

Investigating temporal properties of covert shifts of visual attention using the attentional walk task

ELISABETH HEIN AND CATHLEEN M. MOORE
University of Iowa, Iowa City, Iowa

The *attentional walk task* was used to investigate the temporal properties of covert shifts of attention. Observers shifted attention within arrays of identical items in response to a series of auditory commands and reported the color of the final disk. The density of the arrays and the timing of the shift commands varied. Performance decreased as density increased, and the minimal amount of time needed to shift attention depended on the density of the display, varying from 350 to 750 msec. In addition, the observers were able to maintain attentional focus for at least 3,500 msec without a decline in performance, regardless of density. Thus, although the ability to maintain attention at a given position was found to be independent of the precision with which that location was defined, more precise attentional shifts required more time to execute.

Selection is a ubiquitous problem for sensory systems. When one is reading, for example, many lines of text fill the visual field. Selection processes must be engaged to filter out irrelevant parts of the text and process relevant parts. When one searches within a complex scene for a desired object (e.g., when one searches for a book on a shelf) or when one monitors a scene for critical events (e.g., when one serves as a lifeguard at a busy pool), parts of the scene must be processed while other parts are ignored. Upon thinking about examples like these, it becomes apparent that success at these tasks requires a flexible focus of selection. Perception cannot be understood in terms of fixed sensory mechanisms. Rather, observer-dependent selection—generally referred to as *selective attention*—must be included in the explanation. Moreover, an account of how the focus of selection is controlled must be part of that explanation.

We investigated some of the temporal properties of voluntary covert control of visual attention, using the *attentional walk task* (Intriligator & Cavanagh, 2001). The temporal properties of overt shifts of attention—eye movements—have been investigated extensively and are thought to reflect specific underlying cognitive processes (for reviews, see, e.g., Carpenter, 1988; Henderson & Hollingworth, 1999). Covert shifts of attention, however, occur between eye movements (e.g., Henderson, Pollatsek, & Rayner, 1989), and they tend to precede eye movements to the new point of fixation (e.g., Hoffman & Subramaniam, 1995; Kowler, Anderson, Doshier, & Blaser, 1995). An effective tool for measuring the temporal properties of covert shifts of attention, therefore, could provide finer grained insight into the cognitive processing that underlies overt scanning behavior.

In the attentional walk task, observers shift attention, but not fixation, from one item to another within arrays of

identical items (e.g., disks) in response to a series of commands, such as tones (see Figure 1). The reliability with which observers can execute these *attentional walks* is probed by asking them to report something about the final item in the walk, such as its color. The density of the array is varied from sparse, where reliable shifts of attention are easily achieved, to dense, where reliable shifts of attention are difficult or impossible. Previous work has shown that performance in the attentional walk task reflects limitations in the precision with which attention can be localized within an array of items (Moore, Lanagan-Leitzel, Chen, Halterman, & Fine, 2007; Moore, Lanagan-Leitzel, & Fine, 2008). In particular, rather than reflecting only the spatial resolution of selection, which can be thought of as the size of the metaphoric *spotlight of attention* (Cave & Bichot, 1999), it seems to reflect the spatial uncertainty with which attention can be focused, which can be thought of as the tremor, so to speak, that occurs while the metaphoric spotlight is moved from one point of focus to another.

Although the attentional walk task was designed to investigate the spatial properties of attention (Intriligator & Cavanagh, 2001), it can also be used to investigate the temporal properties of attentional control by manipulating the time between shift commands. This was the focus of the present study. In particular, we asked whether the minimal time needed to endogenously shift attention within arrays of items varies systematically with the density of the display and whether maintaining attention on a given item for extended periods of time varies with density.

EXPERIMENT 1

In Experiment 1, we varied the time between commands to shift attention—*interbeep interval* (IBI)—within an at-

