This article was downloaded by: *[Harvard College]*
On: *4 March 2009*
Access details: *Access Details: [subscription number 908310102]*
Publisher *Psychology Press*
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t716100704

The speed of free will

Todd S. Horowitz [a]; Jeremy M. Wolfe [a]; George A. Alvarez [b]; Michael A. Cohen [c]; Yoana I. Kuzmova [c]
[a] Brigham & Women's Hospital and Harvard Medical School, Cambridge, MA, USA [b] Harvard University, Cambridge, MA, USA [c] Brigham & Women's Hospital, Cambridge, MA, USA

First Published on: 02 March 2009

To cite this Article Horowitz, Todd S., Wolfe, Jeremy M., Alvarez, George A., Cohen, Michael A. and Kuzmova, Yoana I.(2009)'The speed of free will',The Quarterly Journal of Experimental Psychology,

To link to this Article: DOI: 10.1080/17470210902732155

URL: http://dx.doi.org/10.1080/17470210902732155

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
i*First*, 1–27

Psychology Press
Taylor & Francis Group

The speed of free will

Todd S. Horowitz and Jeremy M. Wolfe
Brigham & Women's Hospital and Harvard Medical School, Cambridge, MA, USA

George A. Alvarez
Harvard University, Cambridge, MA, USA

Michael A. Cohen and Yoana I. Kuzmova
Brigham & Women's Hospital, Cambridge, MA, USA

Do voluntary and task-driven shifts of attention have the same time course? In order to measure the time needed to voluntarily shift attention, we devised several novel visual search tasks that elicited multiple sequential attentional shifts. Participants could only respond correctly if they attended to the right place at the right time. In control conditions, search tasks were similar but participants were not required to shift attention in any order. Across five experiments, voluntary shifts of attention required 200–300 ms. Control conditions yielded estimates of 35–100 ms for task-driven shifts. We suggest that the slower speed of voluntary shifts reflects the "clock speed of free will". Wishing to attend to something takes more time than shifting attention in response to sensory input.

Keywords: Selective attention; Visual search; Volition.

Downloaded By: [Harvard College] At: 19:58 4 March 2009

How quickly can we change our focus of attention? Certainly we can rapidly orient to salient stimuli, such as a flash of light, a sudden movement, or a loud noise. However, in such cases, it would be more accurate to say that the focus of attention was changed for us. A large body of work in contemporary psychology has explored the mechanisms by which attention may be "captured" exogenously (Cole, Gellatly, & Blurton, 2001; Luck & Thomas, 1999; Ludwig & Gilchrist, 2002; Theeuwes & Godijn, 2002; Yantis, 1993). Here we are interested in endogenous shifts of attention, when we deliberately change our focus from one object to another more or less equally salient object without any change in the stimulus itself. The literature suggests that endogenous changes in attentional focus are more sluggish than exogenous changes. That is, it is faster to

---

Correspondence should be addressed to Todd S. Horowitz, Visual Attention Laboratory, Brigham & Women's Hospital, 64 Sidney Street, Suite 170, Cambridge, MA 02139, USA. E-mail: toddh@search.bwh.harvard.edu

We would like to thank Aude Oliva for graciously allowing us to use her eye tracker, Winston Chang and Barbara Hidalgo-Sotelo for assistance with setting up the eye-tracking experiments, and Naomi Kenner for piloting the eye-tracking experiments. We would also like to thank Ulrich Ansorge, Artem Belopolsky, David Rosenbaum, Shui-I Shih, Sandy Pollatsek, and an anonymous reviewer for helpful comments on various versions of the manuscript. This work was funded by Grant F49620–01–1–0071 from the Air Force Office of Scientific Research to J.M.W.

© 2009 The Experimental Psychology Society

http://www.psypress.com/qjep
DOI:10.1080/17470210902732155

Downloaded By: [Harvard College] At: 19:58 4 March 2009

have the world attract your attention than to shift attention by an effort of will.

One line of research supporting this notion is the literature on endogenous orienting. In orienting studies, developed by Posner (1980), observers are presented with a cue that tells them where to expect a target. The target usually requires a simple speeded detection response; sometimes a discrimination response is required (e.g., Müller & Rabbitt, 1989). Jonides (1981) distinguished between exogenous and endogenous cues. Exogenous cues directly indicate the target location; if you attend to the location of the cue, you will be attending to the target location (if the cue is valid). Endogenous cues are symbolic (e.g., an arrow) and are typically presented at a location where targets are never presented (e.g., fixation). If you attend to the cue, you must first interpret it and then initiate the attentional shift.

In order to measure the time course of attention in orienting studies, researchers typically measure performance as a function of the stimulus onset asynchrony (SOA) between the cue and the target. With exogenous cues, performance peaks at SOAs of around 100 ms, while endogenous cues require roughly 300 ms to reach maximal efficacy (Cheal & Lyon, 1991; Nakayama & Mackeben, 1989).

