number. Fixation times have been considered to reflect the amount of processing that must be completed at a particular location in a visual scene, on the basis of the assumption that a fixation does not end until processing has been completed (Henderson et al., 1999; Just & Carpenter, 1980; Rayner, 1998; but see Irwin, 2004, for an alternative view). If each fixation a subject makes during VSTM consolidation and maintenance is used equally to process the display—in this case, to allocate attention to particular components of the grid—fixation times should not vary as a function of fixation number. That is, the information processing at Fixation 1 should be at least equal to the processing at Fixation 2, and so on. Mean fixation durations are illustrated in Figure 7 by fixation number (1–6) and ISI. An overall main effect of ISI was observed, since fixation durations generally increased with increases in ISI. At longer ISIs, the subjects may have felt less “hurried” to make their eye movements, thus causing this general slowdown. More important, a main effect of saccade number was observed. In general, fixation durations were shorter for later saccades. This suggests that progressively less processing occurred during later fixations, indicating less processing at each subsequent fixation.

**Discussion**

The fact that only one third of all the fixations occurred in unique grid positions, that on average, only 2.5 grid positions were fixated even when the ISI was 3,000 msec, that no bias for fixating Array 1 positions



**Figure 5.** Percentage of all fixations within each grid space by fixation number (through 6) in Experiment 1. Darker shades indicate greater percentages of fixations.

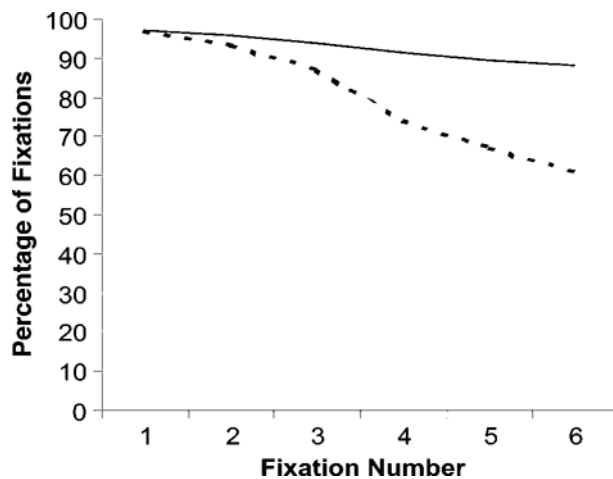


Figure 6. The percentage of trials on which the first six fixations were on the center four positions of the grid (dotted line), as well as the cumulative number of fixations on those positions across fixation number (solid line) in Experiment 1.

existed, that the subjects displayed little tendency to fixate grid positions along the periphery of the grid, and that successive fixations in a trial exhibited shorter latencies strongly suggests that the subjects did not systematically scan the grid during the ISI separating the arrays. Instead, it seems that quite the opposite was true: The subjects tried to move their eyes as little as possible. Thus, it appears that the overt orienting of attention is not required for VSTM consolidation and rehearsal and that the minimum amount of temporal separation between the to-be-integrated arrays that allows for optimal integration performance is not determined by a minimum number of saccades required to drive overt atten-

tion to all of the elements of Array 1. Rather, the ISI separating the arrays appears to be used to assign covert attention to the grid locations maintained in memory.

The fact that memory-percept integration can occur without the deployment of overt attention does not, however, necessarily indicate that the integration system is completely independent of eye movements. Consider two extreme positions that illustrate very different possible interactions between integration and eye movements. On the one hand, it may indeed be that integration is completely *independent* of eye movements. Because the eyes do not need to fixate to-be-remembered information, it is therefore possible that they may be free to scan aspects of the visual scene that are not ultimately included in the maintained memory representation. This would add some flexibility to the integration system, since various aspects of a display could be attended to without being integrated. On the other hand, integration may be completely *dependent* on eye movements—more precisely, on a *lack* of eye movements. Perhaps the eyes do not move much during the VSTM consolidation period because eye movements would force attention to be shifted around the visual display, perhaps overriding covert attention shifts and thereby disrupting the integration process. Another possibility is that overt eye movements to locations in the visual field draw attention away from memory. In other words, integration may be dependent on the eyes' not moving, since movements may cause interference. Indeed, eye movements have been shown to disrupt the retention of sequentially presented visual stimuli, as compared with stimulus-driven covert shifts of attention of equal spatial magnitude (Pearson & Sahraie, 2003). Tacit knowledge of this disruption might lead subjects to keep their eyes still when viewing the sequential arrays presented in the temporal integration paradigm. Experiment 2 assessed integration when

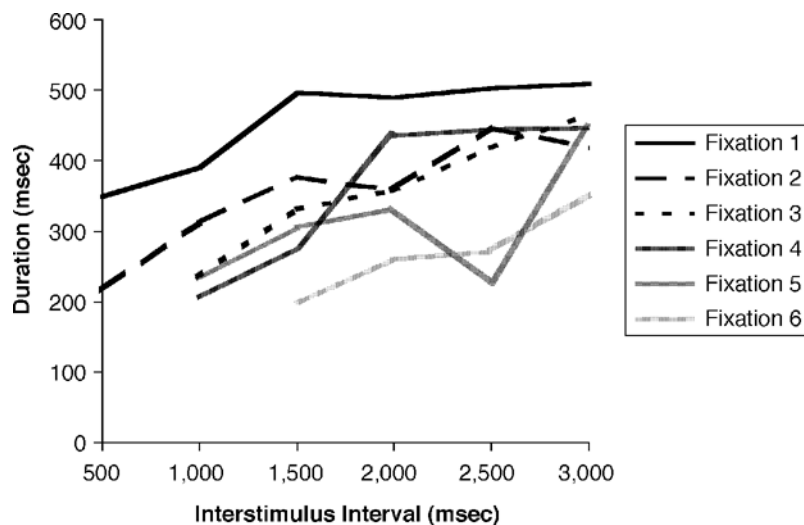


Figure 7. Durations of the first six fixations as a function of interstimulus interval in Experiment 1.



subjects were required to move their eyes to visually acquire the information necessary for integration, a manipulation that distinguished between these possibilities.

## EXPERIMENT 2

Experiment 2 tested memory-percept integration in cases in which the to-be-remembered stimulus and the to-be-integrated stimulus were not presented in the same spatial location; rather, the first and the second arrays were separated by 20° of visual angle. Although eye movements played little or no role during integration when the first and the second arrays were presented in the same grid in Experiment 1 (where no eye movements were required to view both arrays), in Experiment 2 eye movements were critical to integration, insofar as a saccade was required to shift fixation from one grid to the other. That is, Experiment 2 required a lengthy (in terms of spatial displacement) shift of overt attention. This requirement may alter the manner in which eye movements interact with VSTM consolidation and maintenance. The critical question in Experiment 2, then, is the following: Does the need to make eye movements to obtain the relevant information for integration influence the role eye movements play in the generation and maintenance of the memory representation of Array 1?