E. Hein, elisabeth-hein@uiowa.edu

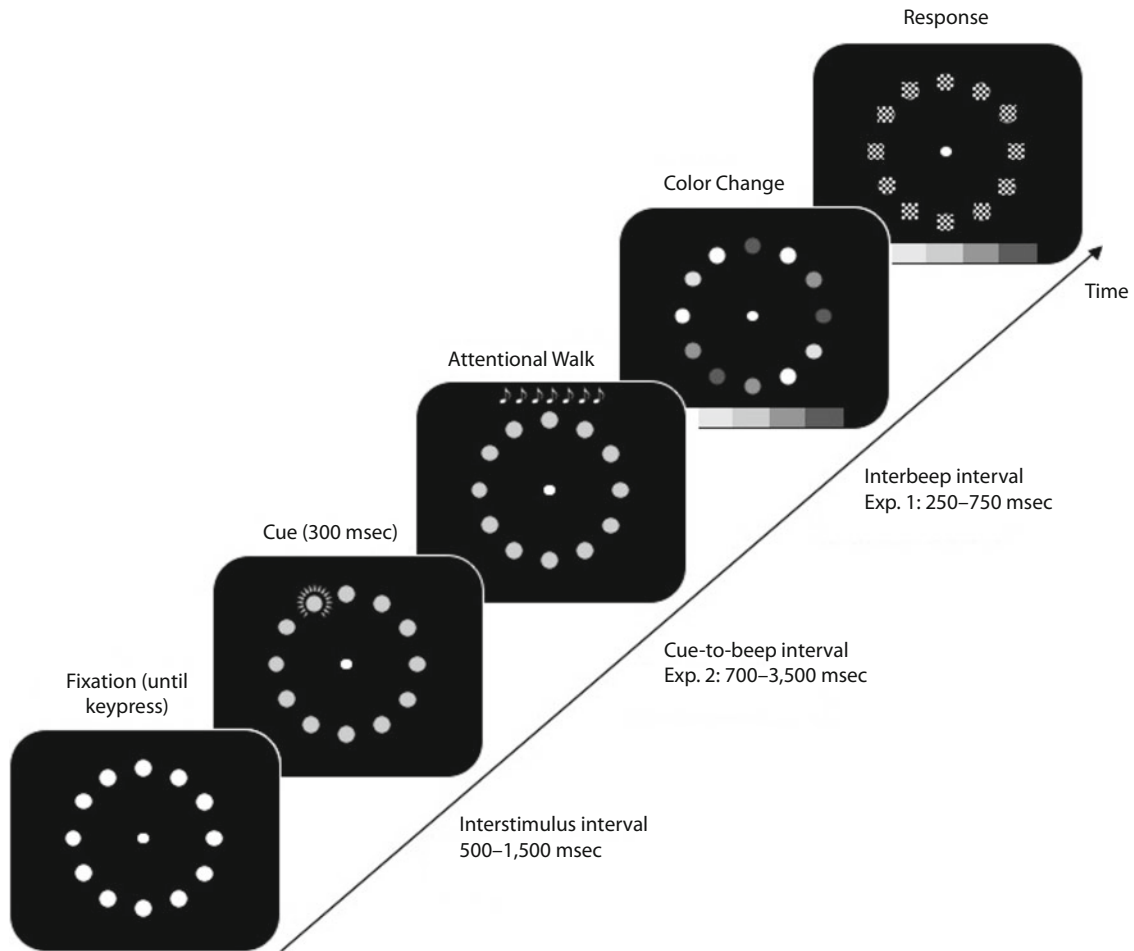


Figure 1. Illustration of the basic trial sequence in both experiments. After the participant indicated that he was fixating the center, a cue appeared that defined the starting point of the attentional walk. During the walk, the participants shifted their attention from dot to dot according to a series of beeps. At the end of the trial, the participant indicated the color of the disk that their attention had shifted to after the last beep by clicking with the mouse on the response field of the same color (for illustration purposes, here indicated in different gray levels). Note—In Experiment 2, no masks were presented at the end of the trial.

tentional walk task. We expected that performance would be poor at the very shortest intervals and that performance would increase to some maximum level as the IBI increased. The IBI at which performance increases beyond chance would provide an estimate of the minimal time needed to covertly shift attention within arrays of that particular density. The question was whether this minimal time would depend on density. Given that the displays were constantly present and the locus of the next shift was known, minimal shift time might not depend on density. If, however, control depends on the precision with which the shift must be made—as, for example, the control of movements of the body does (e.g., Fitts & Peterson, 1964)—then the minimal shift time might vary systematically with density.

Method

Observers. Six observers (mean age, 23 years; 2 female, 4 male) participated in the experiment. All were naive as to the purpose of the experiment. Half of the participants had participated in other experiments using the attentional walk task; the other half were new to the task.

