A new estimation of the duration of attentional dwell time

JAN THEEUWES and RICHARD GODIJN
Vrije Universiteit, Amsterdam, The Netherlands

and

JAY PRATT
University of Toronto, Toronto, Ontario, Canada

How rapidly can attention move from one object to the next? Previous studies in which the dwell time paradigm was used have estimated attentional switch times of 200–500 msec, results incompatible with the search rate estimates of 25–50 msec shown in numerous visual search studies. It has been argued that dwell times are so long in the dwell time paradigm because the attentional shifts measured are unlike those used in visual search. In the present experiment, a variation of a visual search task was used, in which serial endogenous (volitional) deployments of attention were measured directly by means of a probe reaction time task. The experiment revealed a dwell time of about 250 msec, consistent with the faster estimates from other dwell time studies. This result suggests that endogenous shifts of attention may be relatively slow and that the faster attentional shifts estimated from visual search tasks may be due to the involvement of bottom-up processes.

In visual search, observers search through a variable set of nontarget items for a specific target item that may or may not be present. Reaction time (RT) is measured as a function of the number of items in the display (set size), and the rate at which attention is shifted from item to item is inferred from the slopes of RT × set size functions. Typically, search slopes are 5–10 msec/item for easy (parallel) search and 40–50 msec/item for difficult (serial) search (Wolfe, 1994). Thus, when attention is moved serially from one item to the next, it is generally estimated that attention switches every 50 msec (e.g., Treisman & Gelade, 1980). Moreover, there is good evidence that even the slowest of visual searches are quite fast. An exhaustive literature review in which 2,500 sessions involving 1,000,000 trials over a 10-year period were examined indicated that all slopes are less than 150 msec/item (Wolfe, 1998).

Although indirectly estimating the rate at which attention shifts from one item to the next on the basis of search slopes is common in the visual search literature, an alternative paradigm developed by Duncan, Ward, and Shapiro (1994; Moore, Egeth, Berglan, & Luck, 1996; Ward, Duncan, & Shapiro, 1997) provided more direct estimates of the attentional switch times of 200 msec (Moore et al., 1996) to 500 msec (Ward et al., 1997) per item. In this so-called dwell time paradigm, observers

had to identify two objects presented in close temporal separation. The critical measurement in the dwell time paradigm is how long the first object continues to interfere with the second object. This is thought to represent the time course of the first object's attentional demand, because in order to identify the first object, attention must be engaged at the location of the first object. Performance (percentage correct) on the second target is assessed as a function of the stimulus onset asynchrony (SOA) between the first and the second targets, and this gives an estimate of the attentional dwell time. The results of Duncan et al. (1994) have shown that the interference of the first target with the second target lasts up to 450–500 msec, suggesting that attention moves from one object to the next as much as 10 times more slowly than typical estimates from visual search studies. Moore et al. (1996), using the same paradigm, have shown that this dwell time may be reduced to about 200 msec/item when the targets are not masked (as was the case in Duncan et al., 1994).

Clearly, there is a large discrepancy in the attentional dwell time estimates obtained from the two paradigms. On the basis of search slopes from traditional visual search experiments, it can be estimated that attention switches from one object to the next in about 50 msec/item (with an absolute maximum of 150 msec/item; see Wolfe, 1998), whereas in attentional dwell time studies, identification accuracy of the second target indicates an absolute minimum of 200 msec/item (see Moore et al., 1996). This is not a trivial difference, since the estimates of the speed of attention obtained from dwell time studies have been used as evidence against high-speed serial

---

models of attention (see Duncan et al., 1994; Ward et al., 1997). In this argument, it is implicitly assumed that the estimates obtained from dwell time studies can be applied directly to visual search. However, the dwell time paradigm is quite different from the typical visual search paradigm. The most important difference is that the dwell time paradigm is really a variant of the attentional blink paradigm (i.e., two visual targets, appearing in close temporal proximity, that need to be identified). It is well known that in the attentional blink paradigm, the operation of attention in working memory, but not of attention in perception, plays a key role (e.g., Shapiro, Driver, Ward, & Sorensen, 1997; Vogel, Luck, & Shapiro, 1998). This implies that the speed of attention shown in the dwell time studies may be quite different from that in visual search because "nonperceptual" processes, such as access to working memory, play a key role.