Several outcomes are possible. First, the overt shift of attention may not alter the eye movement patterns (or lack thereof) in Experiment 1 if covert attention can reassemble the first array after the eyes land on the second array. Second, if covert attention cannot reassemble the initial stimulus in the new spatial location, evidence for an overt selection process may result. Third, the overt shift of attention from one grid to the other may interfere with integration generally if eye movements are detrimental to the maintenance of VSTM representations in this task. In previous integration studies in which the to-be-integrated arrays did not spatially align, the misalignment was created by either scaling the arrays or rotating them. Both arrays, however, were anchored to the same point in space and were presented so that both could be viewed in a single fixation (Brockmole & Wang, 2003). Indeed, research on transsaccadic memory in the 1980s indicated that very little visual information is carried across a saccade and that integration of pattern information is not possible when a saccade intervenes between stimulus presentations (Bridgeman & Mayer, 1983; Henderson, 1997; Irwin, Brown, & Sun, 1988; Irwin, Yantis, & Jonides, 1983; McConkie, 1991; McConkie & Rayner, 1976; O'Regan & Levy-Schoen, 1983; Rayner & Pollatsek, 1983). However, these studies assessed integration abilities only at very short ISIs (a few hundred of milliseconds), and on the basis of our current knowledge of memory-percept integration, much more time may be needed. As such, a secondary question addressed in Experiment 2 is whether or not integration is even possible if the arrays are presented in separate fixations.

## Method

### Subjects

Twenty members of the University of Illinois community participated after providing informed consent. All the subjects were naive with respect to the experimental hypotheses and were paid \$8 for participating.

### Stimuli

The stimuli were the same as those in Experiment 1, save that the first array was presented approximately 9° of visual angle to the left of the center of the display and the second array was presented 9° of visual angle to the right of the center of the display. No portion of the left and right grids overlapped in space; the rightmost boundary of the left grid and the leftmost boundary of the right grid were separated by 5°, the equivalent of one grid space.

### Apparatus

The same apparatus as that in Experiment 1 was used.

### Design and Procedure

Each trial consisted of two sequentially presented dot arrays within enclosed square grids separated by a variable ISI. During the ISI separating the arrays, the grid changed location. The first and the second arrays together filled all but one position in the grid. Under the constraint that no members of the two arrays could occupy the same position in the grid, the dot patterns were randomly generated on each trial. Thus, the location of the missing dot was equally likely to occur in each grid position. The subjects were instructed to identify the position of the missing dot.

There are two time intervals that may be critical to memory-percept integration. First, the time that elapses between the offset of the first array and the displacement of the grid may be important for extracting information from the initial array required to generate an accurate memory representation. If the grid changes location too quickly, so that the first array cannot be adequately encoded, little information may be retained from which to generate a memory trace. Second, the time that elapses between the displacement of the grid and the onset of the second array may be important for processing the information committed to memory in the new spatial location. Each of these time components, which will be referred to as *early components* and *late components*, respectively, was independently manipulated in Experiment 2 to determine how much time was necessary before and after grid displacement to produce optimal performance. In addition, this manipulation allowed the relationship between eye movements and memory consolidation both prior to and following the displacement of the grid to be examined. That is, depending on the duration of the early and late components, the VSTM representation had to be formed before the displacement of the grid (e.g., long early component, short late component) or following the displacement of the grid (short early component, long late component).

Two conditions were created that contrasted the contributions of the time available for processing prior to grid displacement and the time afforded for processing following grid displacement (see Figure 8). In the *early component condition*, performance was examined as a function of the time available prior to grid displacement. This was accomplished by varying the time that elapsed between the offset of the first array and the presentation of the second, spatially displaced grid. This will be referred to as the *predisplacement delay*. In order to examine performance as a function of the predisplacement delay, this condition severely limited the amount of time available for postdisplacement processing by holding it constant at 100 msec. That is, after the second grid had been presented, it was filled with a dot array 100 msec later. This time allowed the subjects to recover from any saccadic suppression, which can last up to 100 msec postsaccade (Volkman, Schick, & Riggs, 1968), associ-

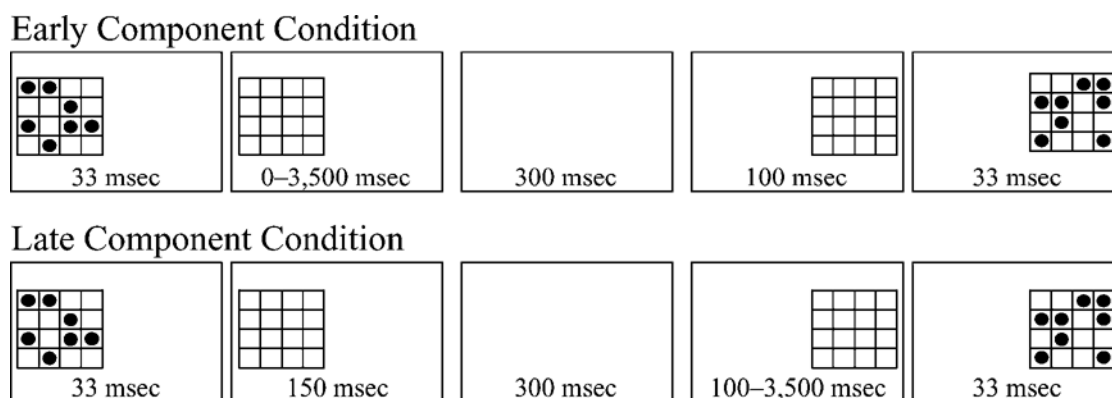


Figure 8. Schematic illustration of trial events for the early and late component conditions in Experiment 2.

ated with moving the eyes from the left to the right side of the display and to acquire a new fixation. Unless postdisplacement processes require less than 100 msec, this time course does not, however, allow such processes to be completed prior to the onset of the second array. As a result, performance in this condition should largely reflect the time course of predisplacement processing and should then reveal how well the arrays can be integrated, given only those processes at work.