Apparatus. The experiment was conducted using an IBM-compatible PC driving a 17-in. CRT monitor at a spatial resolution of $1,024 \times 768$ pixels and a vertical refresh rate of 85 Hz. Software was generated in MATLAB (Version 6.5, Release 13; MathWorks, Natick, MA), using the Psychtoolbox extensions (Version 2.5; Brainard, 1997; Pelli, 1997). Viewing distance was fixed at 51 cm with a chin- and headrest.

Stimuli. Displays consisted of circular arrays of 12, 18, or 24 disks (0.67° in diameter) presented on a dark background, centered around a central white fixation dot (0.45° in diameter), and distributed evenly on an imaginary 6.84° radius circle (measured from the center of fixation to the center of the disks). The cue consisted of one of the disks expanding to 0.87° in diameter for 300 msec. All the disks were white during fixation. They then changed to a single color (red, green, cyan, purple, or blue, randomly) during the walk. For the response displays, the disks changed to individual colors from this set, chosen pseudorandomly with the constraints that the disks to either side of the target disk (i.e., the one on which the attentional walk should have ended) were different colors and no two adjacent disks were the same color. Finally, each disk was replaced with color-noise masks, which were $0.67^\circ \times 0.67^\circ$ rectangles filled randomly with small squares of the five possible colors. A response display was shown 11.75° below the center of the response array.

It consisted of a linear array of five $2.81^\circ \times 7.85^\circ$ rectangles, one in each of the colors used. Each walk step was indicated using the computer's default beep at medium volume. The experiment was conducted with the room lights on.

Task. The participants were asked to remain fixated on the central fixation marker throughout the trial. The task was to shift attention to the cued disk in the array and then, if a series of tones occurred (*multiple-step* condition), to shift attention clockwise/counterclockwise (fixed for a given participant) following each tone from the current disk to the next one. When the series of tones ended, the array changed to a response array, and the observers reported the color of the disk on which the attentional walk had ended by clicking on the appropriate rectangle in the response menu. On some trials (*zero-step* condition), no tones occurred following the cue. Instead, the cue display was followed immediately by the response array, and the participants had to indicate the color of the cued disk.

Design. A 6 (IBI: 250, 300, 350, 400, 450, or 750 msec) \times 3 (density: 12, 18, 24) within-subjects design was used. Walk direction was fixed for a given participant as clockwise or counterclockwise and was counterbalanced across participants. Interstimulus interval (ISI) and density were mixed within blocks of trials. In addition, 20% of the trials (randomly intermixed) were zero-step trials on which no attentional walk was required. Trials on which the mouse click was outside of the response menu were repeated. A total of 36 zero-step observations per density condition and 24 multiple-step observations for each density and IBI condition were collected for each participant.

Procedure. The participants completed two 1-h sessions, usually on 2 consecutive days, but in all cases within a week. The first session began with a set of written instructions describing the walk task. The observers then completed about 20 practice trials until they felt comfortable with the task. Breaks were provided throughout each session every 30 trials.

Trial events are illustrated in Figure 1. Each trial began with the presentation of a central fixation dot, surrounded by an array of white disks. The participants started a trial by fixating the central fixation point and pressing the left mouse button to indicate that they were ready to start the trial. The disks then all changed to a single color, and after an ISI of 500–1,500 msec, one of the disks blinked for 300 msec to indicate the starting point of the attentional walk. After a randomly chosen ISI of 725, 750, or 775 msec, either the response display (in which all the disks changed colors and the response menu appeared) was presented (*zero-step* condition) or a series of beeps (6, 7, 8, or 9 beeps, randomly chosen) started (*multiple-step* condition). IBI was fixed within a given trial but was variable across trials (250, 300, 350, 400, 450, or 750 msec). After a randomly chosen variable interval from 250 to 750 msec following the final beep, the response display was presented. All the disks were replaced by masks 1,500 msec later, which remained on the screen until a response was made. Incorrect responses and responses outside of the menu area were indicated with written feedback at the center of the display, indicating a wrong response or a mouse click outside of the menu area. The participants clicked the mouse to continue to the next trial.

Results and Discussion

Figure 2A shows mean accuracy as a function of IBI and density in Experiment 1. Mean accuracy in the zero-step and multiple-step conditions was analyzed in separate ANOVAs, with density (three levels: 12, 18, or 24) as the only variable for the zero-step condition and both IBI (six levels: 250, 300, 350, 400, 450, or 750 msec) and density (three levels: 12, 18, or 24) as the variables for the multiple-step condition. Alpha was set at .05 for these and all subsequent analyses. Whenever appropriate, p values were Greenhouse–Geisser corrected.

