# Attentional pursuit is faster than attentional saccade

**Todd S. Horowitz** Visual Attention Laboratory, Brigham & Women's Hospital, & Department of Ophthalmology, Harvard Medical School, Boston, MA, USA

**Alex O. Holcombe** School of Psychology, Cardiff University, Cardiff, Wales, United Kingdom

**Jeremy M. Wolfe** Visual Attention Laboratory, Brigham & Women's Hospital, & Department of Ophthalmology, Harvard Medical School, Boston, MA, USA

**Helga C. Arsenio** Center for Women's Mental Health, Massachusetts General Hospital, Boston, MA, USA

**Jennifer S. DiMase** Department of Psychology, Yale University, New Haven, CT, USA

How quickly can we shift the focus of visual attention? We compared the rates of two types of attentional shifts: *attentional saccades* (shifts between objects) and *attentional pursuit* (shifts along with a moving object). Instead of measuring the time required for a single shift, which confounds shift time with cue interpretation time, we measured the pace at which observers could make multiple successive shifts in a predictable order. We find that successive attentional saccades between objects are quite slow (300-500 ms). The object-based theory of attention predicts that attention should shift between locations more quickly when in pursuit of a moving object. Our results support this theory. Attentional pursuit is substantially faster - taking only 200-250 ms to cover the same distance. "Indexing" a moving object (keeping track of one object) can be done at even faster rates, supporting a distinction between attending to and indexing objects.

Keywords: attention, attentive tracking, apparent motion

## Introduction

How quickly can we shift the focus of visual attention? This is a fundamental issue in the study of attention. Estimates of the rate at which attention can shift are central to the debate over serial and parallel processing (Duncan, Ward, & Shapiro, 1994; Moore, Egeth, Berglan, & Luck, 1996), the relationship between overt and covert attention (Mackeben & Nakayama, 1993), and have been used to argue for multiple attentional subsystems (Cheal & Lyon, 1991; Nakayama & Mackeben, 1989). Here we find different rates for two distinct types of attention shifts: *attentional saccades*, when attention is commanded to shift between objects, and *attentional pursuit*, when attention can shift along with a moving object. The use of the terms "saccade" and "pursuit" underlines the obvious similarity between these shifts of attention and analogous shifts of the eyes. It may be that the types of attentional shifts and the corresponding eye movements share common mechanisms (Khurana & Kowler, 1987; Kowler, Anderson, Dosher, & Blaser, 1995). However, the distinction between these shifts of different rates stands even if attentional "saccades" and "pursuit" do not turn out to share mechanisms with the corresponding eye movements.

### Attentional saccades: How fast can attention shift from object to object?

Previously, the attentional shifting rate has been estimated from three types of experiments: orienting experiments, attentional reaction time experiments, and visual search experiments. All three methods require observers to shift their attention from one object to another, "attentional saccades" in our terminology. Use of the orienting method is probably most common (Posner, 1980). Observers respond to a target whose location is varied. A precue indicates the target location with varying reliability and the cue to target interval is varied. The function relating performance to cue-target interval is used to estimate the time required to shift attention. Results from this literature have motivated a distinction between two kinds of attentional saccades, sometimes termed endogenous and exogenous, sometimes central and peripheral. Essentially, they are defined by the type of cues that elicit them. Exogenous (or peripheral) cues are presented at the target location and are

doi:10.1167/4.7.6 Received November 18, 2003; published July 20, 2004 ISSN 1534-7362 © 2004 ARVO

said to automatically draw attention (Jonides, 1981). Endogenous (or central) cues are usually presented at fixation and symbolically indicate the cued location; for example, an arrow that points to the to-be-attended location. Exogenous and endogenous conditions yield substantially different estimates. Exogenous shifts can be completed in less than 100 ms, whereas endogenous shifts require more than 300 ms (Cheal & Lyon, 1991; Nakayama & Mackeben, 1989). Mackeben and Nakayama (1993) have also shown evidence for "express attentional shifts," which they explicitly relate to express saccades, which require only 33 ms.

A different logic underlies the attentional gating method (Reeves & Sperling, 1986). In this method, streams of rapidly changing stimuli are presented in two different locations. Observers monitor one stream for a trigger stimulus (such as a "2" in a stream of digits) and then shift their attention to a second stream, reporting the first stimulus they detect in the second stream (e.g., the first letter). The time to shift attention can be inferred from the time between the presentations of the trigger and of the reported stimulus. Estimates of the attention shift time from this paradigm are typically in the 300-ms range (Klein & Dick, 2002; Reeves & Sperling, 1986; Sperling & Weichselgartner, 1995). If there is a single stream, and the trigger stimulus is a bright frame around the stream, observers can often report the letter presented simultaneously with the trigger, indicating that the "attentional gate" can be opened almost instantly (Weichselgartner & Sperling, 1987). Therefore, the additional 300-ms delay when shifting from one stream to another is clearly a cost for shifting the locus of attention. Furthermore, Shih and Sperling (2002) argue that this cost is independent of the distance between the two streams; the 300-ms cost can be interpreted as the time necessary to shift the locus of attention, rather than the time for attention to cover a particular distance in some analog fashion.

Both of the methods reviewed above are intended to measure a single attentional saccade. However, it is important to realize that the time needed to detect the cue and to determine the next locus of attention is incorporated into the estimate of the shift time. Hence, it is unsurprising that using shift cues which take longer to interpret results in longer estimates of shift times (Reeves & Sperling, 1986). Cue interpretation time can be reduced by using exogenous cues, but it is uncertain whether it is eliminated. Another issue is that processing the cue may interfere with processing the target, resulting in further inflation of the time needed to shift attention (Peterson & Juola, 2000).

To eliminate the contribution of cue interpretation time to the estimate, one can measure the total time required for multiple attentional shifts that are elicited by a single cue. If the number of shifts is varied, the added time per shift should reflect the shift time alone. This logic has been used to estimate the shift time from the results of visual search experiments.

In visual search experiments, observers are presented with an array of multiple items and asked to discriminate some aspect of the target item (e.g., its color, identity, presence, or absence). The number of items (the set size) is varied, which typically results in a linear relationship between set size and reaction time (RT). To estimate the attentional shift time, researchers assume that items are processed serially, with a single attentional saccade after each item is processed. A further assumption is that other mental processes, including initial processing of the display, decision, and motor processes, do not interact with the number of items. The increase in RT with a one-item increase in set size (the slope) is then the rate of attentional deployment. Typical estimates are on the order of 50 ms/item (e.g., Wolfe, 1998).

However, the assumption that the stimuli are processed serially may not be warranted. Results of visual search experiments may be explained just as well with the assumption that multiple items are processed in parallel (Eckstein, Thomas, Palmer, & Shimozaki, 2000; McElree & Carrasco, 1999; Palmer, Verghese, & Pavel, 2000). It has even been suggested that individual items may be processed more than once during search (Horowitz & Wolfe, 1998; Horowitz & Wolfe, 2001, 2003), further complicating this estimate.

Not only the number but also the nature of attentional saccades in search is undetermined. Recall that researchers estimated much slower shift times with endogenous cues than with exogenous cues. Which of these cueing situations is more like the situation during visual search? Because there are no external events to capture attention during visual search, all shifts might be considered endogenous. On the other hand, attention presumably shifts to the most salient items in the field and in this sense may be stimulus-driven or exogenous (Briand, 1998; Briand & Klein, 1987; Klein, 1988).

To investigate the rate of attentional saccades and their role in visual search, Horowitz, Wolfe, and Alvarez (2004; see also Wolfe, Alvarez, & Horowitz, 2000) sought to bridge the orienting and visual search paradigms. Their variant of the search task measured the rate of attentional saccades under a "commanded" condition and an "anarchic" condition. The stimulus was a series of eight frames, each consisting of a circular array of eight letters. The letters changed on each frame. In commanded search, observers were required to step their attention around the circle with each frame, stepping to the next, *N*th location each time a new frame appeared. The target was presented only once. It appeared in the *N*th location on the *N*th frame, with *N* varying from trial to trial. Thus, if the observer had been successfully shifting attention by one step on each frame, the target would appear at the attended location. If the observer had not been able to deploy attention as commanded, the target would be elsewhere and the post-mask would prevent perception of it. A staircase procedure was used to determine the maximum frame rate that still allowed reasonable accuracy (66.7% in a condition where chance performance was 25%). The anarchic condition was similar, except that a target, varying in position across trials, was present on each frame and observers were told to search for the target but not instructed to shift attention in any particular fashion. In this way, the anarchic condition was like a conventional visual search task. The anarchic procedure produced an estimate of 84 ms for each anarchic deployment and shift. The commanded search procedure produced an estimate of 274 ms for each attentional saccade.

These slow shifts in the command condition are comparable to those in endogenous cueing studies. However, the commanded search task requires observers to identify a letter on each frame, and therefore the shift duration estimate includes the time to identify letters, rather than just the time to shift attention. In the present experiments, we seek converging evidence for the rate of attentional saccades using tasks that do not require that an act of identification accompany each shift of attention. We compare attentional saccade rate to the rate at which attention can track an object through space (attentional pursuit).

## Attentional pursuit: How fast can attention follow an object?

A substantial body of data supports the theory that attention can select objects (Chen, 2000; Davis, Welch, Holmes, & Shepherd, 2001; Tipper, Jordan, & Weaver, 1999; Tipper, Weaver, Jerreat, & Burak, 1994; Wolfe & Bennett, 1997), in addition to locations (Egly, Driver, & Rafal, 1994; but see Tsal & Lamy, 2000). Here we point out a novel prediction of that theory. When a selected object moves, the focus of attention should move with the object to the new location. It would not need to be selected again in its new location, having never been unselected. It follows that the rate that limits attentional saccades might not limit attentional pursuit. Attention might, therefore, move with an object from point A to point B more quickly than it can be deployed from an object at A to and then to an object at B. One way to test this hypothesis is to use a task where the motion is ambiguous unless attention remains attached to specific objects.

Verstraten, Cavanagh, and Labianca (2000), developed such a method. Observers are presented with a counterphase motion display (Cavanagh, 1992; Wertheimer, 1912) that consists of two alternating frames of objects arranged in a circle (Figure 1). Objects (small disks in this case) are displaced from one frame to the other in such a way that apparent motion can be seen either clockwise or counterclockwise, according to the bias of the observer. Observers were instructed to perceive motion in one direction and to use attention to track a single disk as it moved. After the motion stopped, observers indicated which disk they had been tracking. The velocity of the disks was adjusted using both staircase and method of adjustment techniques to determine the maximum velocity at which observers could reliably track. Because observers were asked to fixate, and the disks were identical, observers could only distinguish the target disk from the others by successfully pursuing the target with covert attention.