In the *late component condition*, on the other hand, performance was examined as a function of the time available for processing after grid displacement. This was accomplished by varying the time that elapsed between the displacement of the grid and the presentation of the second array. This will be referred to as the *postdisplacement delay*. As opposed to the early component condition, however, it was essential to allow early processing to be optimally completed prior to grid displacement. If predisplacement processing is not complete, either it will be continued postdisplacement or it will not be completed at all, which may reduce general performance. Thus, if predisplacement processing is complete, trends in performance as a function of postdisplacement time should largely reflect the contribution of postdisplacement processing on memory–percept integration.

The early and late component conditions constituted a between-subjects factor. The early component condition was completed prior to the late component condition. In both conditions, prior to the onset of each trial, an empty  $4 \times 4$  square grid was presented on the left side of the display. When ready to begin the trial, the subjects pressed the space bar. Immediately, the first array of seven dots was presented within the grid for 33 msec. The events following the offset of the dot array depended on the condition (early component vs. late component) of the trial.

For the subjects in the early component condition, following Array 1 offset, the grid remained on the left side of the display (the predisplacement delay) for 0, 33, 67, 100, 300, 500, 700, 1,500, or 3,500 msec, whereupon it was erased. Following grid erasure, the display remained blank for 300 msec. The (empty) second grid then appeared on the right side of the display for 100 msec. The second dot array was then presented for 33 msec. Thus, the ISI separating the arrays was 400, 433, 467, 500, 700, 900, 1,100, 1,900, or 3,900 msec. For the subjects in the late component condition, following Array 1 offset, the grid remained present on the left side of the display for 150 msec, whereupon it was erased (the determination of this time will be discussed in detail in the Results section). This was followed by a 300-msec interval, during which the display was blank. The second grid then appeared on the right side of the display, followed by a variable delay of 100, 200, 300, 500, 750,

1,000, 1,500, 2,000, or 3,500 msec (the postdisplacement delay). Finally, the second dot array was presented for 33 msec. Thus, the ISI separating the arrays was 550, 650, 750, 950, 1,200, 1,450, 1,950, 2,450, or 3,950 msec. In both conditions, the 300-msec blank interval was inserted between the offset of the first grid and the onset of the second grid in order to avoid any potential backward, lateral, or object substitution masking effects, which are known to dissipate within 300 msec (Averbach & Coriell, 1961; Breitmeyer, 1980, 1984; Breitmeyer & Ganz, 1976; Enns & Di Lollo, 1997; Kahneman, 1968; Lefton, 1973; Matin, 1975; Turvey, 1973).

The trials were divided into blocks on the basis of the pre- or postdisplacement delay. The subjects were informed of the delay prior to the start of each block and were told to move their eyes away from the first array as soon as it disappeared. Thus, the disappearance of the first grid provided a cue for the subjects to shift their gaze across the display. The subjects completed 252 trials, divided into blocks of 7. Each of the nine delays occurred equally often. For each subject, the blocks of trials were presented in a different random order. Prior to beginning the experimental trials, the subjects completed two practice sessions to familiarize themselves with the experimental procedure. In the first session, the subjects completed 32 trials in which both the first and the second grids were presented in the same spatial location. This session consisted of 8 trials each at ISIs of 0, 100, 750, and 1,500 msec. In the second session, the subjects completed 28 trials identical to the experimental trials in the condition in which they participated. For the subjects in the early component condition, this session consisted of 7 trials at predisplacement delays of 33, 300, 700, and 1,000 msec each. For the subjects in the late component condition, this session consisted of 7 trials at postdisplacement delays of 33, 300, 700, and 1,000 msec each. During the practice sessions, the subjects were given feedback concerning the accuracy of their responses. However, during the experimental trials, no feedback was given.

## Results

The results will be reported in several sections. First, the behavioral results of the early component condition will be discussed; then the behavioral results of the late component condition will be discussed. The focus of these analyses was on trends in accuracy, error, and eye movement behavior as a function of the delay associated with each condition. That is, performance in the early component condition was assessed in terms of the predisplacement delay, and performance in the late compo-

ment condition was assessed in terms of the postdisplacement delay. The accuracy and error data for each condition are illustrated in Figures 9 and 10. Second, eye movement behavior in both the early and the late component conditions was assessed as it was in Experiment 1, with a focus on fixation number, location, and duration. These data are illustrated in Figures 11 and 12.

### Data Trim

Prior to data analysis, the trials were required to meet several selection criteria based on eye movement behavior. First, the eyes were required to be fixated on the left grid for a total duration at least equal to the duration of the predisplacement delay for that trial. Second, the eyes were required to have acquired fixation on the right grid prior to the onset of the second array. These timing requirements ensured that the subjects executed a saccade across the display from the left grid to the right grid during the 300-msec blank interval separating grid presentations. Therefore, the subjects viewed each grid for a period at least equal to that of the pre- or postdisplacement delay for the trial, and the subjects were looking at both the first and the second arrays when they were presented. In essence, this trim ensured that the subjects indeed moved their eyes away from the first grid as soon as it disappeared. In the early component condition, this trim excluded 27% of the trials. The majority of the trials excluded were those for which the delay was 100 msec or less, where approximately 40% of the trials were excluded. The exclusion rate fell to approximately 20% of the trials when the delay was greater than 100 msec. In the late component condition, 25% of the trials were excluded. The trim did not differentially exclude trials at each delay. This trim did result, however, in empty cells

for 1 subject, and that subject's data are not included in the analyses.

### Integration Performance

**Early component condition.** A one-way repeated measures analysis of variance with predisplacement delay as the independent factor demonstrated that accuracy varied as a function of this delay [ $F(8,72) = 4.20$ ,  $MS_e = 73.2$ ]. Single degree of freedom polynomial tests indicated a reliable positive slope linear trend [ $F(1,9) = 10.5$ ,  $MS_e = 114.7$ ] in the accuracy data. Planned comparisons showed that accuracy improved from 30% to 41% as the predisplacement delay increased through 100 msec (500-msec ISI). Surprisingly, accuracy then decreased through 500 msec, falling to 34% before recovering to 42% by 1,500 msec. Statistically, performance at 100 msec was equal to performance when the predisplacement delay was 1,500 msec or greater. The temporary decrease in accuracy when the predisplacement delay was approximately 500 msec was examined in depth in a separate experiment and was found to be theoretically uninteresting, since it was the result of non-systematic variation or Type II statistical error.<sup>1</sup> Therefore, for purposes of the following analyses, it is simply observed that the earliest point at which performance reached a maximum level coincided with a predisplacement delay of approximately 100 msec.