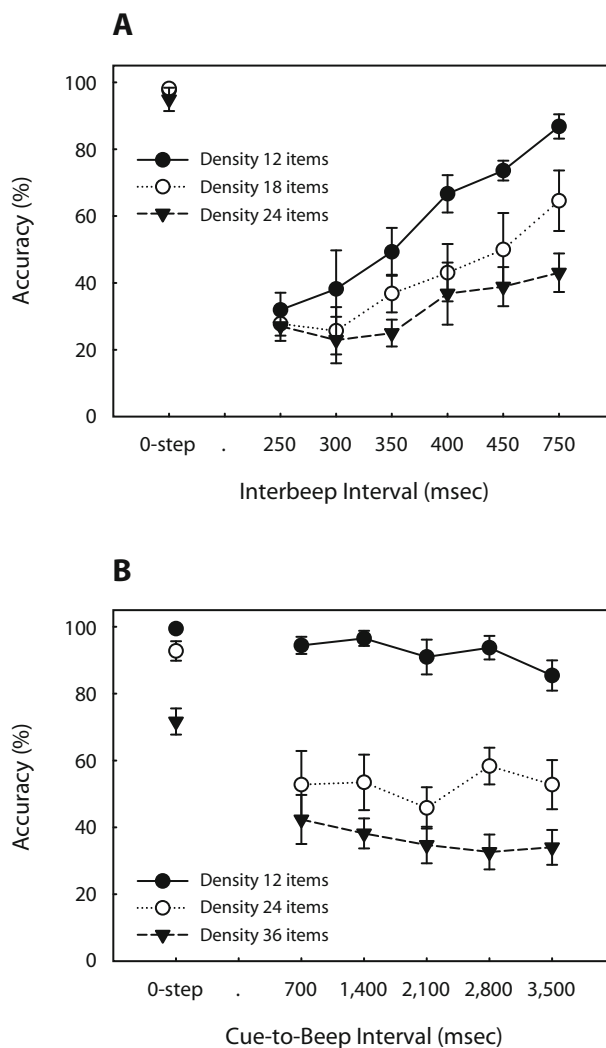


Figure 2. Mean accuracy shown as a function of density (number of items in the display) and interbeep interval (250–750 msec) for Experiment 1 (A) and mean accuracy shown as a function of density and cue-to-beep interval (700–3,500 msec) for Experiment 2 (B). In both panels, the single points represent performance for the zero-step control conditions for the different densities, and the extended functions represent the multiple-step condition for the different densities. The error bars represent the standard error for each condition.

In the zero-step condition, performance was generally very high and changed very little with density [97.7%, 98.2%, and 94.9%, from lowest to highest density; $F(2,10) = 0.58$, n.s.], indicating that the disks were clearly separated at all densities, with little or no effects of masking or crowding (e.g., Pelli, Palomares, & Majaj, 2004). In the multiple-step condition, there was a reliable main effect of density [$F(2,10) = 20.97$, $p < .001$] and a reliable effect of IBI [$F(5,25) = 21.69$, $p = .001$], as well as a trend for an interaction between the two factors [$F(10,50) = 3.08$, $p = .06$]. Accuracy decreased with increasing density from 57.8% for the lowest density condition, to 41.3% for the medium density, to 32.3% in

the highest density condition. In addition, performance increased from the shortest to the longest IBI condition (28.9%, 28.9%, 37.0%, 48.8%, 54.2%, and 64.8%). Thus, performance in the walk task not only was greatly affected by density, but also was affected by the speed with which the attentional walk had to be executed.

As the trend for the interaction between density and IBI suggests, the increase in performance with IBI was more pronounced the lower the density was. We conducted a further ANOVA, looking at the different densities separately, with IBI (six levels: 250, 300, 350, 400, 450, and 750 msec) as the only factor. These analyses revealed significant main effects for each of the densities [$F(5,25) = 17.34, p < .01$, for density 12; $F(5,25) = 10.45, p < .01$, for density 18; and $F(5,25) = 4.23, p < .05$, for density 24]. Separate post hoc comparisons (Fisher's least significant difference) of the different IBI conditions showed that lower densities yielded shorter IBIs at which performance improved beyond that in the shortest IBI condition (i.e., about chance performance). For density 12, IBIs of 350, 400, 450, and 750 msec were all reliably different from the IBI of 250 msec. For density 18, only IBIs of 400–750 msec were reliably different from the IBI of 250 msec. Finally for density 24, only the IBI of 750 msec was reliably different from the IBI of 250 msec.