Reeves and Sperling (1986) introduced a different method. In their attentional reaction time (ART) paradigm, the display consists of two rapidly changing character streams (so-called rapid serial visual presentation, or RSVP streams): one at fixation, one peripheral. Observers are asked to monitor the peripheral stream, which consists of digits, for a trigger stimulus (e.g., "2"). When the trigger is detected, the observer reports the first letter detected in the central stream.[1] The time between the presentation of the trigger and the presentation of the reported letter is interpreted as the time required to shift attention voluntarily from one stream to the next. In agreement with the orienting studies, the mean ART was roughly 300 ms (see also Sperling & Weichselgartner, 1995).

One conceptual problem with these studies is that the time required to process the cue (or trigger) is incorporated into the estimate of attention-shifting time. Thus, the difference between automatic and voluntary shifts of attention might be that exogenous cues are simply interpreted more rapidly. One way of dealing with this problem is to estimate cue interpretation time and subtract it from the total ART (Peterson & Juola, 2000; Shih & Sperling, 2002). Instead we sought a design that would minimize the need to interpret cues. We based our approach on a variant of the visual search task.

In a visual search task, observers are asked to find a specified target object presented in a display cluttered with a number of distractor objects. Except in the special (if widely studied) case when the target is sufficiently different from the distractors to attract attention immediately ("pop-out", Maljkovic & Nakayama, 1994; Nakayama & Joseph, 1998; Nothdurft, 1991; Treisman & Gelade, 1980), a number of attentional shifts are required in order to locate the target. If we vary the number of distractors and measure reaction time (RT), the slope of the function relating RT to set size (total number of items) is proportional to the time required for each attentional shift. The relative time to shift attention can be easily compared across observers or conditions. Deriving an absolute measure of the shifting time requires a model of the sampling regime (e.g., with or without replacement, or something in between, see Horowitz & Wolfe, 2005) and the number of objects processed in a single attentional episode. Estimates derived from this method can vary from 25 to 100 ms.

The time between successive shifts of attention or between successive acts of selection is sometimes referred to as the attentional "dwell time" (Duncan, Ward, & Shapiro, 1994; C. M. Moore, Egeth, Berglan, & Luck, 1996; Ward,

---

[1] Actually, Reeves and Sperling (1986) asked observers to report the first four letters they detected in the central stream. The order in which letters were reported was of interest to them. However, these effects are more germane for a discussion of visual memory; the time course results are similar whether one uses four letters or just the first letter.

Downloaded By: [Harvard College] At: 19:58 4 March 2009

Duncan, & Shapiro, 1996). This raises a difficulty because "dwell time" implies that attention is deployed to a location and *dwells* there and does not depart until the attended object is identified. The problem is that the estimates of search rate in visual search are shorter than even the most optimistic estimates of the time required to identify an object (Thorpe, Fize, & Marlot, 1996; VanRullen & Thorpe, 2001).

It is useful to distinguish between the rate at which items can be selected and the amount of time required to process the selected items. It appears that the interval between shifts of attention is shorter than the time required to identify an item. This could not be true if selection and identification were occurring strictly sequentially. Thus, it appears that there is some parallel processing of the selected objects. This combination of serial and parallel processing can be thought of as a pipeline or a "car wash" (C. M. Moore & Wolfe, 2001; Wolfe, 2003) in which cars enter the car wash one after the other, but where multiple cars are being washed at any one time. In this paper, we are concerned with the rate of shifting. We return to the relationship of shifting time to dwell time in the Discussion.

Visual search has two key advantages as a method for measuring the speed of attentional shifts. First, no cue processing is required at all; attention is simply moving at its natural pace. Second, the procedure elicits multiple shifts of attention in a single trial, so that any initial start-up costs are relegated to the intercept of the RT × Set Size function.

In standard visual search paradigms, however, observers do not deliberately shift attention from one object to another. Rather, attention is driven by the attentional priority given to objects (Serences & Yantis, 2006), where priority is determined by both the "bottom-up" differences between an object and its surround (stimulus salience) and the object's "top-down" similarity to a target template (Duncan & Humphreys, 1992). Shifts of attention in search are therefore not purely voluntary. On the other hand, neither are they purely stimulus driven. We describe this type of attentional shift as "priority driven", which is meant to capture the fact that they are driven by the aforementioned priority, derived from the interaction of the properties of the world and the goals of the observer. In this paper, we contrast this class of attention shifts with voluntary or volitional shifts of attention, which depend explicitly on the will of the observer (albeit influenced by the experimenter's instructions), rather than the properties of the stimulus.

In order to measure voluntary shifts of attention, we have developed a set of paradigms in which observers can only respond correctly if their attention is in the right place at the right time (Experiments 1–3) or if they deploy attention to items in a specified order (Experiments 4 and 5). These paradigms were introduced in Wolfe, Alvarez, and Horowitz (2000), where we reported partial results for Experiment 2 and described the method used in Experiments 4 and 5.