It is worth noting that the usefulness of examining attention dwell times is not limited to determining why different paradigms produce different results. In a broader sense, it is important to obtain independent and direct estimates of the speed of visual attention, because it is a crucial factor in all theories of visual attention. Particularly, theories such as those of Treisman (Treisman & Sato, 1990) and Wolfe (1994) rely heavily on the assumption that a high-speed serial search of 25–75 msec is possible. In addition, speed of attention plays an important role in interpreting findings from a variety of paradigms, such as cuing, attentional capture, visual search, change blindness, and RSVP, and it has implications for neurophysiological models of attention (see Ward, 1999, for an overview).

In the present study, we measured the deployment of attention directly by means of probe RT. The task was designed so that observers had to switch attention serially from one location to the next. Figure 1 presents the sequence of events on a given trial. In this example, the participants first shifted attention to the upper right location. The arrow presented at this location indicated the location that needed to be attended next (in this example, the bottom right location). The participants made a nonspeeded discrimination response to the orientation of the letter E. On some trials, at different SOAs, a probe dot was presented at one of the locations, and the participants had to make a speeded response to the probe dot. Earlier studies have shown that the speed with which observers respond to a probe is directly related to the amount of attention allocated to that location (see, e.g., Cave & Zimmerman, 1997). By presenting a probe at different locations (e.g., first and second attended locations) at different SOAs, we are able to map out how attention moves in time and space. As long as attention is allocated to the first attended location, we expect the fastest RTs for probes presented at this location. As soon as attention moves to the second location, RTs to a probe presented at the first location should go up, while RTs to a probe presented at the second location should go down. The SOA at which this occurs (i.e., probe RT to first attended location goes up and probe RT to second attended location goes down) represents an estimate of the attentional dwell time for the location first attended. To ensure that our dwell time estimation represented covert attentional allocation, we recorded eye movements as a fixation control. Trials in which the eyes moved were excluded from the analysis.

The present experiment was designed to provide a direct measure of the serial deployment of attention in a paradigm, rather than the interference measure from dwell time studies, and to have a strong resemblance to visual search. As in visual search, all the items were presented at the same time, to allow a faster engagement and disengagement of attention. It has been argued that attentional dwell estimates from dwell time studies may have been overestimated because the dwell time paradigm requires sequential presentation of two targets (Moore et al., 1996). If this argument holds, the present experiment, in which all the search items were presented simultaneously, should show much shorter dwell times. In addi-
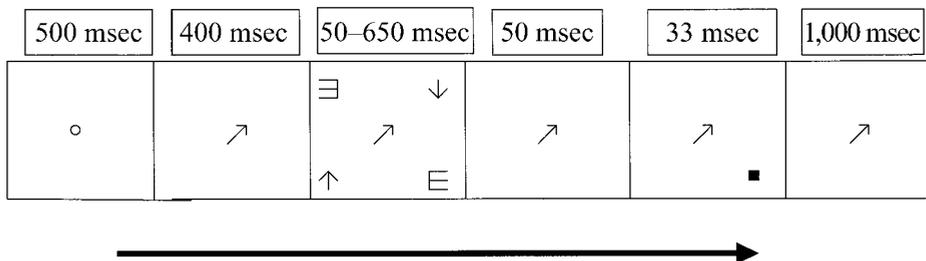

Figure 1. Sequence of frames on a given probe trial. After 500 msec, an arrow was presented, pointing to a location to which attention had to be shifted. After another 400 msec, the target display was presented. An arrow at the location that was attended first pointed to the location that had to be attended next. The letter presented at this to-be-attended location had to be discriminated (E or reversed-E). On half of the trials, a small square probe was flashed for 33 msec. The participants made a speeded response to the probe. Stimulus onset asynchronies between the target display and the probe were either 100, 200, 350, 500, or 700 msec. At the end of the trial, the participants made a nonspeeded discrimination to the target letter (E or reversed-E).

tion, as is assumed by serial search models (cf. Wolfe, 1994), we ensured that attention moved from one item to the next in a serial fashion by structuring the task so that spatial attention had to be allocated to one location before it could move on to the next location. The consequence of this control procedure was that, unlike in visual search, attention was very much driven in a top-down fashion. In standard visual search, it is possible to examine the next location to which attention should switch while encoding the current item. In our paradigm, this was not possible, because the encoding of the first item determined which item had to be attended next. Thus, our new procedure provided a more accurate estimate of endogenous, or volitional, shifts of attention than are possible with typical visual search tasks.

METHOD

Participants

Seventeen participants with normal or corrected-to-normal vision took part in the experiment.