Verstraten et al. (2000) were primarily concerned with determining the conditions under which one could perceive motion while attentively tracking. As a result, their data are expressed in terms of the flicker rate of the stimuli, rather than the attentional shifting rate. However, a flicker rate corresponds to a particular rate at which attention must shift to follow the motion. With a two-frame stimulus, one cycle corresponds to two attentional fixations. In their Experiment 1, using the counterphase motion stimulus, they found an upper limit of 7Hz for the flicker rate. This means that a two-frame cycle took 143 ms; and, therefore, attention was able to move every 71 ms, roughly three times faster than the rate of attentional saccadic shifts reported by Wolfe et al. (2000).

Even faster estimates can be obtained from the Verstraten et al. (2000) Experiment 2, which employed an unambiguous apparent motion display with four disks presented in each of the frames. The independent variable is the number of intervening positions that a single disk would occupy in traveling one quarter of the way around the circle (90°). Because there are four disks, every 90° rotation returns the display to its initial configuration, and constitutes one cycle. So the inverse of the flicker rate is the time taken to travel 90°. Figure 6 in Verstraten et al. (2000) indicates that when the number of intervening positions is eight, the maximum flicker rate at which observers could attentively track was 8Hz, corresponding to 125 ms per 90° cycle. Because there were eight intervening positions, if attention was actually directed to each intervening position, the effective shifting rate is 125/8, or 15.6 ms/step. This is approximately an order of magnitude faster than the rates reported in Wolfe et al. (2000).

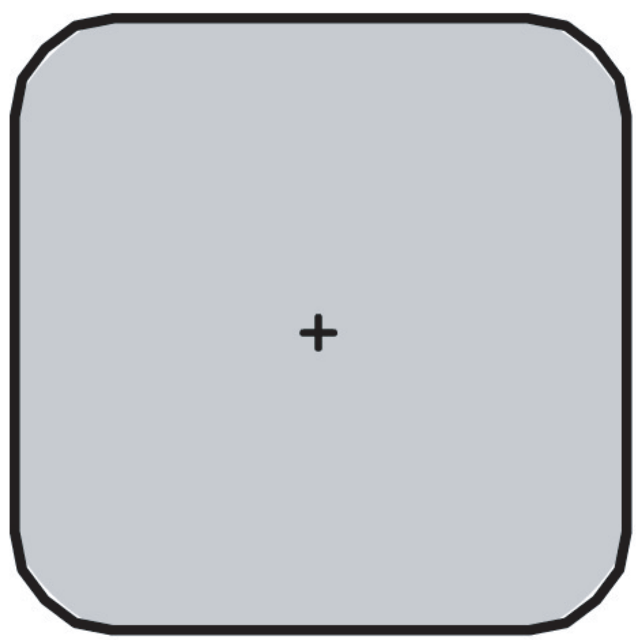

Figure 1. An animated cartoon to illustrate the attentional saccade condition.

Verstraten et al. (2000) argue that breakdown of attentional pursuit in their display was not caused by a shifting limit, but rather a limit to the temporal resolution of attention. Tracking failed when disks were repeatedly presented in the same location more than 5 or 6 times a second. It seems that in such circumstances, a disk groups with the next disk presented in the same location, preventing apparent motion.

Thus, the results of Verstraten et al. (2000) support the notion that attentional pursuit of a moving object, in this case a circle in apparent motion, is faster than attentional saccades between objects. However, there are three reasons to be cautious about this conclusion. First, as noted above, the attentional saccade method of Wolfe et al. (2000) may underestimate the maximum rate of attention shifts between objects because it requires identification of a letter with each shift. Second, Verstraten et al. (2000) appear to have tested only experienced psychophysicists, whereas the Wolfe et al. (2000) observers did not have extensive training. Differences in expertise may have contributed to the difference in estimates. Third, and most significantly, it is quite possible that the attentional pursuit task of Verstraten et al. (2000) did not require attention, in the classical sense, to each location along the object's trajectory.

Verstraten et al. (2000) observers were asked to report which disk they had been tracking. This tests object continuity: the ability to indicate which object at the end of the trial corresponds to which object at the beginning of the trial. There are two prevailing theories about how the visual system maintains object continuity. According to Pylyshyn and his colleagues (Pylyshyn, 1989; Schmidt, Fisher, & Pylyshyn, 1998; Scholl & Pylyshyn, 1999; Trick & Pylyshyn, 1993), object continuity is mediated by a "visual index" (or "FINST" – Finger of INSTantiation), which functions much as a pointer does in computer science. The visual index is assigned to an object, and then attention can later be directed to the object by querying the index. The index itself carries no information about the object except its current location, which presumably is constantly updated by low-level motion processes.

Kahneman and Treisman (1984; Kahneman, Treisman, & Gibbs, 1992) suggested that when an object is attended, a representation called an "object file" is opened, which mediates between the visual representation and memory. When attention is redirected to that same object at a later time, object files previously associated with that item are retrieved, a process termed "impletion." Object file theory is similar to visual index theory in that object continuity is mediated solely by the spatiotemporal history of the object. However, there are a number of differences in the treatment of object continuity under visual index theory and object file theory. Critically, the object file contains a wealth of information about the featural and semantic qualities of the object, unlike the visual index.

According to both accounts, then, determining object continuity does not actually require continuously attending to the object from start to finish. Instead, the task can be accomplished either by assigning a FINST or via impletion at the end of the motion sequence. Note that Pylyshyn (1989) explicitly distinguishes between "attending" to an object and "indexing" that object. Attention entails access to all of an object's properties (Kahneman et al., 1992; Treisman & Gelade, 1980), including its identity. Indexing is a narrower operation, entailing only the ability to distinguish an object on the basis of its spatiotemporal history. Using multiple object tracking and enumeration paradigms, Pylyshyn and his colleagues have argued that the visual system is capable of indexing four or five objects (Pylyshyn & Storm, 1988; Scholl, Pylyshyn, & Feldman, 2001; Trick & Pylyshyn, 1994). Similarly, Kahneman et al. (1992) proposed that a limited number (3–4) of object files could be open at any one time. While Pylyshyn's visual indexes (or "FINSTs") are often described as a form of attention, Pylyshyn has been careful to note that an indexed object is not necessarily an attended object in the conventional sense. In support of this distinction, Scholl, Pylyshyn, and Franconeri (2004) found that although observers report the motion and location of tracked objects accurately, reports of other featural properties such as color or shape are inaccurate. Because Verstraten et al. (2000) required their observers to report which disk they had been tracking, rather than a featural property of the target disk, their results may reflect limits on object continuity rather than attention per se.

## Experiment 1

Our method of measuring the time required for different attentional shifts was designed with two goals in mind:

1. To use similar stimuli and similar methods to measure attentional saccades and attentional pursuit.

2. To avoid the problematic assumptions of previous methods.

The procedure directly compares the movement of attention while pursuing a moving object with shifts of attention between objects. In the attentional saccade condition, observers view a display of 12 circular placeholders, arrayed in a circle around the fixation point (see Figure 1). The task is to step attention from placeholder to placeholder around the circle, in a specified direction and at a rate paced by a regular series of tones. Because the tones occur at regular intervals, observers could anticipate the cue and shift attention without taking the time to process each tone. In other words, once observers began moving their attention at the right pace, they could simply continue to do so, minimizing the role of cue processing time. At a randomly chosen time after the tone had sounded between 12 and 24 times, six probe letters were briefly presented, then masked. The observer was asked to report which letter was presented in the attended placeholder.

In contrast to the attentional saccade condition, the attentional pursuit condition was designed to reveal the temporal limit for shifting attention with objects, rather than between them. In this condition, the 12 locations were split into two sets of six placeholders. These were presented in two alternating frames (see Figure 2). Such a display can be seen rotating either clockwise or counterclockwise. The direction can be governed by the intention of the observer (Shioiri, Yamamoto, Kageyama, & Yaguchi, 2002; Verstraten et al., 2000). Observers were asked to use their attention to follow a particular object moving either clockwise or counterclockwise. Six probe letters were then briefly presented just as in the command condition, and the observer reported which letter appeared in the attended placeholder. Success on this task indicated that observers did have access to the features in the critical location, as opposed to merely indexing the location as might have been possible in the Verstraten version of the task.

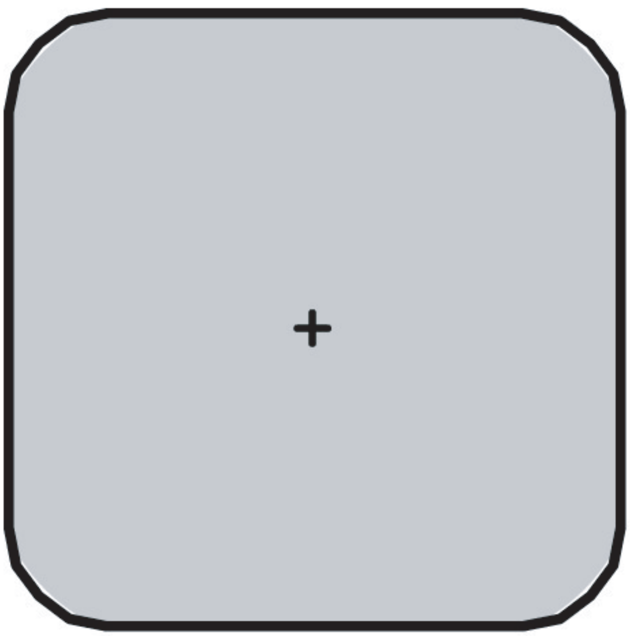

Figure 2. An animated cartoon to illustrate the attentional pursuit condition.

It might be argued that attentional pursuit might better be measured using a continuous stimulus, such as a sinusoidal radial grating (Cavanagh, 1992). However, our concern here was to match the attentional saccade and attentional pursuit conditions as closely as possible.

To determine the fastest rate at which observers could shift their attention, the tone rate in the attentional saccade condition and the frame alternation rate in the attentional pursuit condition were independently staircased to achieve 66.7% accuracy. In both conditions, the measure of successful shifting of attention at the intended rate was the identification of the target letter. Note that observers do not need to identify a stimulus with *each* shift of attention in either condition. The critical difference between conditions is that, in the attentional saccade condition, observers had to shift attention from one object to the next, whereas in the attentional pursuit condition, the observer could track an object in apparent motion around the circle.