In the first three experiments, the stimulus consists of a series of frames, each frame containing a circular array of letters. Observers are instructed to start the trial with attention at the top of the circle (i.e., "12 o'clock") and then shift attention one letter in a specified direction (clockwise or anticlockwise) on each frame. The letters in the array change on each frame. The target is presented on only one frame. The location of the target is determined by the frame, so that the target always appears in the *n*th position on the *n*th frame. Therefore, observers can only identify the target if they are attending to the *n*th position on the *n*th frame, which means that they had been shifting attention at the same rate as the frames are presented. A staircase method is used to estimate the frame duration at which observers can perform the task with 66.7% accuracy. Thus, the staircase asymptote is the estimate of the minimum time needed to shift attention. An auditory cue is presented at the onset of each frame. However, since the frame rate is constant during a trial, observers do not need to process the cue and can generate attentional shifts endogenously (see Large & Jones, 1999, for an example in the auditory domain).

Downloaded By: [Harvard College] At: 19:58 4 March 2009

In Experiment 1 this "commanded search" method was used to compare the command condition to a control condition in which the search task was the same, but observers were not required to shift attention in any particular order. In this "anarchic" control condition, attention is priority driven. The rate of processing in the commanded search condition was substantially slower than that in anarchic search. Experiment 2 replicated this finding while controlling for the total time of the target on the screen and including a control condition measuring standard search rates. Experiment 3 replicated the finding again using a visual pacing cue rather than an auditory one, to control for possible cross-modal demands. Experiment 4 provided converging evidence using static displays and no pacing cues at all. Instead, there are multiple examples of the target stimulus; the true target is the first one in order around a circular display. This forces observers to shift attention in an orderly fashion. Experiment 5 replicated the critical condition of Experiment 4 with eye tracking.

EXPERIMENT 1: COMMAND VERSUS ANARCHY

Method

Participants

A total of 10 participants were recruited from the Visual Attention Laboratory's paid participant panel. In this and all subsequent experiments, all participants passed Ishihara's Tests for Color-Blindness and had 20/25 corrected vision or better. In this and all subsequent experiments, participants gave informed consent and were paid for their time.

Apparatus and stimuli

Stimuli were presented on a 21″ CRT running at a refresh rate of 75 Hz, driven by a PowerMacintosh G3 computer running MacProbe (Hunt, 1994). Participants were seated at a distance of approximately 57 cm from the monitor, so that 1 cm on the monitor subtended 1 degree of visual angle (°).

Stimuli were presented on a black background, inside a white square border measuring 19° on each side, centred on the screen. A white fixation cross was presented at the centre of the square. The search array consisted of 12 upper-case coloured letters, evenly spaced on the circumference of an imaginary circle of radius 6.5° centred on fixation. Letters were 36-point Helvetica bold font, subtending roughly 1.3°. Letter colours were evenly distributed among red, green, blue, and purple.

The mask array was a circular array of white rings, 2.3° in diameter, each centred on a letter position. Auditory cues consisted of 15-ms beeps. The initial beep was middle C, and pitch increased by a semitone on each frame.

Procedure

There were two conditions, command and anarchy. In the command condition, each trial consisted of a series of 12 frames. Each frame consisted of three phases (see Figure 1, Panel A): a 26-ms prestimulus mask array, followed by a 53-ms search array, followed by a poststimulus mask array for a variable duration. Frame onset was accompanied by a beep. The distractor letters comprising the search array were selected randomly on each frame from the alphabet, omitting "I", "J", and "Y", without replacement. The target letter "Y" was presented only on one randomly selected frame. It replaced the distractor letter presented in the *n*th position, where *n* denotes the frame number, and positions are numbered starting with 12 o'clock and moving clockwise. For example, if the target was presented on the 4th frame, it would appear at the 4th position, corresponding to 3 o'clock (Figure 1, Panel B).

The task was to report the colour of the "Y" (red, blue, green, or purple; all colours were equally likely) by pressing appropriately labelled keys on the computer keyboard. Participants were instructed to shift attention clockwise between each frame (keeping pace with each beep), beginning at the 12 o'clock position. They were told that it would be impossible to determine the colour of the "Y" unless they were attending to