Apparatus

A Pentium II computer with a 21-in. color monitor controlled the timing of the events and generated stimuli. Eye movements were recorded by means of an Eyelink tracker with a 250-Hz temporal resolution and a 0.2° spatial resolution.

Task and Stimuli

The basic trial sequence is shown in Figure 1. At the start of each trial, the participants fixated a central fixation point and pressed a key to initiate the trial. After 500 msec, the fixation point was removed, and a central arrow (length, 0.4°) appeared for 400 msec, pointing to one of four locations at which a target arrow would appear. The target display was then presented, consisting of two peripheral arrows (length, 0.4°) and two letters (0.4° × 0.4°) presented at the four corners of an imaginary square centered around the fixation point

(eccentricity, 6.4°). The target arrow to which attention was shifted first pointed to the location that had to be attended next. The letter at the next-to-be-attended location was the target letter (E or reversed-E) that had to be discriminated. The other two locations were filled by a distractor arrow and a distractor letter (E or reversed-E). The target display was removed after 50, 150, 300, 450, or 650 msec. On half of the trials, a small square probe (0.2°) was flashed for 33 msec at one of the four locations 50 msec after the removal of the target display. This resulted in an SOA between the target display and the probe of 100, 200, 350, 500, or 700 msec. All the stimuli were black on a blue background.

Design and Procedure

When a probe was presented, the participants were required to press the space bar as quickly as possible. At the end of each trial, they were required to make a nonspeeded discrimination response to the identity of the target letter. If it was an E, they had to press the "z" key; if it was a reversed-E, they had to press the "x" key. They were not allowed to make eye movements during the trials. After a practice block of 80 trials, the participants performed nine blocks of 80 trials. All the conditions were randomized within blocks.

RESULTS

Discarded Data and Errors

Trials with saccades (3.1%) were discarded. For the probe RT analyses, trials with RTs shorter than 200 msec (0.5%) and trials with RTs longer than 800 msec or no response (3.8%) were excluded. On 3.0% of the trials on which no probe was presented, a false alarm response was given.

Probe RTs

An analysis of variance (ANOVA) was conducted on mean probe RT, with SOA (between target display and probe: 100, 200, 350, 500, or 700 msec) and probe loca-
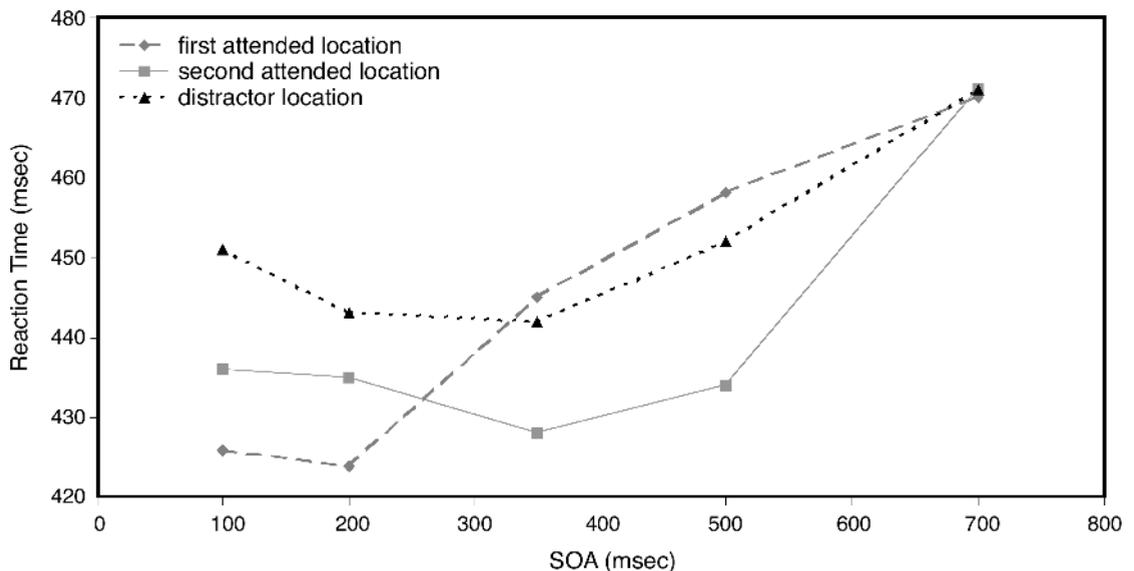

Figure 2. Mean reaction times for a probe detected at the first attended location, the second attended location, and a nonattended (distractor) location.