In addition, the source of errors was assessed, with an emphasis on the strength of relationship between different types of error and accuracy. As a percentage of total trials, Array 1 errors occurred, on average, on 49% of the trials and accounted for 74% of all the errors. The Pearson correlation between Array 1 errors and correct responses across ISIs was  $-.84$ , revealing a very high level

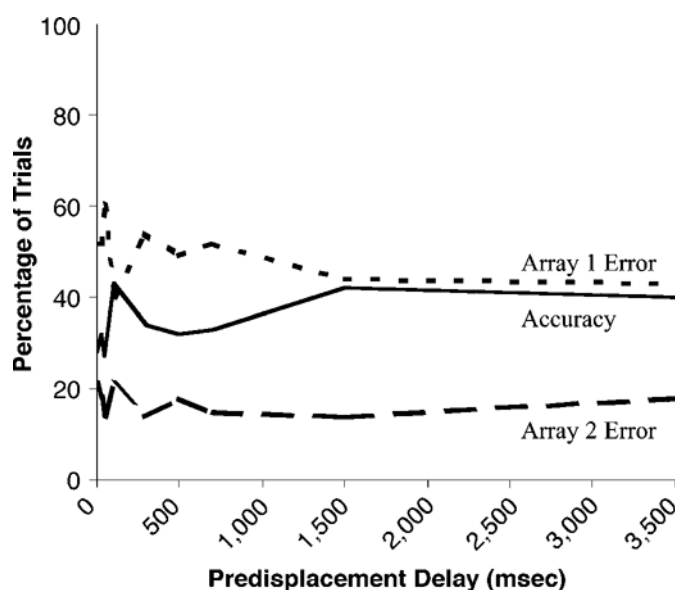


Figure 9. Accuracy and error results from the early component condition in Experiment 2.

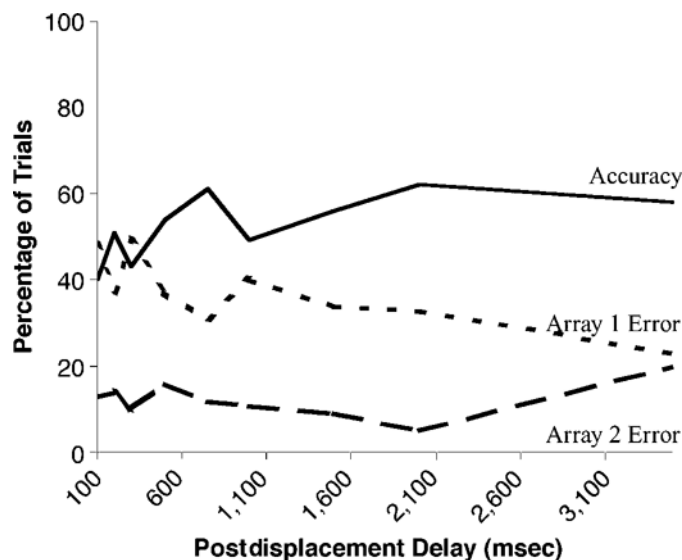


Figure 10. Accuracy and error results from the late component condition in Experiment 2.

of dependence between accuracy and Array 1 error. This means that the improvement in accuracy was heavily due to the formation of a more complete memory representation of the first array. Array 2 errors, however, as a percentage of total trials occurred, on average, on 17% of the trials and accounted for 26% of all the errors. The Pearson correlation between Array 2 errors and correct responses across ISIs was  $-.32$ , revealing relative independence between accuracy and Array 2 error.

**Late component condition.** The aim of the late component condition was to measure performance as a function of time available for postdisplacement processing, given that the predisplacement delay was long enough for all predisplacement processing to be completed. To keep overall ISI durations reasonable, ideally, the time between the offset of the first array and the removal of the grid in the late component condition should be the shortest duration possible that still allows all predisplacement processes to be completed. On the basis of the results of the early component condition, this duration is 100 msec (the time after which performance in the early component condition reached maximum). However, because different subjects completed the late component condition and because there was certainly individual variance around this 100-msec estimate, the predisplacement time in the late component condition was increased to 150 msec, to be more certain that enough time was provided for predisplacement processing.

A one-way repeated measures analysis of variance with postdisplacement delay as the independent factor demonstrated that accuracy varied as a function of this delay [ $F(8,64) = 2.52$ ,  $MS_e = 208.0$ ]. Single degree of freedom polynomial tests indicated a reliable positive slope linear trend [ $F(1,8) = 4.26$ ,  $MS_e = 575.0$ ] in the accuracy data. The quadratic trend was not reliable. How-

ever, planned comparisons showed that accuracy improved from 40% to 60% as the postdisplacement delay increased through 750 msec (1,200-msec ISI) but did not differ thereafter, averaging 57%. A linear spline regression was performed on these data to interpolate the point at which accuracy reached a ceiling level between the selected delay values (see Brockmole et al., 2002). Accuracy improved until the postdisplacement delay was 650 msec (1,100-msec ISI).

An error analysis showed that, as a percentage of trials, Array 1 errors occurred, on average, on 37% of the trials and accounted for 76% of all the errors. The Pearson correlation between Array 1 errors and correct responses across ISIs was  $-.80$ , revealing a high level of dependence between accuracy and Array 1 error. This means that the improvement in accuracy was heavily due to the formation of a more complete representation of the first array. Array 2 errors, however, as a percentage of trials occurred, on average, on 12% of the trials and accounted for 24% of all the errors. The Pearson correlation between Array 2 errors and correct responses across ISIs was  $-.34$ , revealing relative independence between accuracy and Array 2 error.

**Early component versus late component.** Performance in the early component condition was compared with that in the late component condition by collapsing across pre- and postdisplacement delays that exceeded the point at which accuracy no longer varied and performing a two-sample  $t$  test. Asymptotic accuracy in the late component condition was reliably higher than that in the early component condition [ $t(18) = 2.25$ ]. The magnitude of this advantage was, on average, 20%. During this same time frame, Array 1 errors were reliably lower in the late component condition [ $t(18) = 2.87$ ], as were Array 2 errors [ $t(18) = 3.55$ ]. The magnitude of