Downloaded By: [Harvard College] At: 19:58 4 March 2009

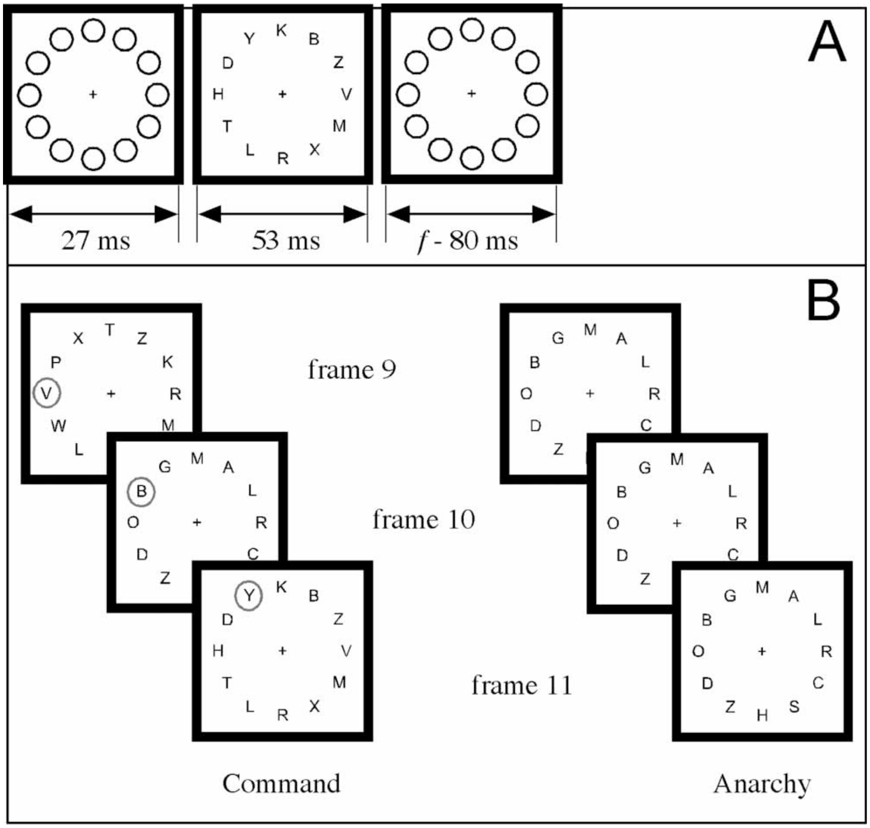

Figure 1. *Command versus anarchy method. Panel A shows the time course of a single frame, from left to right. An initial metacontrast mask array is presented for 27 ms, then replaced by a search array for 53 ms, followed by a mask array for the remainder of the frame duration, f. Minimum frame duration was 80 ms. Panel B shows the search arrays only from three consecutive example frames near the end of a trial. In the command condition (left), the search array changed on each frame, and the target "Y" appears only in Frame 11 at Position 11. The grey circle (not part of the stimulus) denotes the attentional locus of an participant successfully following instructions. In the anarchy condition (right), the search array is the same from frame to frame, and the target is always present. Participants reported the colour of the "Y". The figure is illustrative, and stimuli are not shown to scale. Actual stimuli were presented on a black background. The square border and metacontrast masks were white. Letters could be coloured blue, red, purple, or green.*

the right letter at the right time. They were asked to respond as quickly and accurately as possible.

The anarchy condition was similar, except that the 12 search arrays were identical, and there was always a target present allowing participants to search in any fashion until they found the target. In both conditions, the assumption was that participants could process a single item in the 53-ms exposure period. The variable period of time between stimuli could be used to deploy attention to a new location—in fixed order in the command condition, in any order in the anarchic condition. If participants could process more than one item per frame, then the estimates of time to shift attention would be conservative. Note, however, that there is no reason to assume that participants could process a larger number of items per frame in one condition than in the other.

For each condition, participants first completed 100 trials during which the total frame duration was controlled by a staircase and then 100 trials at a fixed frame duration. The two staircases

Downloaded By: [Harvard College] At: 19:58 4 March 2009

were run first, then the fixed conditions. Within a procedure, participants ran the command and anarchy conditions in counterbalanced order.

In the staircase procedure, the initial frame duration was 263 ms, or 20 refreshes of the monitor. Following a correct response, frame duration was reduced by 13 ms; following an error, it increased by 27 ms. This staircase rule will converge on the frame duration allowing 66.7% accuracy. The frame duration was varied by changing the duration of the poststimulus mask. Note that the minimum possible frame duration was 80 ms (premask + stimulus).

A preliminary asymptote was computed for each condition as the average frame duration for the last 20 trials. The frame duration for the fixed blocks was computed by adding two thirds of the difference in asymptotes to the anarchy condition asymptote. Thus the frame duration in the fixed blocks was identical for both command and anarchy, but different for each participant.

This design provides two ways to compare deployment rates in the command and anarchy conditions. The staircase procedure produces an asymptote for each condition, computed as the average frame duration at the last 10 reversals of the staircase. The fixed-duration blocks produce accuracy measures for each condition.

Results

The central finding of Experiment 1 is that anarchic shifts of attention were much faster than commanded shifts. As shown in Figure 2, the average asymptotic value for the anarchic condition was 85.2 ms while the average in the command condition was 274.5 ms. The difference was significant, $F(1, 9) = 43.6$, $p < .0001$, $\eta = .83$. Since the method imposed a floor of 80 ms on the frame duration, the anarchy rate of 85.2 ms per shift is probably conservative. Based on the staircase results, the average frame duration for the fixed-duration blocks was set to 207.3 ms (*SEM* 20.4 ms). Accuracy in these fixed-duration blocks is shown in the right-hand panel of Figure 2. Accuracy values were arc-sine transformed before analysis; we report the back-transformed means. Accuracy was significantly higher in the anarchy condition than in the command condition, $F(1, 9) = 281.1$, $p < .0001$, $\eta = .97$. Again, the difference between the conditions is conservative because accuracy in the anarchy condition is essentially at ceiling.