tion (first attended location, second attended location, or distractor location) as factors (see Figure 2). We used the distractor location as our baseline. It is fair to assume that attention was not directed at this location, regardless of the speed of attention. We conducted all tests relative to this "neutral" baseline. There was a main effect of SOA [$F(4,64) = 9.59$, $p < .001$], a main effect of probe location [$F(2,32) = 4.61$, $p < .02$], and an interaction between SOA and probe location [$F(8,128) = 3.35$, $p < .005$]. At the 100- and 200-msec SOAs, probe RTs were shorter at the first attended location than at the distractor location [$t(16) = 2.99$, $p < .01$, and $t(16) = 2.46$, $p < .03$, respectively], but probe RTs did not differ between the second attended location and the distractor location [$t(16) = 1.64$, $p > .10$, and $t(16) = 1.14$, $p > .25$, respectively]. At the 350- and 500-msec SOAs, the pattern was reversed: Probe RTs did not significantly differ between the first attended location and the distractor location (ts < 1), but probe RTs were shorter at the second attended location than at the distractor location [$t(16) = 2.59$, $p < .02$, and $t(16) = 2.30$, $p < .04$, respectively]. At the 700-msec SOA, probe RTs at the distractor location did not differ from those at the first or the second attended location (ts < 1). Note that the differences in probe RT between the first and the second attended locations at SOAs of 100 and 200 msec failed to reach significance. When we collapsed over SOAs of 100 and 200 msec, the difference between the first and the second attended locations was marginally significant ($p = .08$).

Discrimination Accuracy

An ANOVA was conducted on discrimination accuracy, with SOA and probe (present or absent) as factors. There was a main effect of SOA [$F(4,64) = 52.92$, $p < .001$], suggesting that with increasing exposure duration, discrimination accuracy of the target letter increased. There was also a main effect of probe [$F(1,16) = 45.28$, $p < .001$], indicating that the participants were slightly better at discriminating the target letter than they were when they did not have to do the probe RT task (mean accuracy of 83.4% vs. 77.5%). There was no interaction between SOA and probe [$F(4,64) < 1$]. An ANOVA for target-present trials showed only a main effect of SOA [$F(4,64) = 46.30$, $p < .001$], with better discrimination accuracy with increasing SOA. There was no effect of probe location, nor did probe location interact with SOA, suggesting that differences in probe RT for the different locations (attended first, attended second, and distractor locations) were not due to a speed–accuracy tradeoff.

DISCUSSION

On the basis of the present experiment, we estimate a dwell time of approximately 250 msec (the point at which the first-attended and the second-attended location probe RTs cross over in Figure 2). The probe RT data provide a direct measure of attentional allocation and show that at the short SOAs of 100 and 200 msec, at-

tention is at the first target location, whereas at the long SOAs of 350 and 500 msec, attention has been switched to the second target location.[1] Even though our paradigm is not similar to those in previous dwell time studies (i.e., it is unlike an attentional blink paradigm and does not involve working memory), our estimation is close to those obtained by Moore et al. (1996) in their *easy* condition. The speculation of Moore at al. that dwell times would go down even further if all items were presented simultaneously does not hold, since simultaneous presentation of all items did not result in any reduction in the dwell time estimate. It is important to note that unlike in earlier studies, the dwell time measured in the present study was due solely to covert shifts of attention, because trials in which eye movements (i.e., overt shifts of attention) were made were excluded from the analysis.

The present study provides converging evidence for the notion that attention moves relatively slowly. Our estimation is in line with the lower estimates based on attentional dwell time studies but remains at odds with the much faster estimates based on the search slopes of visual search experiments. Even though a dwell time of 250 msec seems rather long, studies measuring brain activity during visual search have arrived at similar estimates. For example, Woodman and Luck (1999) measured the N2pc during serial search and found that a shift in the N2pc required about 150 msec. It has been assumed that this time reflects the speed of serial reallocation of attention. Also, on the basis of the steady-state visual evoked potential to a target, Müller, Teder-Sälejärvi, and Hillyard (1998) concluded that attention must have shifted slowly, approximately 450–650 msec.