The number and spatial extent of attention shifts are the same in the attentional saccade and pursuit conditions. However, the attentional saccade condition shows 12 objects on the screen at any one time, compared to 6 in the attentional pursuit condition. It is possible that crowding or load effects might impair attention shifting in the attentional saccade condition. To address this we added a control condition, designated *attentional saccade-6*, which is identical to the attentional saccade condition except that only six places are occupied. To put it another way, the attentional saccade-6 display is identical to one frame of the attentional pursuit condition. If faster shifts are observed in the attentional pursuit condition than in the attentional saccade condition because of the reduced number of stimuli in the attentional pursuit condition, then the same faster shifts should also be observed in the attentional saccade-6 condition.

The second difference between attentional saccade and attentional pursuit conditions is that observers are provided with an auditory cue in the attentional saccade condition but not in the attentional pursuit condition. It may be that tones are particularly effective in triggering attentional shifts. However, note that this would work against our prediction that attentional saccade rates will be slower than pursuit rates.

To summarize, four aspects of the method are critical. First, the task involved multiple shifts of attention on each trial, reducing the proportional contribution of cognitive setup time and initial cue processing time. Second, observers were not required to process an object or object property on each shift. This reduced or eliminated the contribution of the time to identify an object. Third, after a succession of shifts, our probe task required form processing (letter identification). This assures that observers deployed attention to the object and did not merely index it as was possible in the Verstraten version of the pursuit task. Finally, our method experimentally specified the number of attentional shifts required of our observers, rather than requiring a post hoc inference of the number of shifts (as in visual search methods).

## Methods

### *Observers*

Nineteen observers from the paid observer panel of the Visual Attention Laboratory at Brigham & Women's Hospital in Boston participated in this experiment. All observers had normal or corrected-to-normal vision (20/25 or better) and passed the Ishihara test for color blindness. Observers gave informed consent and were compensated $10/hour for their time.

### *Apparatus & stimuli*

Stimuli were presented on 21 in. color monitors with a refresh rate of 75 Hz controlled by Power Macintosh computers running Matlab 5.1 (Mathworks) utilizing the Psychophysics Toolbox (Brainard, 1997) routines.

In the attentional saccade condition, the stimulus consisted of an array of 12 white placeholders, circles of diameter 2.12° visual angle, evenly spaced along an imaginary circle of radius 5.29° visual angle centered on fixation (see Figure 1). In the attentional pursuit condition, two frames were generated, each containing one set of six placeholders (see Figure 2). In the attentional saccade-6 condition, the stimulus consisted of one set of six placeholders, arranged as shown in the first frame of Figure 2. Probe stimuli were black 60-point Arial letters contained within 2.36° visual angle bounding boxes. The letters "I," "O," and "W" were omitted. Masks were constructed by superimposing the characters "+," "X," and "O" in 60-point Arial font. Masks were white. Cue disks were 1.41° visual angle in diameter and white. Stimuli were presented on a gray background.

### *Procedures*

In the attentional *saccade* condition, observers were instructed to step their attention around the circle of placeholders, beginning with the placeholder at the top of the display (the "12 o'clock" position). Observers initiated each trial with a key press. Each trial began with a presentation of the fixation cross (Arial 60-point "+" within a 2.36° bounding square) for 667 ms. Then the 12 placeholders appeared, with a cue disk in the 12 o'clock placeholder. The cue disk then moved either clockwise or counterclockwise for four steps, indicating both the direction and pace at which observers should move their attention. The direction was selected randomly on each trial. A 750-Hz, 50-ms tone was sounded with each successive presentation of the cue disk, and continued at the same pace after the cue disk disappeared. Observers were instructed to maintain fixation while shifting their attention at the indicated pace. After a random number of steps between 13 and 25 (that is, sometime during the second circuit of the display), the placeholder display was abruptly replaced with a probe frame, consisting of six randomly selected probe letters. One probe letter was presented at the location that the observer would have been attending to, had she been shifting her attention at the correct rate. The remaining probes were presented at alternating locations around the circle (see Figure 1). For example, if the observer should have been attending to the 4 o'clock location, probes were presented at all even-numbered locations. Probes were presented for 133.33 ms and then masked. The observer was then asked to report which letter was at the location she was attending to by pressing the appropriate key on the keyboard. The observer was then asked to confirm her response before going on to the next trial.

In the attentional *pursuit* condition, observers were told that they would see two alternating frames of six circles, which they could see as moving either clockwise or counterclockwise, as they chose. They were instructed that their task was to use their attention to follow the circle that started at the 12 o'clock position. As in the attentional saccade condition, a cue disk was initially presented in this circle. It then moved four steps either clockwise or counterclockwise to indicate the direction of motion that the observer should follow. Aside from the alternation of even clock positions with odd clock positions, the procedure in this condition was precisely the same as in the attentional saccade condition (see Figure 2).

The attentional saccade-6 condition was identical to the attentional saccade condition except that there were only six placeholders on the screen.

In all three conditions, the pace was controlled by a staircase designed to converge on the pace that yielded 66.7% correct. If the observer responded incorrectly, the duration of each interval increased by 26.67 ms. If the observer made two consecutive correct responses, the duration decreased by 13.33 ms. A staircase run began with a block of 20 practice trials. The initial duration of this practice block was set to 667 ms (1.5 Hz). The initial duration for the run of experimental trials was determined by computing the half-way point between the initial and final durations of the practice block. The experimental staircase was terminated after the observer had completed at least 100 trials and the staircase had undergone at least 12 reversals. Asymptotes were computed as the average duration over the last 10 reversals.

Each observer completed one run of each condition, and the order of conditions was counterbalanced. The experiment took roughly half an hour.

## Results

Three observers were eliminated because their staircases in at least one condition increased rather than decreased during the experiment, indicating an inability to do the task at all. The staircase procedure causes the presentation rate to converge on the fastest rate at which observers can shift their attention and perform the task with 66.7% accuracy. Average asymptotes from the remaining 16 observers are shown in Figure 3.

Planned comparisons indicate that the asymptotes in the attentional pursuit condition (mean of 199.1 ms/step) were significantly lower than in the attentional saccade (mean of 331.9 ms/step) ($t(15) = 2.66$, $p < .05$) or attentional saccade-6 (mean of 358.1 ms/step) ($t(15) = 3.34$, $p < .005$) conditions. Asymptotes did not differ between the two attentional saccade conditions ($t(15) = 1.07$, $p > .10$).

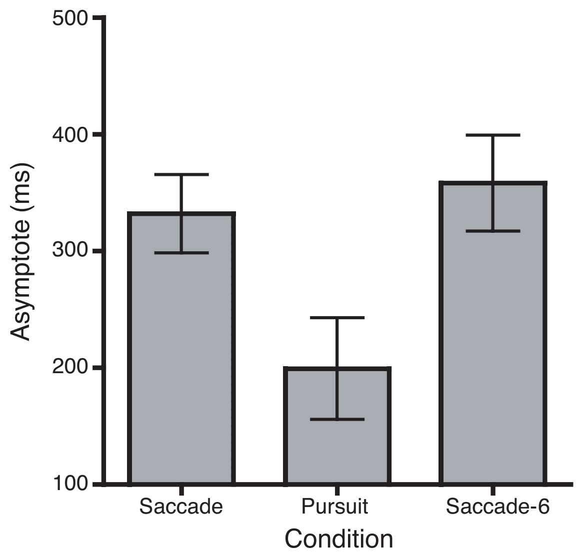


Figure 3. Asymptote data from Experiment 1. Asymptotes were computed as the average of the last 10 staircase reversals. Error bars in this and all subsequent figures indicate the *SEM*.

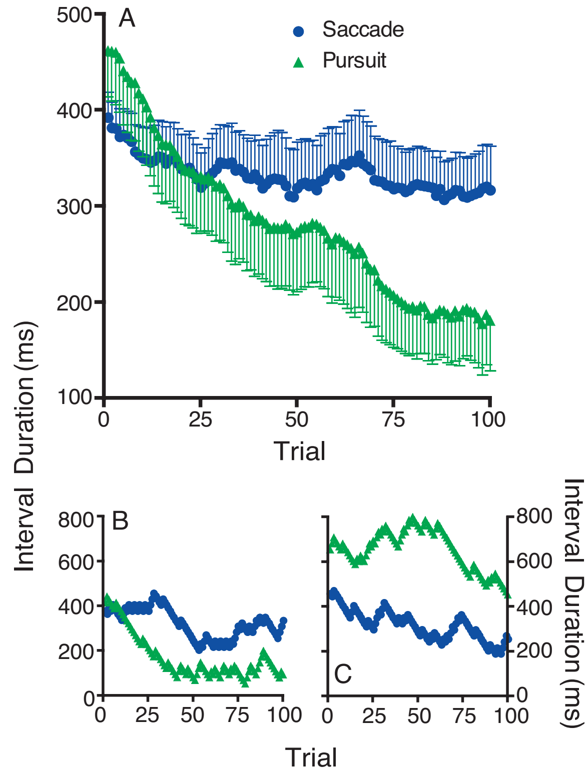


Figure 4. Staircase data from Experiment 1. Blue circles denote the attentional saccade condition, green triangles the attentional pursuit condition. Panel A shows the average staircase across all 16 observers in the attentional saccade and attentional pursuit. Panel B plots the staircases for observer MS, who performed well on the attentional pursuit task. Panel C plots the staircases for observer RM, who apparently had trouble with the attentional pursuit task.

## Discussion

The result in the attentional saccade condition establishes that even when observers do not have to identify a letter with each step, saccadic deployments of attention are quite slow, 331.9 ms per shift in this experiment. This is even slower than the 217 ms/step reported by Wolfe, Alvarez, and Horowitz (2000). The sluggish Wolfe et al. estimate might have been attributed to the need to identify a stimulus with each shift. However, the present result establishes that shifts are slow even without this requirement. Shifts of attention were substantially faster in the attentional pursuit condition. This advantage for pursuit was clearly not due to the smaller number of items on the screen, because asymptotes in the attentional saccade-6 condition were even slower than in the standard attentional saccade condition (though not sufficiently slow as to suggest a cost for distance).

Although the asymptotes in the attentional pursuit condition were faster than those of the attentional saccade condition, they were much slower than the rates found by Verstraten et al. (2000). At first glance, this seems to indicate a difference between the rate of indexing as measured by Verstraten et al. and the rate of selective attention shifts measured in the present experiments. However, Verstraten et al. also used experienced psychophysical observers, whereas our observers were comparatively psychophysically naïve, with no experience in tracking ambiguous motion stimuli. There is reason to believe that at least a few of our observers may have had difficulty in properly perceiving the counterphase stimuli. Figure 4 plots staircases from Experiment 1. Panel A shows the average staircase. Staircases for a typical observer (RS) are shown in Panel B. Panel C, however, shows staircases from an observer (RM) who appeared to have difficulty with the pursuit task. The data of three other observers behaved similarly. This may reflect natural variation in ability to rapidly shift attention, or it may reflect inexperience with these displays.