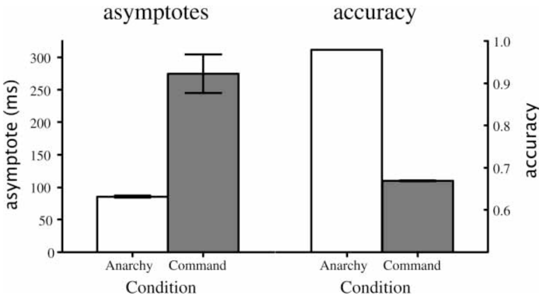

Figure 2. *Results of Experiment 1. Left-hand panel shows staircase asymptotes (average of the last 10 staircase reversals); right-hand panel shows accuracy from fixed frame duration blocks. In this and all subsequent figures, error bars indicate the standard error of the mean.*

Discussion

The results of Experiment 1 demonstrate that voluntary shifts of attention are significantly slower than the priority-driven shifts elicited by the visual search array. When moving attention around the circle in an orderly fashion, participants needed 274 ms per step. Estimating the speed of priority-driven shifts is less straightforward because it depends on assumptions about the nature of search—notably, whether items are sampled from the display with or without replacement. Participants required 85 ms per letter in this condition, which implies a speed of one shift every 85 (assuming a memory-less search in which the same item can be sampled more than once) to 170 ms (assuming some form of memory marks previously selected items so they are not sampled more than once).

At the fixed frame duration, performance in the anarchy condition was near ceiling, indicating that 200 ms was more than enough time for a priority-driven attentional shift. Command performance was 66.7%, higher than might be expected since

Downloaded By: [Harvard College] At: 19:58 4 March 2009

the fixed frame duration was deliberately set lower than the duration estimated to yield 66.7% performance in the staircase blocks. However, participants have two advantages in the fixed-duration blocks. First, these blocks always follow the staircase blocks, so participants are more practised. Second, it is probably easier to adapt to a constant rate for 100 trials than to adjust slightly on each trial, as in the staircase procedure.

Experiment 1 relies on the assumption that observers must be attending to the target location in order to identify the target and thus correctly report its colour. This assumption, in turn, relies on the assumption that the metacontrast masks interrupt processing, such that no information could be accumulated during the masking phase. This seems like a plausible assumption, given that our SOA of 53 ms falls into the range associated with maximal metacontrast masking (for a review, see Enns & Di Lollo, 2000). However, it is known that attention can attenuate metacontrast masking (Boyer & Ro, 2007). Since our observers knew that the target would occur on the right-hand side of the display early in the trial and the left-hand side of the display later in the trial, might they have adopted a strategy of attending broadly to one side of the display or the other, allowing them to accumulate some information about multiple items during the masking interval?

As a direct test, we ran a control experiment in which we presented observers with a single frame (as in Figure 1A) and asked them to report the colour of the Y. The Y was constrained to appear on either the right or the left side of the display, in separate blocks of trials. We then measured accuracy as a function of the duration of the postmask, which was varied over the range from 14 ms to 108 ms to 201 ms. For postmask durations $\leq$ 108 ms, accuracy was consistent with observers processing at most a single item. This means that we may have been overestimating the time required to shift attention in the anarchy condition. At faster speeds, it may have become harder to identify even the one, attended item.

At the longest, 201-ms mask duration, observers did somewhat better, producing results consistent with the processing of two items. This suggests that we may have underestimated the time required to shift attention in the command condition. in the command condition, the staircase asymptote was well over 200 ms. If observers could process two items in that time, then the true shift time would be even slower. Thus, the improvement in performance with postmask duration in this control experiment strengthens the main conclusion of Experiment 1. The difference between the “speed of volition” and the speed of anarchic search may be greater than we estimated.

EXPERIMENT 2: COMMAND VERSUS ANARCHY AND RANDOM ANARCHY

One problem with the design of Experiment 1 is that in command condition, the target is present only on one frame, while in the anarchy condition the target is present throughout the trial. An alternative account of Experiment 1, derived from the work of Palmer and his colleagues (Eckstein, Thomas, Palmer, & Shimozaki, 2000; Palmer, Verghese, & Pavel, 2000) might assume that the two conditions differ not in the rate but in the scope of attentional deployment. As in our story, this account assumes that attention is deployed serially following the tones in the command condition, and that the command asymptote represents the minimum attentional shift time. However, instead of a serial focus of attention moving in a rapid, priority-driven fashion in the anarchy condition, suppose that attention is deployed globally to the display, simultaneously monitoring all 12 locations. Letter detectors operating in parallel at each location accumulate information until one detector signals “Y” with sufficient confidence, at which point attention narrows to that location, and the colour is retrieved. On such an account, the total time available to process the target in the anarchy condition is 53 ms (the actual target presentation time) × 12 frames, or 640 ms, irrespective of the frame rate. If 640 ms is sufficient to produce at least 66.7% accuracy, then the staircase would naturally reduce the frame rate to the floor.