We followed the response display with masks in order to prevent the participants from using an alternative strategy. In particular, they might try to count the number of beeps and then select the final disk on the basis of that count during the response period, rather than shifting attention in response to individual beeps. Masking the response displays within 1,750–2,250 msec following the final beep, we reasoned, would render this strategy unlikely. First, the observers would have to realize that a given beep was the final beep, which would require the time of the IBI of the current trial or at least 250 msec, because that is the shortest time between the final beep and the response display that was used. They would then have to use their count information to find the relevant disk in the remaining time. Although unlikely, we sought to confirm that this was too difficult an option by conducting an experiment with 6 naive participants, in which we asked them to use the counting strategy without mentioning the attentional shift task at all. During instruction and practice, the observers complained that they could not do the task but they were asked to do the best that they could. Following completion of the experiment, we asked them how they had done the task. All 6 participants described a strategy that involved shifting attention while the beeps were playing. Thus, it seems unlikely that when the explicit task was the attentional shift task, the observers resorted, instead, to a counting strategy that this group reported being unable to do.

In summary, Experiment 1 demonstrates that the attentional walk task is limited not only by spatial density, but also by the time available for making shifts. In particular, the higher the density was, the more time was needed to attentionally navigate within arrays of items. This is the pattern of results that would be expected if the attentional walk reflects the precision with which attention can be

localized and the focusing of attention, like the targeting of effectors (e.g., Fitts & Peterson, 1964), requires more time when greater precision is necessary.

EXPERIMENT 2

Whereas in Experiment 1 we asked whether there is a minimum time between shift commands for supporting above-chance performance in the attentional walk task, in Experiment 2 we asked whether there is a maximum time between shift commands beyond which performance begins to decrease. The idea is that maintaining attention to a particular item within a dense array might require active maintenance, which would be subject to increasing error with increasing time. This hypothesis predicts that there should be a duration at which performance begins to decrease from asymptotic levels.

Experiment 2 included longer intervals during which observers were required to maintain their attentional focus. This was done by manipulating the time between the initial cue and the first step in the attentional walk from 700 to 3,500 msec, referred to as the *cue-to-beep interval* (CBI). We manipulated the CBI, rather than the IBI, during the walk because multiple steps with very long IBIs would have resulted in impractically long trials.

Method

The method was the same as that in Experiment 1, except where noted.

Observers. Six observers (mean age, 24 years; 3 female, 3 male) participated in the experiment. None had participated in Experiment 1, but all the participants were familiar with the attentional walk task paradigm in general. All the participants reported normal or corrected-to-normal visual acuity and color vision. Each participant was tested in advance for the ability to maintain fixation as monitored by the eyetracker.

Apparatus, Stimuli, and Task. The computer equipment and software used to develop the experiment were the same as those in Experiment 1. Eye movements were monitored with a video-based monocular eyetracker (ViewPoint Eye Tracker by Arrington Research, Inc.) controlled with Viewpoint software (Version PC60). Viewing distance was fixed at 51 cm with a chin- and headrest. The stimuli and task were the same as those as in Experiment 1, with the exception that the circular arrays consisted of 12, 24, or 36 disks.

Design. A 5 (CBI: 700, 1,400, 2,100, 2,800, or 3,500 msec) \times 3 (density: 12, 24, or 36) within-subjects design was used. Each observer participated in six experimental sessions of 75 trials each (60 multiple-step and 15 zero-step trials), resulting in 30 zero-step observations for each density condition and 24 multiple-step observations for each density and CBI condition.

Procedure. In the first session, the observers were familiarized with the eyetracker and practiced fixating during the walk task. The window of tolerance around fixation was set to either 1.5° or 2.5° around the fixation dot, depending on the participant's ability to fixate. Participants who lost fixation on more than 50% of the trials during the second half of this training phase were discontinued. Following the initial session of familiarization with the eyetracker and practice of the task, the participants completed six 30- to 40-min sessions on different, usually consecutive, days. Breaks were provided throughout each session.