Another issue is that in our experiment, the probe task itself, identification of a brief, masked letter, may have been difficult for some observers. This would further reduce the rate estimates in comparison to those of Verstraten et al. Indeed, this may also explain why threshold rates in the attentional saccade condition were slower than those found in the comparable conditions of Wolfe et al. (2000). Experiment 2 was designed to address these problems.

# Experiment 2

In Experiment 2, we replicated Experiment 1, but added additional training procedures. The practice and screening procedure had two parts. First, we attempted to teach each subject to attentionally track an ambiguous motion stimulus.

Second, we practiced observers on the attentional saccade and attentional pursuit tasks at a very slow rate, with the cue disk continually present, so we were confident that they would have no trouble attending to the correct location at the time of the probe. The probe duration was the same as in the experiment proper. This procedure had two advantages. First, observers had ample opportunity to familiarize themselves with all aspects of the task without demand of rapid attention shifts. Second, we were able to determine each observer's ability to perform the probe task directly.

## Methods

### Observers

Twenty observers from the Visual Attention Laboratory's paid observer panel participated in this experiment.

### Apparatus & stimuli

The apparatus was the same as in Experiment 1. The main experiment and the masked letter identification training procedure used the same stimuli as in Experiment 1, except that the size of the placeholder circles was increased to 2.36° visual angle. The ambiguous motion training stimuli were generated and presented using Macromedia Flash™ 5. These stimuli consisted of a set of sixteen 1.78° black disks arranged in a circle with a radius of 3.81°. Four equally spaced disks had a quadrant deleted and served as inducers for an illusory Kanisza square. The placement of the inducers from frame to frame varied according to whether ambiguous or unambiguous motion was being presented, and, in the case of unambiguous motion, the direction of motion.

### Procedure

***Ambiguous motion training.*** The purpose of this procedure was to teach each observer to attentionally track an ambiguous motion stimulus.

Observers were shown a fixed sequence of displays intended to demonstrate apparent motion and ambiguous apparent motion, and to train the observers to disambiguate the ambiguous motion displays at will. Observers who failed at least three steps in this procedure were considered to have "failed" this phase of the experiment. Initially, such observers were excluded from further participation at this point. Three observers were eliminated in this fashion. However, we then decided that it would be useful to test such observers in the main experiment, to determine whether or not the training procedure was useful and/or predictive of performance.

1. Observers were shown an unambiguous clockwise motion display (Figure 5) at 2 Hz and asked to report what they saw. All observers spontaneously reported a moving illusory square. If the observer reported a rocking motion percept rather than rotation, the experimenter prompted the observer, asking, "Can you see this as a big white square rotating clockwise?" All observers passed this step.

2. Observers were shown an ambiguous motion display at 1.5 Hz. After making sure that the observer could see the illusory square moving, a blue chevron was added to one of the inducers (Figure 6).

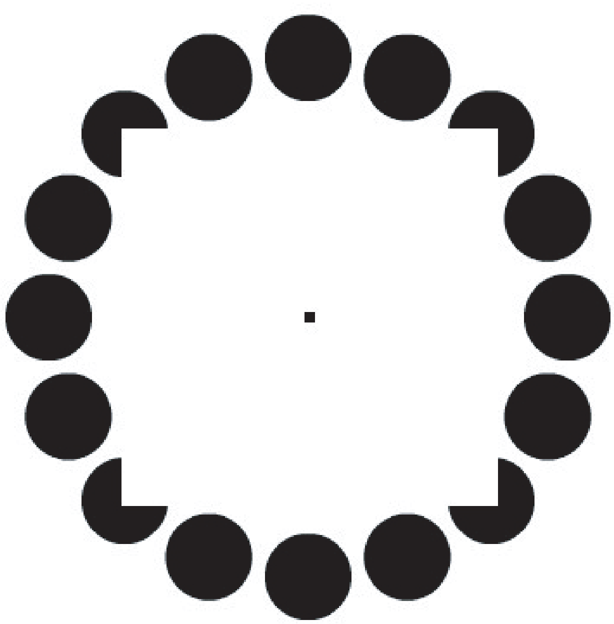

Figure 5. Ambiguous motion training method. This movie was used to demonstrate unambiguous clockwise motion.

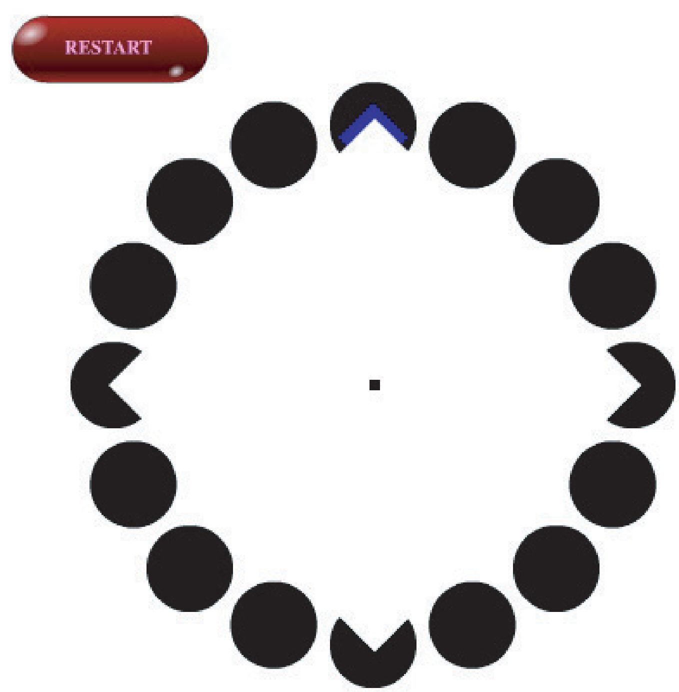


Figure 6. Ambiguous motion training method. This movie was used to demonstrate ambiguous motion, with a cue (blue chevron) to help observers see clockwise motion.

The chevron changed position in such a way as to specify clockwise motion for five consecutive frames, and then disappeared. The ambiguous motion sequence continued at this point. The experimenter instructed the observer to follow the chevron with their mind to produce a percept of clockwise motion, and then to imagine that the clockwise motion continued after the chevron disappeared. This procedure was repeated. The observer reported when the clockwise motion percept stopped. The observer had to report seeing clockwise motion for at least five frames in order to pass this step.

3. The procedure of the second step was repeated using a 2-Hz display and a counterclockwise motion cue.

4. Observers were shown a 1.5-Hz ambiguous motion display (Figure 7), and told, "This figure is ambiguous. Sometimes people see it as rotating clockwise, and sometimes they see it as rotating counterclockwise. Also, some people can use their mind and attention to control which way it appears to rotate for them. Which way are you seeing it now?" The experimenter then asked the observer to reverse the direction of motion. Observers who could not make the motion reverse failed the fourth step.

5. Starting with the fifth step, observers were asked to fixate the central dot while looking at the displays. Here they were shown a 2-Hz ambiguous motion display with a disambiguating blue chevron indicating clockwise motion. Again, observers had to keep the clockwise motion going for at least five frames after the chevron disappeared in order to pass this step.

6. The next display was a 1.5-Hz ambiguous motion display (Figure 7). Observers were again asked to control the direction of motion while maintaining fixation. If they could not control the direction, they failed this step.

7. Step 6 was repeated with a 2-Hz display.

8. Finally, we included a check for demand characteristics. Observers were shown an unambiguous 2-Hz clockwise motion display, and asked to see it moving counterclockwise. Observers who reported seeing counterclockwise motion failed this step.

***Training with masked letter identification.*** The purpose of this procedure was to give subjects experience with maintaining fixation, shifting attention, and reporting a masked item in the periphery. The procedure was similar to the main experiment, except that the pace was a constant 667 ms/step. Also, observers were provided with an unambiguous cue in the form of a white cue disk which stepped around the display at the proper rate. Observers needed only to follow the cue disk in order to be attending to the location of the critical, postmasked letter when it appeared. The order of conditions in training was counterbalanced.

***Main experiment.*** Experimental procedures were identical to those used in Experiment 1, except that only the attentional saccade and attentional pursuit conditions were run. The experiment took less than 1 hr for each observer.

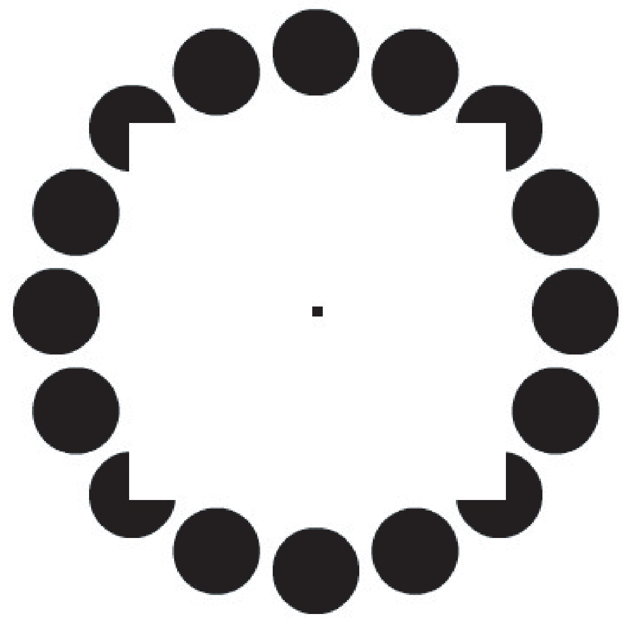

Figure 7. Ambiguous motion training method. This movie was used to demonstrate ambiguous motion.

## Results

Three observers were eliminated after the ambiguous motion training phase of the experiment. One observer was eliminated for having a constantly increasing staircase in the attentional pursuit condition. Results from the remaining 16 observers are shown in Figure 8. The mean attentional pursuit asymptote, 250.2 ms, was significantly lower than the mean attentional saccade asymptote, 500.2 ms ($t(15) = 4.15$, $p < .001$).