Downloaded By: [Harvard College] At: 19:58 4 March 2009

We find this explanation unlikely, primarily because since the processing of all stimuli is repeatedly interrupted with metacontrast masking on each frame, it would be difficult to accumulate information over time. Nevertheless, in order to further discourage such an explanation, in Experiment 2 we added a variant on the anarchy condition called random anarchy. From previous work, we know that randomly replotting the target at different locations on each frame does not slow the rate of search, as measured by RT × Set Size slopes (Gibson, Li, Skow, Salvagni, & Cooke, 2000; Horowitz & Wolfe, 1998, 2003). However, such a manipulation would surely disrupt accumulation of information at a given location. While a parallel system might be able to accumulate evidence about whether or not a target was present in the display, it would have difficulty determining where the target was, and therefore what colour it was. Therefore, in the random anarchy condition the target location changed unpredictably from frame to frame. Of course, if target locations were selected entirely at random, a participant might choose to focus attention on one particular location and wait for the target to appear (von Mühlenen, Müller, & Müller, 2003). In order to thwart this "sit-and-wait" strategy, the target was presented only at 3 of the 12 possible locations (for an in-depth discussion of such strategies, see von Mühlenen et al., 2003).

A second problem with Experiment 1 is that we don't know how the asymptotic value in the anarchy condition is related to the standard visual search slope measurements. This makes it difficult to compare our results with those of other experiments. Therefore, in Experiment 2 we also added a standard search condition in which we varied set size and measured RT.

Finally, Experiment 1 used a version of search in which the target is identified by one feature (form), but a different feature of the target (colour) is reported (Maljkovic & Nakayama, 1994). For simplicity, in Experiment 2 we changed the task to a two-alternative forced choice (2AFC) form judgement.

Method

Participants

A total of 12 participants were recruited from the Visual Attention Laboratory's paid participant panel.

Apparatus and stimuli

The same apparatus as that used in Experiment 1 was employed here.

Stimuli in Experiment 2 were presented on a grey background, within a 19° square black border. The fixation cross was black, as were the mask rings. Search stimuli consisted of eight letters arranged evenly along an imaginary circle of radius 6.5°. Letters could be either black or white, Helvetica or Times font, and upper case or lower case. Upper-case letters subtended roughly 2.5°, lower-case letters 2.0°. Colour, font, and case were the same for all letters on a given frame, but could vary across frames, as described below.

Procedure

There were four conditions: command, anarchy, random anarchy, and standard search (Figure 3). In the first three conditions, trials consisted of a series of eight frames. As in Experiment 1, each frame started with a 26-ms prestimulus mask array, followed by presentation of the search array for 53 ms, followed by a poststimulus mask array for the remainder of the frame duration, if any. Frame onset was accompanied by a beep, as described in Experiment 1. The distractor letters comprising the search display were sampled without replacement from the alphabet, excluding "I", "J", "N", and "Y".

In all conditions the task was to determine whether there was a "Y" or an "N" present. In the anarchy, command, and random anarchy conditions, the font, case, and colour of the letters alternated from frame to frame. This ensured that there was a stimulus change at every location on every frame, helping to minimize the possibility that the appearance of the target (which replaced a distractor, see below) would draw attention to itself. The font for Frame 1 was randomly Times