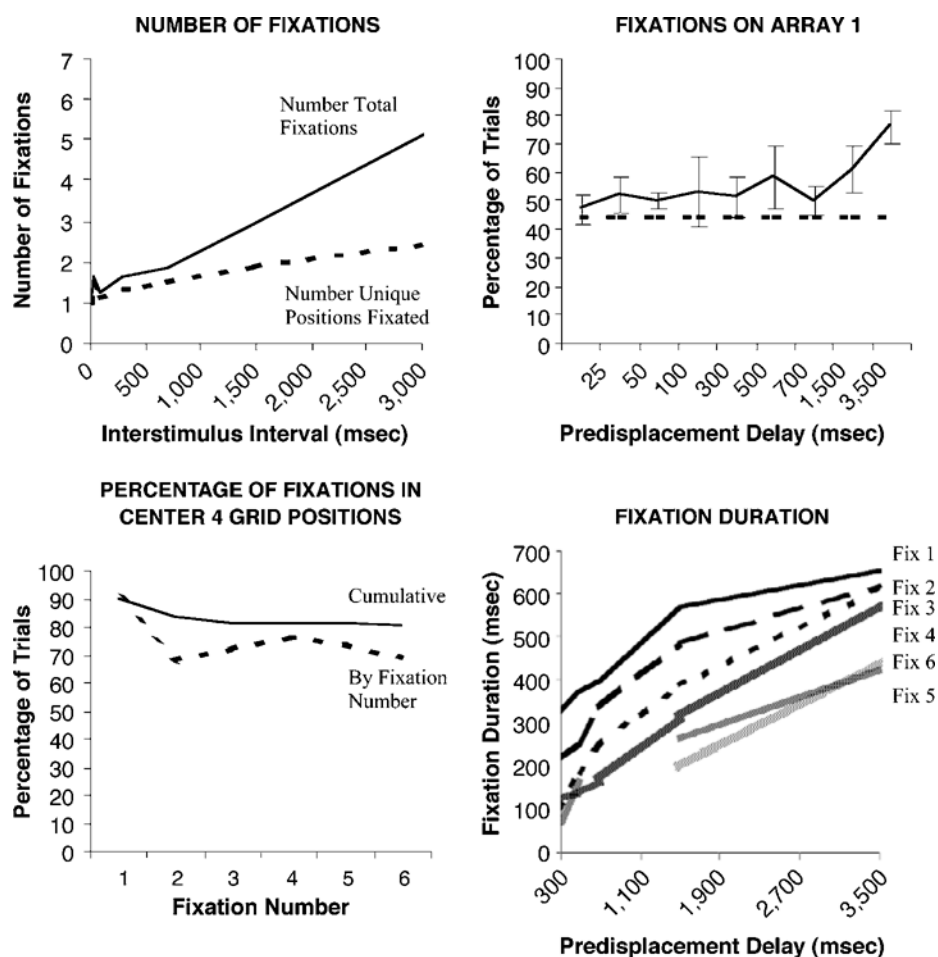


Figure 11. Fixation number, location, and duration in the early component condition in Experiment 2.

these differences was, on average, 15% and 5%, respectively, indicating that the increase in accuracy between the early and the late component conditions was due to a better representation of the first array.

### Eye Movement Behavior

The same analyses of eye movement behavior (fixation number, location, and duration) were conducted as those in Experiment 1 for the pre- and postdisplacement delays (early and late component conditions, respectively).

**Number.** For both the early and the late component conditions, the number of fixations observed on a particular trial increased with ISI. Linear regression analyses indicated that in the early component condition, 1.3 additional fixations were made for every 1,000-msec increase in ISI [ $F(1,88) = 377$ ,  $MS_e = 0.444$ ]. In the late component condition, 1.2 additional fixations were made for every 1,000-msec increase in ISI [ $F(1,79) = 123$ ,  $MS_e = 0.857$ ].

**Location.** The absolute number of grid positions fixated increased with ISI in both the early and the late component conditions. In the early component condi-

tion, the slope of this increasing function was 0.4 items for every 1,000-msec increase in ISI [ $F(1,88) = 141$ ,  $MS_e = 0.135$ ]. In the late component condition, the slope was also 0.4 items for every 1,000-msec increase in ISI [ $F(1,79) = 47.9$ ,  $MS_e = 0.224$ ]. These rates are about one third those of the increase in the number of fixations observed, indicating that approximately one out of every three fixations was on a unique grid position. In both the early and the late component conditions, the rate at which fixations corresponded to locations previously occupied by a dot from Array 1 was greater than chance. Across predisplacement delay, Array 1 locations were fixated at an average rate of 55%. Across postdisplacement delay, Array 1 locations were fixated at an average rate of 60%. Furthermore, these rates varied linearly as a function of the corresponding pre- or postdisplacement delay. In the case of the early component condition, the probability of fixating an Array 1 location increased from 47% when the predisplacement delay was 0 msec to 76% when it was 3,500 msec. Nearly identical rates of increase were observed in the late component condition, in which the probability of fixating an Array 1 location increased

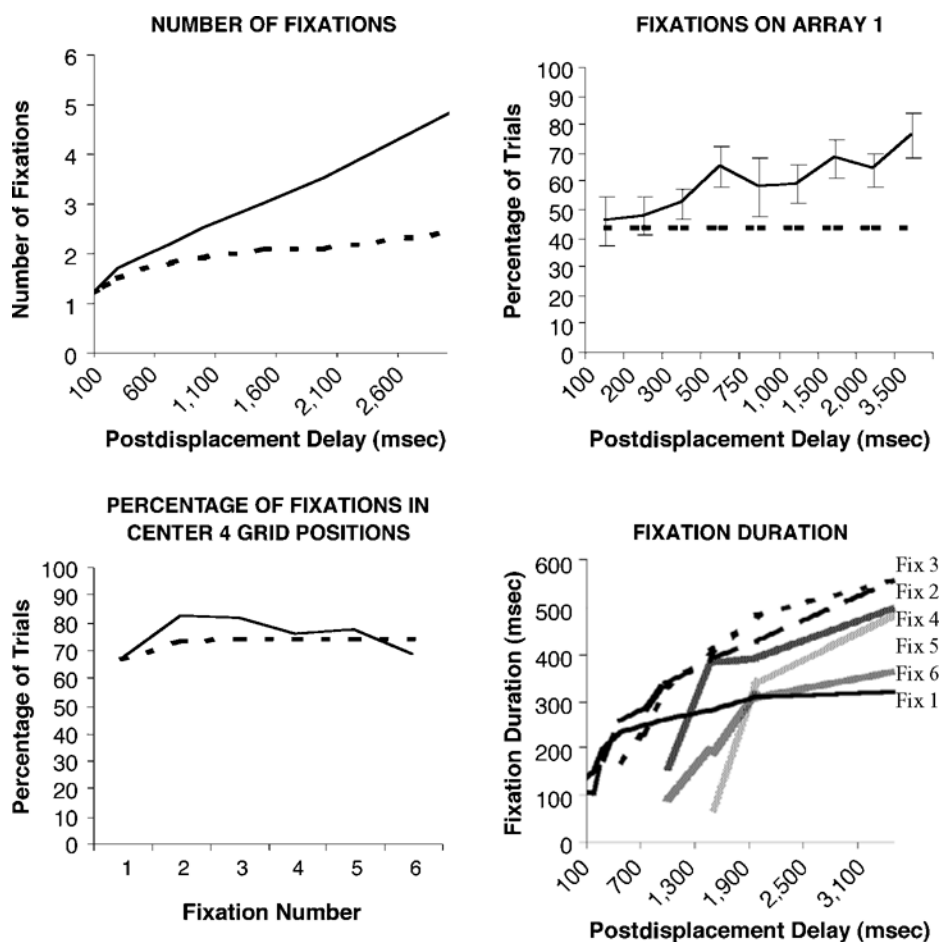


Figure 12. Fixation number, location, and duration in the late component condition in Experiment 2.