Performance in the training session was marginally negatively correlated with asymptote for the attentional saccade condition ($r = -.47$, $t(14) = 2.01$, $p = .064$), and unrelated to asymptote for the attentional pursuit condition ($r = -.23$, $t(14) < 1$). We split the observers into two groups, based on training performance in the attentional pursuit condition, and subjected the data to a 2 (attentional saccade vs. attentional pursuit condition) × 2 (high vs. low attentional pursuit training performance) repeated measures ANOVA. The effect of condition was, of course, significant ($F(1, 14) = 16.66$, $p < .005$, $MSE = 30017$). The

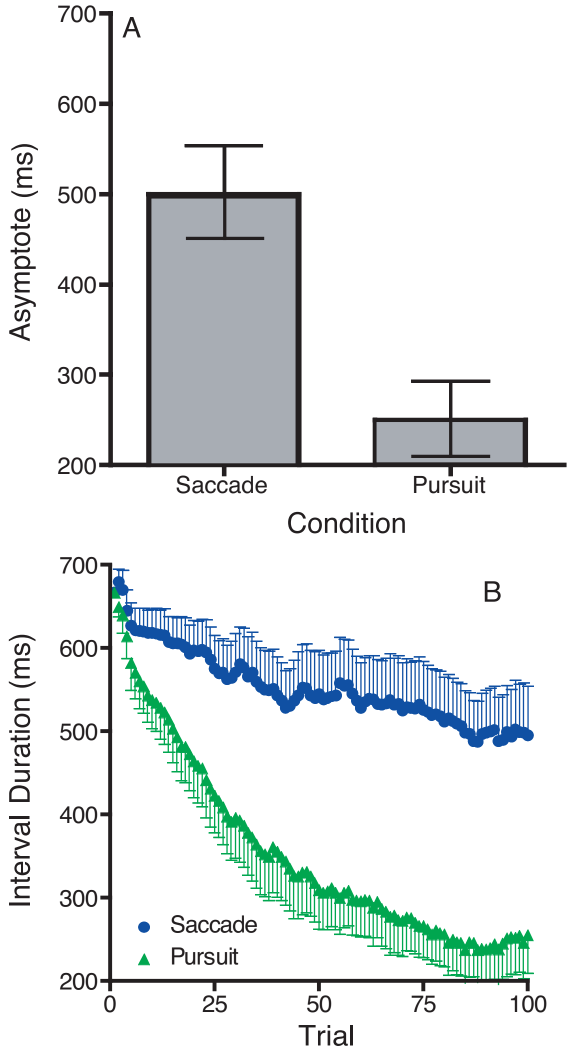


Figure 8. Data from Experiment 2. Blue circles denote the attentional saccade condition, green triangles the attentional pursuit condition. A shows the average asymptote in the attentional saccade and attentional pursuit conditions, whereas B plots the average staircases across all 16 subjects.

observers who did well in the initial training on the attentional pursuit task produced somewhat lower asymptotes overall in the main experiment ($F(1, 14) = 4.59$, $p = .050$, $MSE = 30809$). However, training performance did not predict the difference between the attentional saccade and attentional pursuit conditions ($F(1, 14) < 1$, $MSE = 30017$). Even for the high performance group, the mean asymptote was similar to that obtained in Experiment 1 (Table 1).

| | Saccade | | Pursuit | |
|---|---|---|---|---|
| | Mean | *SEM* | Mean | *SEM* |
| Overall | 500.16 | 50.75 | 250.18 | 40.36 |
| High | 411.33 | 48.66 | 206.01 | 45.03 |
| Low | 588.99 | 80.10 | 294.34 | 66.29 |

Table 1. Asymptotes (ms per frame) from Experiment 2 according to performance on counterphase training.

Two observers who "failed" the ambiguous motion training procedure were included in the analysis. With a group size of 2, it is not possible to come to any statistically valid conclusions. However, their results in both conditions appear to be within the range set by the remaining 14 observers. Observer AK produced an asymptote of 800.0 ms in the attentional saccade condition and 172.0 ms in the attentional pursuit condition, whereas observer MK's results were 809.3 and 502.7 ms, respectively. Removing these two observers would not change the basic results of this experiment. It appears that either the outcome of the ambiguous motion training procedure is not predictive of observers' attentional pursuit performance or, conversely, that the ambiguous motion training procedure was successful in bringing observers' performances up to par.

Asymptotes were longer overall in this experiment than in Experiment 1. We analyzed the effect of Experiment 2 on asymptotes from the attentional saccade and attentional pursuit conditions from these two experiments using a mixed ANOVA, and obtained a significant main effect ($F(1, 30) = 5.75$, $p < .05$, $MSE = 33461$). However, the difference between the two conditions remained constant (interaction $F(1, 30) = 2.24$, $p > .10$, $MSE = 24491$).

### Discussion

This experiment closely replicated the results of Experiment 1. Shifts of attention between objects took about twice as long as shifts of attention pursuing object movement. However, the rate of attentional pursuit was still fairly slow, somewhat slower than in Experiment 1, and substantially slower than observed by Verstraten et al. (2000). This difference persisted from the previous experiment despite the fact that we provided our observers with instruction and training on tracking ambiguous apparent motion stimuli. Of course, it is probable that had we used still more experienced psychophysical observers, performance in the attentional pursuit condition would have improved. However, even our best observer could not follow the attentional pursuit display faster than 90 ms/step, and recall that this is the rate necessary to achieve only 66.7% performance (Verstraten et al. adjusted flicker rate until performance was "just possible").

## Experiment 3

Experiment 3 was designed to provide converging evidence for the conclusions of Experiments 1 and 2 using the method of constant stimuli. There are three drawbacks to the adaptive staircase technique used in Experiments 1 and 2. First, observers may have difficulty changing the pace of attention on a trial-by-trial basis; and, therefore, we may be underestimating the maximum rate that observers could achieve if allowed to maintain the same rate over many trials. This factor might also interact with the attentional saccade/pursuit manipulation.

Second, the staircase method provides only a single point along the psychometric function, in this case the 66.7% point. The full function may provide more information about the relationship between the two conditions. Furthermore, interpretation of staircase data rests on certain assumptions, including the monotonicity of the underlying psychometric function (Leek, 2001). The functions here could conceivably be nonmonotonic, as certain rates might be more natural, and attempting to move attention at faster or slower rates might yield worse performance. Although the average staircase functions are well behaved (see Figures 4A and 8B), this is not true for all individual functions (e.g., Figure 4B and 4C), suggesting a possible violation of the assumptions underlying the staircase method.

Third, because the average asymptote in the attentional saccade condition in Experiments 1 and 2 was approximately equivalent to the rate of ocular saccades when viewing natural scenes, we wished to measure performance under conditions of enforced fixation (see Experiment 4). However, the adaptive method is not suited to the eyetracker. Trials on which the observer made eye movements would have to be eliminated online so as not to affect the staircase track, and we did not have the technology available to do this.

For these reasons, in Experiment 3 we measured performance at three fixed frame durations: 107, 307, and 507 ms, corresponding to rates of 9.4, 3.3, and 2.0 Hz, respectively. The primary dependent variable becomes accuracy, rather than the staircase asymptote.

## Methods

### Observers

Fourteen participants from the Visual Attention Laboratory's paid observer panel participated in this experiment.

### Apparatus & stimuli

Apparatus and stimuli were identical to those used in Experiment 2.

### Procedure

The sequence of events on a given trial was identical to that in Experiments 1 and 2. However, the interval duration was fixed for a block of trials, instead of being controlled by a staircase procedure. There were six blocks total, generated by crossing attentional saccade/pursuit conditions with three interval durations (107 ms, 307 ms, and 507 ms). Each block began with 20 practice trials, followed by 100 experimental trials. Observers ran all three blocks of one condition, followed by three blocks of the other condition. Order of conditions was counterbalanced across subjects, and the order of the interval duration blocks was also counterbalanced.

At the beginning of each condition, the experimenter stressed the importance of maintaining fixation on the fixation cross and illustrated how observers should move their attention, pointing with a pen to the appropriate placeholder on the screen. The experimenter remained with the observer until satisfied that the observer was performing the task properly. The experiment took approximately 2 hr.

## Results

Data from two observers were discarded due to difficulty with the task and failure to return to complete the study. Performances of the remaining 12 observers are plotted in Figure 9. Accuracy data were analyzed using a two-way within-subjects ANOVA with condition (attentional saccade vs. pursuit) and duration (107 ms, 307 ms, and 507 ms) as factors. Accuracy was higher in the attentional pursuit condition compared to attentional saccade ($F(1, 11) = 14.3$, $p < .005$), and increased with interval duration ($F(2, 22) = 24.4$, $p < .0001$). The two factors interacted ($F(2, 22) = 5.1$, $p < .05$) such that performance was identical at 507 ms, but the two conditions diverged as interval duration decreased.

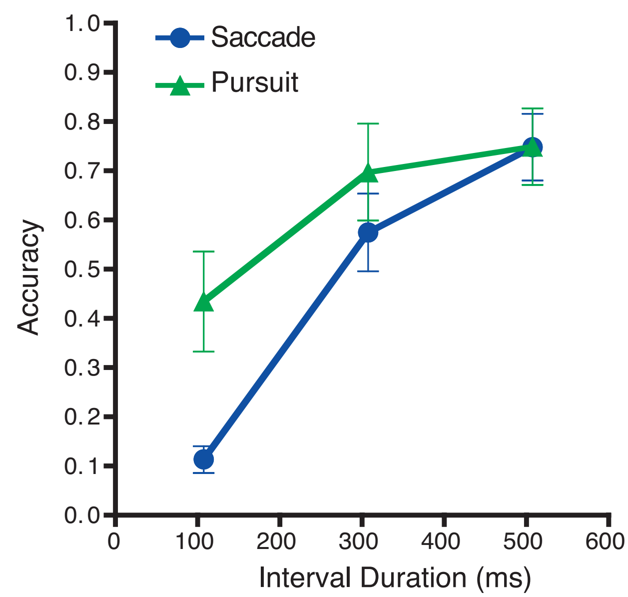


Figure 9. Data yielded by the method of constant stimuli (Experiment 3). Blue circles denote the attentional saccade condition, green triangles the attentional pursuit condition.

## Discussion

Despite the changes to the experimental procedure, between-object attentional shifts were still slower than pursuit accompanied by motion. The attentional pursuit condition yields better performance at all but the longest interval durations. At the shortest interval duration (107 ms), attentional saccade performance is near chance (11.3%)[1], whereas attentional pursuit performance is still quite respectable (43.4%). These data reinforce the conclusions of Experiments 1 and 2, and show that the difference between the conditions obtains even without the rate variability introduced by the staircase procedure.

Nor was the staircase procedure responsible for the slowness of the rate estimates observed in our previous experiments. In fact, if we interpolate the 66.7% thresholds, we find values of 459.3 ms for the attentional saccade condition and 346.4 ms for the attentional pursuit condition, both slower than those derived from the staircase procedure. Both of the functions in Figure 6 are monotonic, validating the staircase procedure used in Experiments 1 and 2. One caveat, however, is that the functions appear to reach a ceiling around 75%. The staircase method assumes that all errors can be attributed to failure to keep up with the desired pace on a given trial. If some errors are caused by the failure to recognize the letter even when attention is properly deployed, then the asymptotic durations will be artificially long.