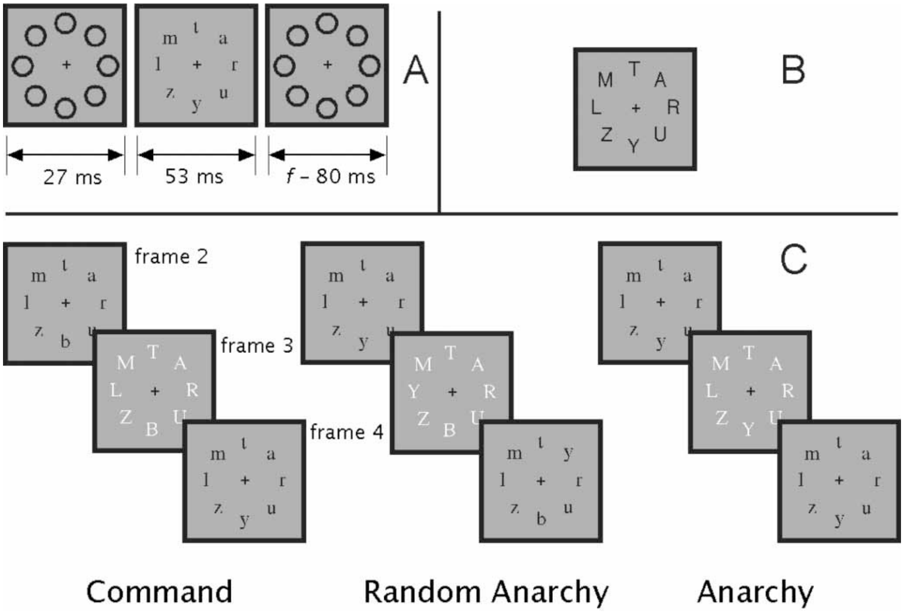

Figure 3. *Methods for Experiment 2. Panel A depicts, from left to right, the time course of a single frame. An initial metacontrast mask array is presented for 27 ms, then replaced by a search array for 53 ms, followed by a mask array for the remainder of the frame duration, f. Minimum frame duration was 80 ms. Panel B shows a sample stimulus from the standard search condition. Panel C shows the search arrays only from three consecutive example frames (2, 3, 4). Case, contrast polarity, and font were selected randomly for Frame 1 and alternated on every frame afterward. In the command condition (left), the target "Y" appears only in Frame 4 at Position 4. In the random anarchy condition (centre), the target is present on every frame but moves among three randomly selected positions. In the anarchy condition (right), the search array is the same from frame to frame, and the target is always present in the same position.*

Downloaded By: [Harvard College] At: 19:58 4 March 2009

or Helvetica, the case for Frame 1 was randomly upper case or lower case, and the colour for Frame 1 was randomly black or white. Subsequent frames alternated between the two possible values for font, case, and colour. In the standard search condition, letters were always Helvetica, upper case, and black.

In the command condition, the letter at each position was the same on each frame except the critical frame. On the critical frame, a target letter ("Y" or "N") was presented in the *n*th position, where *n* refers to the number of the critical frame. Thus, as in Experiment 1, the target was only present for one frame in the command condition. The anarchy condition was identical to the command condition, except that the target was present on every frame. The random anarchy condition was similar to the anarchy condition, except that the target location was not constant from frame to frame. In order to thwart a "sit and wait" strategy, three of the eight locations were selected at random on each trial. The target was randomly assigned to one of these locations on each frame, with the constraint that it not appear at any one location on two successive frames.

The command, anarchy, and random anarchy conditions were all run first using the staircase procedure described in Experiment 1 to control the frame duration. Each participant then ran a block of 100 trials in each condition where the frame duration was fixed at the asymptote of their random anarchy condition (here asymptote was computed using the last 10 reversals of the staircase). Within a procedure, order of conditions was counterbalanced.

Finally, participants also completed two 100-trial blocks of standard search. In the standard

Downloaded By: [Harvard College] At: 19:58 4 March 2009

search condition, a single frame was presented, without masking of any kind, until response. RT was the primary dependent measure in this condition. In one block, set size was 8, as in the other three conditions. In the other block, set size was 4. Order was again counterbalanced.

Results

Data from the command, anarchy, and random anarchy conditions are plotted in Figure 4. The left-hand panel of Figure 4 shows the average staircase asymptotes in the three conditions, and the right-hand panel shows the accuracy data. The basic pattern of results mirrors that in Experiment 1. The commanded deployments of attention are markedly slower than anarchic deployments. For the measures of asymptote in the staircase portion of the experiment, a one-way repeated measures analysis of variance (ANOVA) indicated a significant effect of condition, $F(2,\ 22) = 11.4$, $p < .0005$, $\eta = .51$. The command condition asymptote was significantly slower than both of the anarchic condition asymptotes, which did not differ from one another (Fisher's protected least significant difference, PLSD, critical difference = 63.0 ms; $p < .0005$ for anarchy vs. command, $p < .005$ for random anarchy vs. command, $p > .10$ for anarchy vs. random anarchy). For the accuracy measures in the fixed-duration conditions, again the ANOVA revealed a significant condition effect, $F(2,\ 22) = 111.0$, $p < .0001$, $\eta = .91$. Accuracy in the anarchy condition was superior to that in the random anarchy condition, which was in turn better than command condition performance (Fisher's PLSD, critical difference = 0.4%; all $p < .0001$).

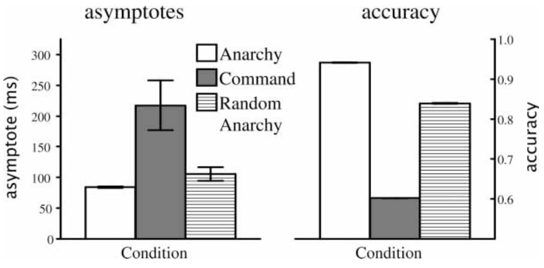

Figure 4. *Results from the staircase portion of Experiment 2. Left-hand panel shows staircase asymptotes as a function of condition, while the right-hand panel shows accuracy in the fixed-duration blocks as a function of condition.*

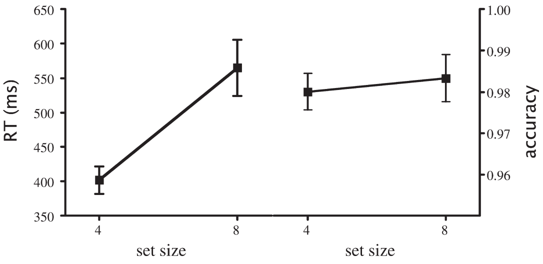

Figure 5. *Standard search results from Experiment 2. The left-hand panel plots RT against set size; the right-hand panel plots accuracy against set size.*

Figure 5 depicts RTs from correct trials and accuracy in the standard search condition as a function of set size. The mean slope was 40.8 ms/item (*SEM* 6.7 ms/item). Accuracy did not differ as a function of set size, $t(11) < 1$.