Overall, then, what is the time it takes to shift attention from location to location? On the basis of dwell time studies and the present probe detection study, one would conclude that attention moves relatively slowly, with a speed of approximately 200–250 msec. Obviously, such an estimation cannot be consistent with models that assume fast serial allocation of attention, such as those of Wolfe (1994) and Treisman and Gelade (1980; Treisman & Sato, 1990). The only way to accommodate fast serial search slopes of 25–75 msec/item is to assume that processing is not serial but (partly) parallel. This is a plausible explanation, because it is well known that limited-capacity parallel models can easily mimic serial search behavior when searching for a target among multiple nontargets (e.g., Bundesen, 1990; Townsend & Ashby, 1983).

On the basis of an extensive literature review, Ward (1999) also argued in favor of parallel models of attention to explain fast serial search slopes. According to Ward's review, it is quite conceivable that attention moves slowly; in fact, he claimed it moves as slowly as eye movements do (indeed, his title is "visual attention moves no faster than the eyes"). This comparison with eye movements is striking. Indeed, there is a consensus that fixation durations of the eyes are around 200–250 msec, a measure that is very similar to the estimation of dwell times.

Moreover, there is a strong positive correlation between overall RT and the number of fixations in visual search (e.g., Shen, Reingold, & Pomplun, 2000). However, recent studies have shown that fixation durations of the eye may be as short as 50–75 msec after an exogenous oculomotor shift (e.g., Godijn & Theeuwes, 2003; Theeuwes, Kramer, Hahn, & Irwin, 1998). This finding, and the similarity to the speed of eye movements, may make it possible to accommodate the differences in dwell time estimations. As was noted before, in dwell time studies and also in our probe RT task, attention moves very much in a top-down controlled fashion involving an endogenous attentional shift, endogenous processing, and an endogenous disengagement of attention, which is very similar to top-down controlled eye movements. However, in visual search, search is much less controlled and very much driven by the bottom-up salience of the items in the display. Indeed, Wolfe, Alvarez, and Horowitz (2000) have provided evidence that attention can move much faster when allowed to move randomly than when controlled by top-down factors. Overall, it appears that attention and eye movements may move slowly from one object to the next when such a movement is completely controlled in a top-down fashion, whereas both may move much faster when engagement and disengagement are controlled in a bottom-up fashion.

REFERENCES

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, **97**, 523-547.

Cave, K. R., & Zimmerman, J. M. (1997). Flexibility in spatial attention before and after practice. *Psychological Science*, **8**, 399-403.

Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature*, **369**, 313-315.

Godijn, R., & Theeuwes, J. (2003). Programming of endogenous and exogenous saccades: Evidence for a competitive integration model. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 1039-1054.

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.

Müller, M. M., Teder-Sälejärvi, W., & Hillyard, S. A. (1998). The time course of cortical facilitation during cued shifts of spatial attention. *Nature Neuroscience*, **1**, 631-634.

Shapiro, K. L., Driver, J., Ward, R., & Sorensen, R. E. (1997). Priming from the attentional blink: A failure to extract visual tokens but not visual types. *Psychological Science*, **8**, 95-100.

Shen, J., Reingold, E. M., & Pomplun, M. (2000). Distractor ratio influences patterns of eye movements during visual search. *Perception*, **29**, 241-250.

Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our eyes do not always go where we want them to go: Capture of the eyes by new objects. *Psychological Science*, **9**, 379-385.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge: Cambridge University Press.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, **12**, 97-136.

Treisman, A. M., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 459-478.

Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological evidence for a post-perceptual locus of suppression during the attentional blink. *Journal of Experimentalp Psychology: Human Perception & Performance*, **24**, 1656-1674.

Ward, R. (1999). Visual attention is no faster than the eyes. In K. Shapiro (Ed.), *The limits of attention: Temporal constraints on human information processing* (pp. 199-216). Oxford: Oxford University Press.

Ward, R., Duncan, J., & Shapiro, K. (1997). Effects of similarity, difficulty, and nontarget presentation on the time course of visual attention. *Perception & Psychophysics*, **59**, 593-600.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, **1**, 202-238.

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, **9**, 33-39.

Wolfe, J. M., Alvarez, G. A., & Horowitz, T. S. (2000). Attention is fast but volition is slow. *Nature*, **406**, 691.

Woodman, G. F., & Luck, S. J. (1999). Electrophysiological measurement of rapid shifts of attention during visual search. *Nature*, **400**, 867-869.

NOTE

1. The reason that the difference between the first and the second attended location is only marginally significant is that, on a proportion of trials, attention may already be at the "second attended location" at the 100- or 200-msec SOA. This would reduce the probe RTs at the "second attended location" and increase probe RTs at the "first attended location."