from 46% when the postdisplacement delay was 100 msec to 75% when it was 3,500 msec. However, in both conditions, the subjects demonstrated a preference for localizing the majority of their fixations in the center four grid positions, but to a lesser extent than in Experiment 1. Through the first six fixations, in the early component condition, 84% of all the fixations occurred in the center four grid positions; in the late component condition, this value was 74%.

**Duration.** In both the early and the late component conditions, an overall effect of delay was observed, since fixation durations generally increased with ISI; however, only in the early component condition was an effect of saccade number observed where fixation durations were shorter for later saccades. In the late component condition, fixation number had no systematic impact on fixation duration.

### Discussion

The data from Experiment 2 suggest that memory-percept integration is possible when the to-be-remembered

stimulus is not presented in the same spatial location as the subsequent to-be-integrated stimulus. As in the other experiments, timing was found to be critical. The ISI separating the arrays, although an essential factor in allowing memory-percept integration to occur, however, was not the only temporal determinant of performance. The point during the ISI at which the grid changed location was also a major factor in determining performance. Optimal performance was achieved when the time prior to grid displacement was approximately 150 msec and when the time after displacement was approximately 650 msec. If the pre- or the postdisplacement time was less than these values, integration was suboptimal. As a final note, considering the additional 300-msec blank interval, these values produced an ISI of 1,100 msec, which is very similar to that observed by Brockmole et al. (2002) in their original studies, in which the grids did not change position.

The eye movement results generally paralleled those obtained in Experiment 1, but several important differences emerged. In both experiments, ISI had a similar ef-

fect on the number of fixations made, and approximately one third of those fixations occurred in new grid positions. This indicates that the requirement that visual information be acquired in separate fixations did not significantly alter the total number of fixations made or the number of grid positions visited by the eyes. However, Experiment 2 revealed a systematicity in the distribution of fixations in the grid that was not evident in Experiment 1. In both the early and the late component conditions, the subjects were more likely to fixate distal regions of the grid, as well as positions previously occupied by a dot from Array 1, especially at longer ISIs. Notable differences between the early and the late component conditions were evident, however. Fixation times decreased as the predisplacement delay increased but were unaffected by the postdisplacement delay, suggesting that when consolidation was confined largely to a spatial position distant from the site of the initial stimulus, each fixation was used equally to process the display—in this case, to allocate attention to particular components of the grid. That is, the information processing at Fixation 1 was equal to the processing at Fixation 2, and so on. The decreasing fixation durations in Experiment 1 and in the early component condition in Experiment 2, combined with the absence of fixation duration effects in the late component condition in Experiment 2, suggest that eye movements are instrumental in forming and maintaining the structure of the to-be-remembered array only when the memory trace must be reconstructed in a unique position in space. When spatial memory does not have to be moved, or if it can be constructed prior to such movement, eye movements play little to no role in its construction. Rather, eye movements are used to reconstruct VSTM in a new space.

Why might the patterns of fixations have differed between Experiments 1 and 2? One possibility is that in Experiment 1, eye movements were not necessary to acquire information, and so they were not strategically used to do so. This would result not only in a reluctance to move the eyes, but also in randomness in the fixations that were made. In Experiment 2, however, the eye movement system had to be engaged strategically, and once it had, its effects extended beyond moving the eyes between stimuli and into moving the eyes within each stimulus. These subsequent eye movements were not made to acquire new perceptual information, since no information about the arrays was visually available, but more likely reflect a visual rehearsal process: The eyes retraced the pattern formed by the lead array in order to reconstitute the initial stimulus array in a new spatially appropriate memory representation.

So, what kind of relationship is shared between the integration and the eye movement systems? Earlier in Experiment 1, two possibilities were laid out: Integration could be either independent of eye movements or dependent on them, in the sense that eye movements interfered with integration. Experiment 2 clearly indicates that the latter was not the case. Integration performance

was quite high when eye movements were required for integration to be successful. Instead, the evidence points to at least partial independence between integration and eye movements. However, the magnitude of this independence is not clear. When to-be-integrated information was presented in distinct spatial locations, integration relied on eye movements in order for the observer to perceive all the information. Whether eye movements are necessary to integrate information that is presented in different spatial locations is not known, however. For example, if Array 1 and Array 2 were to be presented in different locations, both of which could be perceived without moving the eyes, could integration carry on normally? This question lies outside the scope of the present research, but it is one that is important for developing a complete understanding of the role that eye movements may play in integration.

Finally, the result that integration performance was associated with systematic patterns in gaze control is inconsistent with previous demonstrations that eye movements disrupt memory span for spatial locations. For example, Pearson and Sahraie (2003) found that subjects' memory span for spatial locations was lower if they executed saccades around a visual display, as compared with cases in which they shifted covert attention to the same locations. Their paradigm, however, forced observers to move their eyes to preestablished locations around the periphery of the display. As such, eye movements could not be used strategically to aid in the memory task. The present experimental paradigm, however, placed no inherent restrictions on an observer's gaze. As a result, eye movements could have been used strategically in order to rehearse the structure of the first array. This suggests that eye movements may help and hinder memory performance, depending on whether gaze is under the control of the observer.

## GENERAL DISCUSSION

The visual world contains more information than can be perceived and understood in a single glance, and the visual system is confronted by the physical limitations of the eye, as well as by the cognitive limitations of attention and memory. In order to overcome these limitations during scene processing, information about the world is obtained over time. Observers move their eyes around a scene to accommodate the resolution problems associated with peripheral vision and shift attention around the scene to sequentially process various objects. As a result, what has been seen needs to be analyzed in conjunction with what is being seen.