The specific trial sequence was the same as that in Experiment 1, with the exception that the CBI varied and the IBI for the remainder of the sequence was always 725, 750, or 775 msec, chosen randomly. Furthermore, the participants initiated a trial by fixating and pressing

the left mouse button to indicate that they were fixating. If fixation was within the fixation area, the trial started. If the tracker could not detect fixation within the specified window, a beep indicated failed fixation, and the participant was asked to refixate and restart the trial. Once fixation was accepted, the disks changed color, and the trial began. One thousand milliseconds after the final beep, all the disks changed color, and the response menu appeared at the bottom of the screen. The response display remained visible until a mouse keypress was registered or for 10 sec if no response occurred. Eye position was monitored throughout the trial. If fixation went outside of the fixation window, the trial was aborted and repeated later.

Results and Discussion

Figure 2B shows mean accuracy as a function of CBI and density from Experiment 2. Separate ANOVAs were conducted for the zero- and multiple-step conditions. Density (three levels: 12, 24, or 36) was the only variable for the zero-step analysis. Both CBI (five levels: 700, 1,400, 2,100, 2,800, 3,500 msec) and density (three levels: 12, 24, or 36) were the variables for the multiple-step analysis.

In the zero-step condition, performance declined somewhat with increasing density [99.4%, 92.8%, and 71.7% from lowest to highest density; $F(2,10) = 30.77, p < .001$], revealing some masking or crowding irrespective of having to shift attention within the array.

In the multiple-step condition, there was again a substantial and reliable main effect of density [$F(2,10) = 89.93, p < .001$] but no main effect of CBI [$F(4,20) = 1.97, n.s.$] and no interaction between CBI and density [$F(8,40) = 0.95, n.s.$]. In particular, performance in the attentional walk task decreased with increasing density (92.2%, 52.6%, and 36.4%), whereas performance was not at all affected by the amount of time that attention had to be maintained on the cue before the attentional walk was started (from lowest to highest CBI: 63.2%, 62.7%, 57.2%, 61.6%, and 57.4%). Thus, despite the range of times having been increased to as long as 3.5 sec and very demanding density conditions having been used, there was no evidence of a decline in performance with CBI.

GENERAL DISCUSSION

In this study, the attentional walk task was used to investigate the temporal properties of the control of spatial selective attention. In Experiment 1, we asked whether the minimal time between shift commands that is necessary to support above-chance performance in the attentional walk task, assuming that there is one, varies with the density of the display. If the control of attention is similar to the control of effectors, one might predict, indeed, that the more precise the control of attention must be, the greater will be the amount of time needed to successfully make the shift. This is exactly what was found. Although performance reached above-chance levels with intervals between shift commands as short as 350 msec in density 12 displays, it did not do so until 400 msec for density 18 and somewhere between 450 and 750 msec for density 24 displays.

The results of Experiment 1 provide a rough estimate of the minimal time needed to execute voluntary shifts of

covert attention in response to external commands and demonstrate that this minimal time depends on the density of the display. Thus, the speed with which attention can be *targeted* is dependent on the precision with which the targeting must be done, just as the speed with which aimed hand movements can be targeted is dependent on the precision with which the targeting must be done (e.g., Fitts & Peterson, 1964).

These results—increased time needed to execute covert shifts of attention with increasing density—are in line with those in studies that have shown that saccade latencies depend on the density of items in a visual search display (Motter & Belky, 1998). It will be noted, however, that the estimate of the minimal time that is required to shift attention from one item to the next, even for the lowest density condition, is longer than many previous estimates of shift time. For example, saccades (overt shifts of attention) to the target in a visual search paradigm can occur as quickly as about 150 msec (e.g., Motter & Belky, 1998). Given the close relationship between eye movements and attentional shifts (e.g., Hoffman & Subramaniam, 1995; Kowler et al., 1995), one could ask why the estimate of the minimal time required for covert attentional shifts in the attentional walk task is more than twice as long. Moreover, estimates of shifts of attention on the basis of the slopes of functions relating response time to display size in visual search tend to be on the order of 30–80 msec for searches assumed to require something like serial search (see Wolfe, 1994, for a review). These estimates, however, are probably underestimates (see, e.g., Moore, Egeth, Berglan, & Luck, 1996, for a related discussion). Moreover, the requirements of the attentional walk task are quite different from those in either standard visual search tasks or even free viewing of scenes in which eye movements are tracked. Specifically, the attentional shifts must occur to specified items, whereas in the visual search, shifts are unspecified.