Finally, the converging evidence provided by the method of constant stimuli is quite useful, not only for the additional validation but also because this method is more tractable for certain experimental manipulations. In particular, Experiments 4 and 5 test two alternative explanations for our results: eye movement artifact and a counting strategy.

# Experiment 4

Although observers in this experiment were strongly encouraged and reminded to maintain fixation, it is of course possible that they failed to do so and employed eye movements differentially in the two tasks. Specifically, one hypothesis is that the attentional saccade condition induces observers to make eye movements from placeholder to placeholder, while observers can more easily maintain fixation in the attentional pursuit condition. On this view, the attentional saccade condition produces slower rate estimates not because it requires between-object attentional shifts, but because it induces observers to employ eye movements.

We believe that eye movements are actually more likely in the attentional pursuit condition. First, small eye movements are more difficult than large eye movements (Bahcall & Kowler, 1999), and the distance between placeholders was necessarily smaller in the attentional saccade condition than in the attentional pursuit condition. Second, the attentional pursuit condition has salient offsets and onsets which might be expected to aid the eyes to disengage from one placeholder and move to the next, when compared to the static stimuli in the attentional saccade condition.

To definitively eliminate the possibility that differential eye movements affected the results in Experiment 4, we used an eyetracker.

## Methods

### Observers

Fourteen observers were recruited from the campus of the University of California, San Diego. All observers reported normal or corrected-to-normal vision, and earned extra credit in undergraduate psychology courses in return for their participation.

### Apparatus & stimuli

Stimuli were presented on an 18 in. flat screen CRT monitor with a refresh rate of 75 Hz controlled by Power Macintosh computers running Matlab 5.1 (Mathworks) utilizing the Psychophysics Toolbox (Brainard, 1997) routines. Observers viewed the CRT from a chinrest 50 cm away.

A Skalar Iris™ (http://www.skalar.nl/) infrared eye tracker monitored the horizontal movement of the observers' right eye (Reulen et al., 1988). Eye tracker output was acquired first by an oscilloscope at a sample rate of 4 ms, with acquisition triggered by the beginning of the stimulus train. The oscilloscope buffered the signal until each trial was over, whereupon the contents of the oscilloscope buffer were downloaded to a Wintel machine.

Stimuli were identical to those used in Experiment 3.

### Procedures

***Eye tracking.*** With observers' chins on the chinrest, the eye tracker was calibrated by manual adjustment to create a high gain and approximately symmetric signal when fixating two targets on opposite sides of the screen. Observers then participated in an automated calibration session consisting of repeated verbal commands from the computer to saccade among five dots. Two targets were spaced 1.7° from fixation on either side of the fixation dot, and two more were set 15° from fixation. After the experiment, the lab technician used this calibration data to estimate the average eye tracker position signal change caused by a 1.7° oculomotor saccade. All experimental trials in which the total change in eye tracker position signal was below this figure were allowed. The position signal plot of those trials that exceeded this criterion was visually scrutinized by the lab technician. Blinks are easily recognized and were ignored. Any saccadelike signal pattern which traversed a larger amplitude than that of the 1.7° calibration saccade was flagged and the trial was excluded.

An oculomotor saccade of 1.7° is several times smaller than the distance from fixation to the stimulus circles, as well as smaller than the distance from the center of one circle to the next, hence rejecting eye movements of this size is a fairly conservative measure to prevent any untoward contribution of eye movements. The lab technician also checked for sinusoidal patterns on a coarse timescale, which would be consistent with the eyes moving in a small circle, pacing the disks. Other patterns consistent with movement with each stimulus appearance, even if not exceeding the reference saccade criterion, were also grounds for discarding the trial.

The experimenter visually monitored the oscilloscope intermittently during the experimental trials and, if large eye movements were detected, the observer was reminded to fixate. Eye-tracking data were not acquired in the 507-ms

ISI condition due to a limit on the size of the oscilloscope buffer. Nevertheless, to ensure that from the point of view of the observers the procedure would be the same, participants were not informed of this and the eye-tracking headset was placed on the participant's head and calibrated in this condition just as in the others.

***Experiment.*** Before each block, observers participated in a variable number of practice trials, determined by how quickly they learned to perform the task.

Procedures were substantially identical to those used in Experiment 3, except for the following details. Observers participated in either the attentional saccade condition or the attentional pursuit condition, and ran three blocks of 60 trials, each with a different interstimulus interval: 147, 307, and 507 ms. The order of blocks was pseudorandomized across observers. The experiment took less than 1 hr.

### Results

One observer's data were eliminated because his eye-tracking data were too noisy to estimate the threshold for detecting an eye movement. This left seven participants in the attentional pursuit condition and six in the attentional saccade condition.

First, the data were analyzed using the same procedure as in the previous experiment, except with condition as a between-subjects factor. Attentional pursuit observers were more accurate than attentional saccade observers ($F(1, 11) = 29.0$, $p < .005$), and accuracy increased with interval duration ($F(2, 22) = 13.9$, $p < .0005$). The two factors interacted ($F(2, 22) = 4.0$, $p < .05$), such that the difference between groups decreased with interval duration. These results are quite similar to Experiment 3, except that performance was overall somewhat improved in this experiment. These data are shown as green (pursuit) and blue (saccade) symbols in Figure 10. Linear interpolation yields 66.7% thresholds at 619.2 ms for the attentional saccade condition and 159.0 ms for the attentional pursuit condition.

Next, we excluded trials on which observers made eye movements. Using the rejection criterion described in the methods section, an average of 10.3 trials out of the 60-trial attentional saccade conditions (17.1%) and an average of 9.5 trials out of the 60-trial attentional pursuit conditions (15.8%) were rejected due to eye movement. The number of trials excluded did not vary with group, interval duration, or their interaction (all $F < 1$). The means of the included trials are plotted as red symbols in Figure 10. It is clear from the figure that discarding the trials in which subjects moved their eyes had no effect on the means, because the data with and without the rejected trials fall on top of each other.

### Discussion

If the slower pace in the attentional saccade condition was due to a differential tendency to employ eye movements in that condition, then enforcing fixation should have eliminated, or at least reduced the difference between conditions. If anything, however, it had the opposite effect. The difference between conditions is noticeably larger in Experiment 4. This is partially because the observers in the attentional pursuit condition of Experiment 4 outperformed their counterparts in Experiment 3 at all three interval durations, reaching an apparent ceiling just under 85%. In the attentional saccade condition, performance is similar for the two shorter interval durations, but at the long interval duration, the attentional saccade observers in this experiment did not reach the same level of performance as the attentional pursuit observers. Accuracy for this block was only 63.1%, compared to 74.8% in the previous experiment. Linear interpolation yields thresholds of 619.0 ms for the attentional saccade condition and 159.0 ms for the attentional pursuit condition. We can confidently conclude that the difference between the conditions is not related to eye movements.

## Experiment 5

We can be fairly sure that observers were fixating as asked in Experiments 1–3, because enforcing fixation in Experiment 4 did not substantially change the pattern of results. But can we be sure that observers were shifting covert attention at the appropriate pace? The brief exposure of the probe stimulus ensured that observers were required to attend to the probe location in order to respond correctly. However, an external cue (auditory or visual) was available in both conditions. Might an observer choose to simply

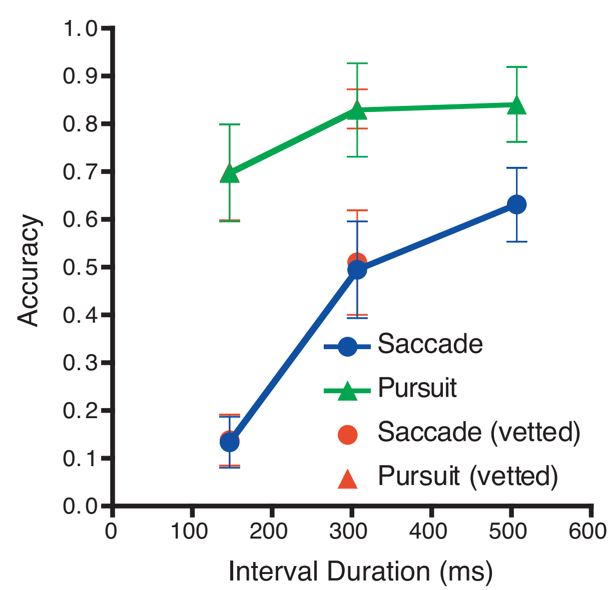


Figure 10. Data with fixation monitoring. Circles denote the attentional saccade condition, triangles the attentional pursuit condition. Blue and green symbols indicate averages based on all trials. Red symbols show the data after removing those trials on which eye movements were made.

count the cues, then shift attention to the correct location when the probes appeared? This would not explain the difference between conditions, but if observers were not following directions, then it would be difficult to draw any conclusions about attention from these data.

Such a strategy would require the observer to maintain a running count of cues in memory, and perform a modulus operation on the result (because the probe never appeared until at least one full cycle was complete), shift attention to the appropriate probe location, and identify a letter, all within 133 ms. This time constraint seems rather severe, given that the time to interpret a cue, shift attention, and identify a letter has been measured at 300–400 ms (Shih & Sperling, 2002). Nevertheless, we wished to empirically rule out this possibility. In Experiment 5, observers performed the attentional saccade and pursuit tasks under a working memory load that should preclude any possibility of counting cues and performing mental arithmetic.

## Methods

### *Observers*

Ten observers from the Visual Attention Laboratory's paid observer pool participated in this experiment.

### *Apparatus & stimuli*

Stimuli in the attentional saccade and pursuit tasks were identical to those used in Experiment 3. Digits for the digit memory task were presented in 12-point Geneva font.

### *Procedures*

In the *no load* condition, the procedure was identical to that of Experiment 3, except that interval duration was fixed at 307 ms. (One observer also ran in blocks of 107 and 507 ms under load conditions; this observer's data for 307 ms did not differ noticeably from the other observers). In the *load* condition, a digit memory task was added. Prior to presentation of the placeholder display on each trial, a sequence of seven digits was presented on the screen. After memorizing the digits, the observer pressed a key to continue to the attention shifting task. After the observer responded to the probe letter, two seven-digit sequences were presented, and the observer had to choose which one of the sequences was presented at the start of the trial. One of the two sequences was always the correct one, while the other, foil sequence differed only by a single digit. Auditory feedback was provided for this task in the form of a low tone for a correct response or a high tone for an incorrect response. Text feedback on the screen also indicated the accuracy of both letter probe and digit responses.

Each block of trials consisted of 20 practice trials followed by 100 experimental trials. Observers either completed the two no load blocks before the two load blocks, or vice versa. The order of attentional saccade and pursuit blocks was the same for the load and no load conditions, but counterbalanced across observers.