Because visual perception has temporal extent, temporally discontinuous input must be linked in memory. Recent research has suggested that this may be accomplished by integrating the content of active VSTM with subsequently perceived information (Brockmole et al., 2003; Brockmole & Wang, 2003; Brockmole et al., 2002). The purpose of the present article was to explore the re-

relationship between VSTM consolidation and eye movements in order to discover how attention selects the information that is to be integrated. Specifically, we addressed the question of whether stimuli had to be overtly attended to be included in the VSTM representation or whether covert attention was sufficient.

Previously, Brockmole et al. (2003) showed that during the delay separating the presentation of the arrays, attention is deployed to those grid locations that were previously occupied by a dot from the first array. But because eye movements were not monitored, whether this deployment was overt or covert could not be determined. Experiment 1 indicated that this deployment does not rely on overt shifts of attention. That is, integration seemingly proceeded independently of the eyes, since the subjects made few eye movements, fixating only two or three grid positions, rarely along the periphery of the matrix. Experiment 2, however, demonstrated that the relationship between integration and eye movement behavior differed in dynamic displays in which the first and the second arrays were presented in different spatial locations. In this case, in which eye movements were directly related to task performance, observers did view the grid systematically. Although the number of eye movements did not differ from that in the case of the static display, the eye movements that were made were more strategic, since positions previously occupied by a dot from the first array were the most common loci of fixation.

These results suggest that the differing use of eye movements in each experiment may be related to different encoding strategies that were engaged on the basis of the different task demands in the static and the dynamic cases. In the static display case, VSTM was maintained in the same spatial location as that in which it was generated. This apparently could be accomplished with covert deployments of attention. In the dynamic case, however, VSTM was generated in a location that did not overlap with one of the to-be-integrated percepts. In order to “move” the memory trace, overt shifts of attention were required.

We hypothesize that these strategy differences may also explain why, given previous evidence for a link between eye movements and memory consolidation and maintenance, a tighter relationship between eye movements and integration was not found. Although the results of Experiment 2 lend support to the hypothesis that VSTM consolidation and maintenance employ eye movements, integration was found to be quite independent of eye movements in Experiment 1. The present results highlight Neisser's (1967) prediction that the relationship between eye movements and the reinstatement of a visual display in memory should not be absolute. At least in some instances, the consolidation and maintenance of VSTM can rely on covert shifts of attention. When covert shifts are insufficient for bringing a stimulus into the focus of attention during perception, observers employ overt shifts by moving their eyes. The same may be true of memory.

As a general conclusion, overt shifts of attention are not necessary for integration, although they appear to play a greater role in integration when temporally and spatially separated stimuli must be integrated. This means that nonfoveated aspects of a display can be included in the memory representation used in integration. Indeed, even in Experiment 2, when the eyes moved systematically to the to-be-remembered array, very few positions in the grid were actually foveated. In order to produce the observed levels of performance, then, many regions that were not foveated must have been remembered. In terms of memory rehearsal, eye movements may be helpful, but this seems limited to cases in which they are required in order to view all the relevant information. Thus, the ISI required for integration (1,000–1,500 msec) is not determined by scan time alone.

Another conclusion that can be drawn from the present experiments is that engaging the eye movement system does not hurt integration if sufficient time is allowed for encoding and VSTM generation. In fact, peak accuracy in Experiment 1 (no eye movements because of spatial overlap) was 61%, and in Experiment 2 (eye movements because of spatial displacement), it was 57%. The results differ strikingly from those in prior studies in which the integration of visual information across saccadic eye movements has been investigated. In the past, a transsaccadic memory store that maintains and accrues visual information across successive fixations was proposed as a mechanism for relating information in one fixation to that obtained later (Breitmeyer, Kropfl, & Julesz, 1982; McConkie & Rayner, 1976). In general, however, little evidence for information accrual over saccades has been found (Bridgeman & Mayer, 1983; Henderson, 1997; Irwin et al., 1988; Irwin et al., 1983; McConkie, 1991; O'Regan & Levy-Schoen, 1983; Rayner & Pollatsek, 1983), although some limited visual information, such as pattern structure and spatial relationships, is remembered from fixation to fixation (e.g., Carlson-Radvansky, 1999; Carlson-Radvansky & Irwin, 1995; Irwin 1991, 1992; Pollatsek & Rayner, 1992). A likely reason for the different results from prior work and the present experiments concerns the role of time. Past researchers investigated integration from one fixation to the next within the time frame of one fixation (a few hundred milliseconds); in the present study, the observers were afforded much more time (thousands of milliseconds). Because of this, integration did not occur from one fixation to the next but, rather, over the course of multiple fixations that could be used to help reconstruct an initial stimulus after its disappearance from view. This presents an interesting contrast to past work on integration and eye movements and should be examined more closely in future research.

The finding that spatially displaced visual stimuli can be integrated suggests that the memory representation of the first array is not tightly bound to a particular region of space. It is not spatial locations that are being remembered but information about the stimulus itself, in-



dependently of where it was presented. What is the nature of the information that is retained during memory-percept integration? One could conceive of memory-percept integration as a process that combines object representations. That is, each dot (or a bound unit, or "chunk," of a few dots) in each array is an independent entity, maintained in VSTM as a unique object. One could also conceive of it as a process that integrates pattern information. Instead of treating the dot arrays as collections of individual objects, they are treated as global patterns akin to a spatial layout. It is this layout information, not object identity information, that is integrated. Currently, we are designing experiments to dissociate these possibilities.

Whether memory-percept integration unites object or layout information, the properties of memory-percept integration observed in the present experiments suggest that this integration process may serve a functional role in scene processing. Memory-percept integration is capable of combining information that is perceived at discrete points in time, as well as information that misaligns in space. In real-world perception, conceptually related pieces of information are rarely viewed under conditions of constant temporal, retinal, and spatial overlap. The ability of the system to account for variability along these dimensions simultaneously is critical if memory-percept integration is to play any role in the development of scene representations. We are now beginning to directly examine memory-percept integration as a means for constructing object and scene representations.