In Experiment 2, we addressed temporal limitations at the other end of the continuum. In particular, we asked whether performance in the attentional walk task begins to decrease as the time during which attention must be maintained on a given item increases, and whether the time at which it does begin to decrease, assuming that it does, depends on density. In fact, within the range that we tested, up to 3.5 sec, there was no evidence of decreasing performance in any of the density conditions. Thus, maintaining attention for 3.5 sec was no more difficult than maintaining it for 750 msec (a shift time that supports asymptotic performance at all densities), independently of the density condition.

In summary, the attentional walk task allowed us to investigate the temporal properties of shifting attention within dense arrays. The minimal amount of time needed to execute such shifts depended on the density of the array. In contrast, no upper limit for the time that attention can be maintained within a dense array was found, at least for the densities (12, 24, 36) and the time range (up to 3.5 sec) we used. These results contribute to the general view that the attentional walk task reflects limitations of

the dynamic process of shifting attention within dense arrays (Moore et al., 2007; Moore et al., 2008), rather than just limitations on the spatial extent of attention, as has been suggested in the past (e.g., Intriligator & Cavanagh, 2001).

AUTHOR NOTE

The work reported here was supported in part by NIH Grant MH067793. We thank Elisabeth Fine and John Palmer for useful discussions of this work, and Sergey Yurgensen for technical assistance. We also thank Steve Luck, Richard Abrams, and two anonymous reviewers for their helpful comments and suggestions, which greatly improved the original manuscript. Correspondence concerning this article should be addressed to E. Hein, Department of Psychology, University of Iowa, E11 Seashore Hall, Iowa City, IA 52242 (e-mail: elisabeth-hein@uiowa.edu).

REFERENCES

- BRAINARD, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, **10**, 433-436.
- CARPENTER, R. H. S. (1988). *Movements of the eyes*. London: Pion.
- CAVE, K. R., & BICHOT, N. P. (1999). Visuospatial attention: Beyond a spotlight model. *Psychonomic Bulletin & Review*, **6**, 204-223. doi:10.1163/156856897X00357
- FITTS, P. M., & PETERSON, J. R. (1964). Information capacity of discrete motor responses. *Journal of Experimental Psychology*, **67**, 103-112.
- HENDERSON, J. M., & HOLLINGWORTH, A. (1999). High-level scene perception. *Annual Reviews of Psychology*, **50**, 243-271. doi:10.1146/annurev.psych.50.1.243
- HENDERSON, J. M., POLLATSEK, A., & RAYNER, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Perception & Psychophysics*, **45**, 196-208.
- HOFFMAN, J. E., & SUBRAMANIAM, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics*, **57**, 787-795.
- INTRILIGATOR, J., & CAVANAGH, P. (2001). The spatial resolution of attention. *Cognitive Psychology*, **43**, 171-216. doi:10.1006/cogp.2001.0755
- KOWLER, E., ANDERSON, E., DOSHER, B., & BLASER, E. (1995). The role of attention in the programming of saccades. *Vision Research*, **35**, 1897-1916.
- MOORE, C. M., EGETH, H., BERGLAN, L. R., & LUCK, S. J. (1996). Are attentional dwell times inconsistent with serial visual search? *Psychonomic Bulletin & Review*, **3**, 360-365.
- MOORE, C. M., LANAGAN-LEITZEL, L. K., CHEN, P., HALTERMAN, R., & FINE, E. M. (2007). Nonspatial attributes of stimuli can influence spatial limitations of attentional control. *Perception & Psychophysics*, **69**, 363-371.
- MOORE, C. M., LANAGAN-LEITZEL, L. K., & FINE, E. M. (2008). Distinguishing between the precision of attentional localization and attentional resolution. *Perception & Psychophysics*, **70**, 573-582. doi:10.3758/PP.70.4.573
- MOTTER, B. C., & BELKY, E. J. (1998). The zone of focal attention during active visual search. *Vision Research*, **38**, 1007-1022.
- PELLI, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, **10**, 437-442.
- PELLI, D. G., PALOMARES, M., & MAJAJ, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, **4**(12, Art. 12), 1136-1169. doi:10.1167/4.12.12
- WOLFE, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, **1**, 202-238.

(Manuscript received September 12, 2008;
revision accepted for publication October 15, 2009.)