## Results

Two observers had undue difficulty with the load condition, and their data were discarded. Data from the remaining eight observers are shown in Figure 11. Accuracy on the digit memory task (blue bars) was quite high, .84 for the attentional saccade condition and .94 for the attentional pursuit condition. A one way within-subjects ANOVA indicated that this difference was significant ($F(1, 7) = 20.0$, $p < .005$). Accuracy on reporting the probe letter in the attention shifting task is shown as a function of condition (attentional saccade vs. pursuit) for single task (green bars) and dual task (red bars) blocks. For the load condition, only trials on which the digit memory task response was correct are included. A two-way within-subjects ANOVA indicated that accuracy was substantially higher for the attentional pursuit condition when compared to the attentional saccade condition ($F(1, 7) = 11.0$, $p < .05$). However, there was no main effect of load, nor did load interact with condition (both $F < 1$).

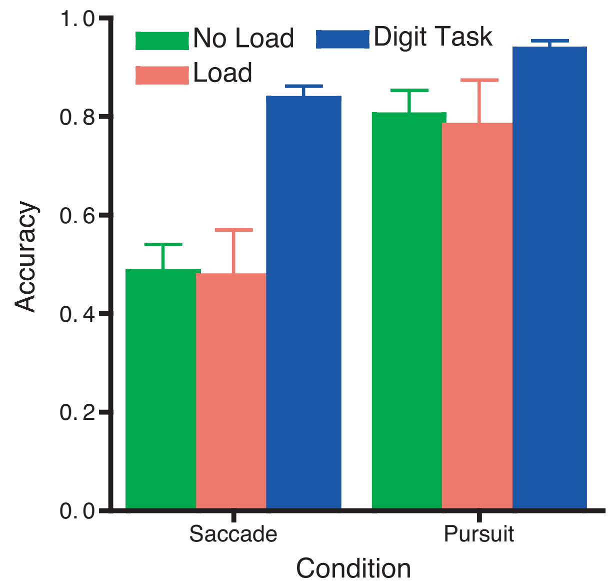


Figure 11. Accuracy data from Experiment 4. Green bars show single task performance, red bars dual task performance, and blue bars indicate the digit task performance in dual task blocks.

## Discussion

If it is unlikely that observers could maintain a running cue count, perform a modulus operation when the probe letter appeared, then shift attention to the appropriate location within 133 ms, then it is highly unlikely that they could do so while holding a 7-digit number in working memory. Yet adding the digit memory task did not noticeably affect performance on the attention shifting task. We conclude that observers were in fact shifting attention in accordance with experimenter instructions in this and previous experiments, rather than employing some alternative strategy based on counting cues.

It is interesting that the attention shifting tasks did have an effect on the digit memory task. Although accuracy was quite high overall, keeping the digits in memory was noticeably more difficult during the attentional saccade task, suggesting that between-object shifts may be more demanding on central resources.

# General discussion

Three basic findings emerge from these experiments. First, attentional saccades between objects are quite slow, on the order of 300-500 ms. Second, attention can shift substantially faster when an observer is attending to an object which moves. We believe this is because when the object moves, it carries the focus of attention with it. Third, these attentional pursuit movements are still substantially slower than the previously observed rates for tracking objects in apparent motion (Verstraten et al., 2000). We discuss the implications of each of these findings in turn.

## Shifts of attention between objects are slow

### *Shifting time and object recognition*

The attentional saccade condition was intended to measure a pure shifting cost and, even if it does not reach this goal, it should come closer than did previous work. In our attentional saccade condition, the time needed to detect, process, and interpret cues has been minimized. Moreover, in contrast to estimates of the shift time from visual search paradigms, our estimates are not dependent on assumptions about the parallel or serial nature of attention or the number of attentional shifts required to complete the task, because these properties are specified by the experimental method itself.

During the successive shifts of attention in this experiment, observers did not have to process a cue, recognize a letter, or perform any other task—they simply had to shift attention at the appropriate pace. Nevertheless, the rates found here are not far from those found by Wolfe et al. (2000) in an experiment in which letter identification was required with each step. If the rate of object identification was the limiting factor, then we would expect that the estimate of attentional dwell time would depend on the complexity of the task required of the observer. Instead, comparing present results with those of Wolfe et al. (2000), we see that dwell times are similar whether or not observers simultaneously perform a difficult letter recognition task. This is consistent with the suggestion (Moore & Wolfe, 2001) that attention does not recognize objects per se, but rather serves to transfer information to later processing stages, which then perform object recognition autonomously.

Of course, the act of shifting attention may not be a unitary process. For instance, in Posner's influential account (Posner & Cohen, 1984), shifting attention requires the observer to disengage attention from one locus, move attention, and then engage at the new locus. On this view, even though we may have removed cue processing time from the shift cost, we have not isolated the time to move attention per se from the disengage and engage times. Some other design may be required to further parse the act of shifting attention,

### *Attention, eye movements, and free will*

It has been known for a long time that covert attention can be shifted while the eyes remain fixed (Jonides, 1981). Similarly, attentional pursuit can be carried out in the absence of eye movements (Verstraten, Hooge, Culham, & Van Wezel, 2001). Nevertheless, a growing body of evidence suggests a substantial overlap between the neural substrates of covert and overt attention (Moore, Armstrong, & Fallah, 2003; Sheliga, Craighero, Riggio, & Rizzolatti, 1997). Rizzolatti's premotor theory (Rizzolatti, Riggio, Dascola, & Umilta, 1987) argues that shifts of covert attention are in fact a consequence of oculomotor programming. In support of this view, a number of studies have demonstrated that attention precedes the shift of gaze to a location or object (Deubel & Schneider, 2003; Godijn & Theeuwes, 2003). Indeed, ocular pursuit and attentional pursuit seem to be tightly yoked (Godijn & Pratt, 2002; Khurana & Kowler, 1987; Van Donkelaar & Drew, 2002). Our results would seem to present a puzzle for this account of the coupling between attention and eye movements. How can an attentional saccade taking 300–500 ms precede an ocular saccade with a typical latency of 200–300 ms?

The resolution to this discrepancy may lie in the way that attentional shifts are controlled in these experiments. Here the visual system is presented with a set of physically identical objects. The only difference between the target of the attentional saccade and the other placeholders is that the observer (following the experimenter's instructions) has selected that object. Another way to put it is that only the will of the observer distinguishes between targets and nontargets. The results of Wolfe, Alvarez, and Horowitz (2000) suggest that such attentional shifts are substantially slower than "anarchic," or stimulus-driven shifts, when the attentional target is defined by physical differences from distractors, and the order of attentional saccades is unconstrained. These stimulus-driven shifts seem to take on the order of 100 ms, rapid enough to precede an ocular saccade.

Conversely, data from Kowler et al. (1995) indicate that when the target of an eye movement in a complex array is the salient item, latencies are much faster than when the target is non-salient, defined only by the will of the observer. In the latter case, saccade latencies are 300–500 ms.

## Attention, motion, and object representations

A fundamental issue in the study of attention is whether attention selects objects or locations, or a combination of the two. These possibilities have been evaluated with increasing refinement because of Duncan's landmark study of object-based attention 20 years ago (Blaser, Pylyshyn, & Holcombe, 2000; Davis et al., 2001; Duncan, 1984). In the present experiments, we have directly compared the maximum rate for attentional pursuit of one object to the rate for attentional saccade between objects. The faster rate of attentional pursuit documented here provides novel support for object-based theories of attention. Attention might be shown to pursue a moving target at still faster rates in a paradigm not limited by the additional demands of attentive tracking.

Alternatively, returning to the Posnerian scheme, the attentional pursuit condition might have allowed observers to shift attention without disengaging, because the object of attention was removed from the screen. On this account, the difference between attentional saccade and attentional pursuit conditions measures the disengage process. However, we think this explanation is unlikely. Mackeben and Nakayama (1993) measured the time required for a single shift of attention from a fixation stimulus to a peripheral cue. When the fixation stimulus remained visible (thus requiring a disengage operation), the shift took 100-200 ms. However, if the fixation stimulus was turned off 200 ms before the cue appeared, they observed what they called "express attentional shifts" of 33 ms. If observers truly did not have to disengage attention in the attentional pursuit condition, then the rates should have approached the rate of these express attentional shifts.

## Temporal architecture of the visual system

The strong form of the object-based attention hypothesis predicts that there should be absolutely no cost to shifting attention among locations linked by a moving object. This would imply that the rate limit on attentional shifts should be the same as the rate of the fastest moving object one can perceive. Empirically, however, this is clearly not the case. A more conservative prediction is that the attentional rate limit should be determined by the limit on object continuity processes (i.e., impletion or indexing). This is essentially the assumption made by Verstraten et al. (2000) when they deduced a limit on the temporal resolution of attention from performance on a tracking task.

But as we noted in the "Introduction," tracking tasks measure the observer's ability to maintain object continuity, which may not be the same as attentional selection. The visual system may be able to determine that an object presented at time $t_1$ is the same as an object presented at $t_0$ even if that object has not been continually attended from $t_0$ to $t_1$. However, our attentional saccade and attentional pursuit tasks specifically required attention to be directed to the selected object at all times during the trial, because probe letter identification might be required at any time. Using this method, we found that the rate limit on attention measured in the counterphase task was much lower than that inferred by Verstraten et al (2000). We argue that the 7-Hz rate measured in the Verstraten et al. study was actually the temporal resolution of object continuity processes, whereas shifts of attentional selection are limited in a different way. Based on the results reported here, we suggest that object continuity cannot be maintained at the highest rates that the motion system can resolve, nor can attention follow an object at the highest rates that object continuity can resolve.

This view leads us to the following prediction for observers tracking moving objects: If the task requires only individuating objects, maximum tracking rates will be much greater than if the task requires focal attention to the object

Our current results are consistent with a processing hierarchy. Early motion processing with high temporal resolution provides input to object continuity processes. These mid-level processes have coarser temporal resolution than the motion system, but can follow faster shifts than selective attention can. Because top-down attentional input appears necessary to construct the motion percept in counterphase motion (see also Battelli et al., 2001; Cavanagh, 1992; Horowitz & Treisman, 1994), the flow of information in the hierarchy must be bi-directional or "re-entrant" (Di Lollo, Enns, & Rensink, 2000; Hochstein & Ahissar, 2002).

## Acknowledgments

This research was supported by a grant from the Air Force Office of Scientific Research to JMW. AOH was supported by a postdoctoral National Eye Institute National Research Service Award at the University of California, San Diego, and EY01711 awarded to DIA MacLeod. We would like to thank Frans Verstraten and an anonymous reviewer for helpful comments on the manuscript.