#### REFERENCES

- AVERBACH, E., & CORIELL, H. S. (1961). Short-term memory in vision. *Bell Systems Technical Journal*, **40**, 309-328.
- BRANDT, S. A., & STARK, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience*, **9**, 27-38.
- BREITMEYER, B. G. (1980). Unmasking visual masking: A look at the "why" behind the veil of the "how." *Psychological Review*, **87**, 52-69.
- BREITMEYER, B. G. (1984). *Visual masking: An integrative approach*. New York: Oxford University Press.
- BREITMEYER, B. G., & GANZ, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review*, **83**, 1-36.
- BREITMEYER, B. G., KROPFL, W., & JULESZ, B. (1982). The existence and role of retinotopic and spatiotopic forms of visual persistence. *Acta Psychologica*, **52**, 175-196.
- BRIDGEMAN, B., & MAYER, M. (1983). Failure to integrate visual information from successive fixations. *Bulletin of the Psychonomic Society*, **21**, 285-286.
- BROCKMOLE, J. R., IRWIN, D. E., & WANG, R. F. (2003). The locus of spatial attention during the temporal integration of visual memories and visual percepts. *Psychonomic Bulletin & Review*, **10**, 510-515.
- BROCKMOLE, J. R., & WANG, R. F. (2003). Integrating visual images and visual percepts across time and space. *Visual Cognition*, **10**, 853-874.
- BROCKMOLE, J. R., WANG, R. F., & IRWIN, D. E. (2002). Temporal integration between visual images and visual percepts. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 315-334.
- BUSWELL, G. T. (1935). *How people look at pictures: A study of the psychology of perception in art*. Chicago: University of Chicago Press.
- CARLSON-RADVANSKY, L. A. (1999). Memory for relational information across eye movements. *Perception & Psychophysics*, **61**, 919-934.
- CARLSON-RADVANSKY, L. A., & IRWIN, D. E. (1995). Memory for structural information across eye movements. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 1441-1458.
- COLTHEART, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics*, **27**, 183-228.
- DI LOLLO, V. (1980). Temporal integration in visual memory. *Journal of Experimental Psychology: General*, **109**, 75-97.
- ENNS, J. T., & DI LOLLO, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science*, **8**, 135-139.
- ERIKSEN, B. A., & ERIKSEN, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, **16**, 143-149.
- ERIKSEN, C. W., & COLLINS, J. F. (1967). Some temporal characteristics of visual pattern perception. *Journal of Experimental Psychology*, **74**, 476-484.
- ERIKSEN, C. W., & ST. JAMES, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics*, **40**, 225-240.
- HEBB, D. O. (1968). Concerning imagery. *Psychological Review*, **75**, 466-477.
- HENDERSON, J. M. (1997). Transsaccadic memory and integration during real-world object perception. *Psychological Science*, **8**, 51-55.
- HENDERSON, J. M., WEEKS, P. A., & HOLLINGWORTH, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 210-228.
- IRWIN, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology*, **23**, 420-456.
- IRWIN, D. E. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 307-317.
- IRWIN, D. E. (2004). Fixation location and fixation duration as indices of cognitive processing. In J. M. Henderson & F. Ferreira (Eds.), *The integration of language, vision, and action: Eye movements and the visual world* (pp. 105-134). New York: Psychology Press.
- IRWIN, D. E., & ANDREWS, R. V. (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., BROWN, J. S., & SUN, J. S. (1988). Visual masking and visual integration across saccadic eye movements. *Journal of Experimental Psychology: General*, **117**, 276-287.
- IRWIN, D. E., YANTIS, S., & JONIDES, J. (1983). Evidence against visual integration across saccadic eye movements. *Perception & Psychophysics*, **34**, 49-57.
- 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 visual arrays in visual short-term memory. *Experimental Psychology*, **52**, 39-46.
- JUST, M. A., & CARPENTER, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, **87**, 329-354.
- KAHNEMAN, D. (1968). Method, findings, and theory in studies of visual masking. *Psychological Bulletin*, **70**, 404-425.
- LAENG, B., & TEODORESCU, D. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science*, **26**, 207-231.
- LEFTON, L. A. (1973). Metacontrast: A review. *Perception & Psychophysics*, **13**, 161-171.
- LUCK, S. J., & VOGEL, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, **390**, 279-281.
- MATIN, E. (1975). The two-transient (masking) paradigm. *Psychological Review*, **82**, 451-461.
- MCCONKIE, G. W. (1991). Perceiving a stable visual world. In J. van Rensbergen, M. Deuijver, & G. d'Ydewalle (Eds.), *Proceedings of the Sixth European Conference on Eye Movements* (pp. 5-7). Leuven: Laboratory of Experimental Psychology.
- MCCONKIE, G. W., & RAYNER, K. (1976). Identifying the span of the effective stimulus in reading: Literature review and theories of reading.

- In H. Singer & R. B. Ruddle (Eds.), *Theoretical models and processes of reading* (pp. 137-162). Newark, DE: International Reading Association.
- NEISSER, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.
- O'REGAN, J. K., & LEVY-SCHOEN, A. (1983). Integrating visual information from successive fixations: Does trans-saccadic fusion exist? *Vision Research*, **23**, 765-768.
- PEARSON, D. G., & SAHRAIE, A. (2003). Oculomotor control and the maintenance of spatially and temporally distributed events in visuo-spatial working memory. *Quarterly Journal of Experimental Psychology*, **56A**, 1089-1111.
- PHILLIPS, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, **16**, 283-290.
- POLLATSEK, A., & RAYNER, K. (1992). What is integrated across fixations? In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 166-191). New York: Springer-Verlag.
- RAYNER, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, **124**, 372-422.
- RAYNER, K., & POLLATSEK, A. (1983). Is visual information integrated across saccades? *Perception & Psychophysics*, **34**, 39-48.
- SHULMAN, G. L., & WILSON, J. (1987). Spatial frequency and selective attention to spatial location. *Perception*, **16**, 103-111.
- SPIVEY, M. J., & GENG, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, **65**, 235-241.
- TURVEY, M. T. (1973). On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review*, **80**, 1-52.
- VOLKMANN, F. C., SCHICK, A. M., & RIGGS, L. A. (1968). Time course of visual inhibition during voluntary saccades. *Journal of the Optical Society of America*, **58**, 562-569.
- WATSON, D. G., & HUMPHREYS, G. W. (2000). Visual marking: Evidence for inhibition using a probe-dot detection paradigm. *Perception & Psychophysics*, **62**, 471-481.
- YARBUS, A. (1967). *Eye movements and vision*. New York: Plenum.

#### NOTE

1. An additional experiment conceptually replicated Experiment 2 but only tested predisplacement delays of 150, 600, and 1,500 msec, to increase statistical power at each ISI. No reliable effect of predisplacement delay on accuracy was observed [ $F(1,10) < 1$ ]. We conclude that the temporary decrement in accuracy in the early component condition in Experiment 2 at predisplacement delays between 500 and 700 msec, although reliable, is not indicative of any differences in the image representations available at those time points but, rather, is likely due to unsystematic variation or Type II error.

(Manuscript received January 26, 2004;  
revision accepted for publication July 15, 2004.)