Commercial relationships: none.
Corresponding author: Todd S. Horowitz.
Email: toddh@search.bwh.harvard.edu.
Address: Brigham & Women's Hospital Visual Attention Laboratory, 64 Sidney Street, Suite 170, Cambridge, MA 02319.

## Footnotes

[1]Chance can be defined as 16.7%, 7.7%, or 3.9%, depending on whether the computation is based on the number of probes presented (6), the number of possible probe letters (13), or the total number of letters in the English alphabet (26); observers were not explicitly given the set of possible probe letters. Because Command performance in the 107-ms condition was significantly below 16.7% ($t(12) = 2.4$, $p < .05$), we suspect that the correct definition must be 7.7% or 3.9%.

## References


Bahcall, D. O., & Kowler, E. (1999). Attentional interference at small spatial separations. *Vision Research, 39*(1), 71-86. [PubMed]

Battelli, L., Cavanagh, P., Intriligator, J., Tramo, M. J., Henaff, M. A., Michel, F., & Barton, J. J. (2001). Unilateral right parietal damage leads to bilateral deficit for high-level motion. *Neuron, 32*(6), 985-995. [PubMed]

Blaser, E., Pylyshyn, Z. W., & Holcombe, A. O. (2000). Tracking an object through feature space. *Nature, 408*(6809), 196-199. [PubMed]

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 443-446. [PubMed]

Briand, K. (1998). Feature integration and spatial attention: More evidence of a dissociation between endogenous and exogenous orienting. *Journal of Experimental Psychology: Human Perception & Performance, 24*(4), 1243-1256.

Briand, K. A., & Klein, R. M. (1987). Is Posner's "beam" the same as Treisman's "glue?" On the relation between visual orienting and feature integration theory. *Journal of Experimental Psychology: Human Perception & Performance, 13*(2), 228-241. [PubMed]

Cavanagh, P. (1992). Attention-based apparent motion. *Science, 257*, 1563-1565. [PubMed]

Cheal, M., & Lyon, D. R. (1991). Central and peripheral precuing of forced-choice discrimination. *Quarterly Journal of Experimental Psychology A, 43*(4), 859-880. [PubMed]

Chen, Z. (2000). An object-based cost of visual filtering. *Perception & Psychophysics, 62*(3), 482-495. [PubMed]

Davis, G., Welch, V. L., Holmes, A., & Shepherd, A. (2001). Can attention select only a fixed number of objects at a time? *Perception, 30*(10), 1227-1248. [PubMed]

Deubel, H., & Schneider, W. X. (2003). Delayed saccades, but not delayed manual aiming movements, require visual attention shifts. *Oculomotor and Vestibular Systems: Their Function and Disorders, 1004*, 289-296. [PubMed]

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General, 129*(4), 481-507. [PubMed]

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113*(4), 501-517. [PubMed]

Duncan, J., Ward, R., & Shapiro, K. L. (1994). Direct measurement of attentional dwell time in human vision. *Nature, 369*(6478), 313-315. [PubMed]

Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception & Psychophysics, 62*(3), 425-451. [PubMed]

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General, 123*(2), 161-177. [PubMed]

Godijn, R., & Pratt, J. (2002). Endogenous saccades are preceded by shifts of visual attention: Evidence from cross-saccadic priming effects. *Acta Psychologica, 110*(1), 83-102. [PubMed]

Godijn, R., & Theeuwes, J. (2003). Parallel allocation of attention prior to the execution of saccade sequences. *Journal of Experimental Psychology: Human Perception and Performance, 29*(5), 882-896. [PubMed]

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron, 36*(5), 791-804. [PubMed]

Horowitz, T., & Treisman, A. (1994). Attention and apparent motion. *Spatial Vision, 8*(2), 193-219. [PubMed]

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature, 394*(6693), 575-577. [PubMed]

Horowitz, T. S., & Wolfe, J. M. (2001). Search for multiple targets: Remember the targets, forget the search. *Perception & Psychophysics, 63*(2), 272-285. [PubMed]

Horowitz, T. S., & Wolfe, J. M. (2003). Memory for rejected distractors in visual search? *Visual Cognition, 10*(3), 257-298.

Horowitz, T. S., Wolfe, J. M., & Alvarez, G. A. (2004). *The speed of free will*. Manuscript submitted for publication.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. L. A. Baddeley (Ed.), *Attention & performance IX*. Hillsdale, NJ: Erlbaum.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In D. R. D. Raja Parasuraman (Ed.), *Varieties of attention*. Orlando: Academic Press.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*(2), 175-219. [PubMed]

Khurana, B., & Kowler, E. (1987). Shared attentional control of smooth eye movement and perception. *Vision Research, 27*(9), 1603-1618. [PubMed]

Klein, R. (1988). Inhibitory tagging system facilitates visual search. *Nature, 334*, 430-431. [PubMed]

Klein, R. M., & Dick, B. (2002). Temporal dynamics of reflexive attention shifts: A dual-stream rapid serial visual presentation exploration. *Psychological Science, 13*(2), 176-179. [PubMed]

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research, 35*(13), 1897-1916. [PubMed]

Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Perception & Psychophysics, 63*(8), 1279-1292. [PubMed]

Mackeben, M., & Nakayama, K. (1993). Express attentional shifts. *Vision Research, 33*(1), 85-90. [PubMed]

McElree, B., & Carrasco, M. (1999). The temporal dynamics of visual search: Evidence for parallel processing in feature and conjunction searches. *Journal of Experimental Psychology: Human Perception & Performance, 25*(6), 1517-1539. [PubMed]

Moore, C. M., Egeth, H., Berglan, L., & Luck, S. J. (1996). Are attentional dwell times inconsistent with serial visual search? *Psychonomic Bulletin & Review, 3*(3), 360-365.

Moore, C. M., & Wolfe, J. M. (2001). Getting beyond the serial/parallel debate in visual search: A hybrid approach. In K. Shapiro (Ed.), *The limits of attention: Temporal constraints on human information processing* (pp. 178-198). Oxford: Oxford University Press.

Moore, T., Armstrong, K. M., & Fallah, M. (2003). Visuomotor origins of covert spatial attention. *Neuron, 40*(4), 671-683. [PubMed]

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research, 29*(11), 1631-1647. [PubMed]

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research, 40*(10-12), 1227-1268. [PubMed]

Peterson, M. S., & Juola, J. F. (2000). Evidence for distinct attentional bottlenecks in attention switching and attentional blink tasks. *Journal of General Psychology, 127*(1), 6-26. [PubMed]

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3-25. [PubMed]

Posner, M. I., & Cohen, Y. (1984). Components of attention. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X* (pp. 55-66). Hillsdale, NJ: Erlbaum.

Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition, 32*(1), 65-97. [PubMed]

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3*(3), 179-197. [PubMed]

Reeves, A., & Sperling, G. (1986). Attention gating in short-term visual memory. *Psychological Review, 93*(2), 180-206. [PubMed]

Reulen, J. P., Marcus, J. T., Koops, D., de Vries, F. R., Tiesinga, G., Boshuizen, K., & Bos, J. E. (1988). Precise recording of eye movement: The IRIS technique. Part 1. *Medical & Biological Engineering & Computing, 26*(1), 20-26. [PubMed]

Rizzolatti, G., Riggio, L., Dascola, I., & Umilta, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia, 25*(1A), 31-40. [PubMed]

Schmidt, W. C., Fisher, B. D., & Pylyshyn, Z. W. (1998). Multiple-location access in vision: Evidence from illusory line motion. *Journal of Experimental Psychology: Human Perception & Performance, 24*(2), 505-525. [PubMed]

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology, 38*(2), 259-290. [PubMed]

Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition, 80*(1-2), 159-177. [PubMed]

Scholl, B. J., Pylyshyn, Z. W., & Franconeri, S. L. (2004). *The relationship between property-encoding and object-based attention: Evidence from multiple object tracking*. Manuscript submitted for publication.

Sheliga, B. M., Craighero, L., Riggio, L., & Rizzolatti, G. (1997). Effects of spatial attention on directional manual and ocular responses. *Experimental Brain Research, 114*(2), 339-351. [PubMed]

Shih, S. I., & Sperling, G. (2002). Measuring and modeling the trajectory of visual spatial attention. *Psychological Review, 109*(2), 260-305. [PubMed]

Shioiri, S., Yamamoto, K., Kageyama, Y., & Yaguchi, H. (2002). Smooth shifts of visual attention. *Vision Research, 42*(26), 2811-2816. [PubMed]

Sperling, G., & Weichselgartner, E. (1995). Episodic theory of the dynamics of spatial attention. *Psychological Review, 102*(3), 503-532.

Tipper, S. P., Jordan, H., & Weaver, B. (1999). Scene-based and object-centered inhibition of return: Evidence for dual orienting mechanisms. *Perception & Psychophysics, 61*(1), 50-60. [PubMed]

Tipper, S. P., Weaver, B., Jerreat, L. M., & Burak, A. L. (1994). Object-based and environment-based inhibition of return of visual attention. *Journal of Experimental Psychology: Human Perception & Performance, 20*(3), 478-499. [PubMed]

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

Trick, L. M., & Pylyshyn, Z. W. (1993). What enumeration studies can show us about spatial attention: Evidence for limited capacity preattentive processing. *Journal of Experimental Psychology: Human Perception & Performance, 19*(2), 331-351. [PubMed]

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101*(1), 80-102. [PubMed]

Tsal, Y., & Lamy, D. (2000). Attending to an object's color entails attending to its location: Support for location-special views of visual attention. *Perception & Psychophysics, 62*(5), 960-968. [PubMed]

Van Donkelaar, P., & Drew, A. S. (2002). The allocation of attention during smooth pursuit eye movements. *Brain's Eye: Neurobiological and Clinical Aspects of Oculomotor Research, 140*, 267-277. [PubMed]

Verstraten, F. A., Cavanagh, P., & Labianca, A. T. (2000). Limits of attentive tracking reveal temporal properties of attention. *Vision Research, 40*(26), 3651-3664. [PubMed]

Verstraten, F. A., Hooge, I. T., Culham, J., & Van Wezel, R. J. (2001). Systematic eye movements do not account for the perception of motion during attentive tracking. *Vision Research, 41*(25-26), 3505-3511. [PubMed]

Weichselgartner, E., & Sperling, G. (1987). Dynamics of automatic and controlled visual attention. *Science, 238*(4828), 778-780. [PubMed]

Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift für Psychologie, 61*, 161-165.

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

Wolfe, J. M., Alvarez, G. A., & Horowitz, T. S. (2000). Attention is fast but volition is slow. *Nature, 406*(6797), 691. [PubMed]

Wolfe, J. M., & Bennett, S. C. (1997). Preattentive object files: Shapeless bundles of basic features. *Vision Research, 37*(1). [PubMed]