Discussion

These data provide additional evidence that voluntary shifts of attention are substantially slower than priority-driven shifts. The command and anarchy conditions of Experiment 2 replicated the salient findings of Experiment 1 using a task where participants searched for and reported the form of the target, as opposed to Experiment 1, where the target-defining feature and reported feature were different. The asymptote in the anarchy condition is again at the floor value for our procedure, while the command condition requires more than 200 ms per frame. The conclusions do not change if we compare command to random anarchy, in which participants did not have the advantage of the target staying in a constant position during the trial. In this condition, the frame duration of 100 ms did not differ reliably from the standard anarchy condition. Finally, the standard search condition allows us to compare search rates measured in the usual fashion (RT × Set Size slopes) to the rates measured with our staircase

Downloaded By: [Harvard College] At: 19:58 4 March 2009

procedure. The observed search slope is well within the typical range for such letter form searches (Wolfe, 1998). The slope is roughly half the asymptotic frame duration. In order to infer the deployment rate, we need to make an assumption about how many items need to be sampled, on average, in order to find the target . If standard search required serial, self-terminating search without replacement from the set of items, then the target would be found after sampling slightly more than half the items, so the inferred rate of deployment would be twice the standard search slope (modulated by a correction for errors). This would make the anarchic and standard search estimates very similar. However, current evidence suggests that visual search is self-terminating but samples *with* replacement from the display, probably with some small memory for recent deployments (Horowitz & Wolfe, 2005). If this is the case, then the number of samples needed to find the target is roughly equal to the set size, so the inferred rate of processing is approximately the observed slope (see Horowitz & Wolfe, 2003, for a mathematical derivation of these estimates). On this assumption, the estimate from standard search is probably faster than the estimate from the asymptote of the staircase in the anarchy condition. This difference could be caused by an underestimation of search rate by the slope measurement or by an overestimation by the staircase method. Since the staircase measurement is at the minimum possible value (note the tight error bars in Figures 2 and 4), overestimation is almost certainly a factor. We address this problem in the next two experiments.

As in Experiment 1, the pattern of results in the fixed-duration blocks matched what we would expect based on performance in the staircase blocks. With a duration fixed at the estimated 66.7% level for the random anarchy condition, performance in both random anarchy and anarchy was quite good, slightly but not significantly greater in anarchy than in random anarchy. Performance was well above 66.7% in these conditions, probably due to the factors discussed in Experiment 1. Performance in the command condition, however, was significantly impaired at this frame duration, indicating that ~100 ms was rarely enough time to execute a voluntary shift of attention, even given substantial practice and a constant frame rate.

EXPERIMENT 3: VISUAL PACING

One difference between the anarchic and commanded conditions of the previous experiments is that the command conditions employed an auditory cue to establish the rate at which attention was to be deployed. The goal, especially in the constant-rate versions of the experiments was that participants would learn to generate a self-paced series of attentional shifts. However, it is possible that the auditory cue introduced cross-modal attentional demands that may have artificially lengthened the required frame duration in the command conditions of Experiments 1 and 2. Given that participants were told that the beeps were an important part of the task during the command conditions, and not the anarchy conditions, they might have devoted more resources to monitoring the auditory modality during the command conditions, which might in turn account for the slower pace of command attentional shifts. Experiment 3 replaced the beeps with a visual pacing cue, a "clock hand" that swept around the display, indicating where the participant should attend.

We also sought to lower the floor in the anarchy condition in order to make it possible to measure faster rates of deployment. The minimum frame duration in the command condition was set to 53 ms (26 ms of stimulus + 26 ms of mask). In the anarchy condition, we dispensed with the interframe masks and simply displayed the search array for a single uninterrupted period lasting 10 times the frame duration; thus we could reduce the minimum staircase value in this condition to one refresh (13 ms).

Finally, we varied set size in the anarchy condition. This manipulation was intended to help estimate the time between successive deployments of attention. Simply dividing the exposure

Downloaded By: [Harvard College] At: 19:58 4 March 2009

duration by the number of stimuli is inadequate, because some set-up time is typically involved in initiating visual search. When set size is varied in a visual search experiment, the intercept of the RT × Set Size slope (the time required even if no items are scanned) is always significantly nonzero. Some of this time is poststimulus decision and response time, of course, but some initial preprocessing of the display is also required. For instance, in speed–accuracy trade-off function studies of search, accuracy does not depart from chance for at least 300 ms after stimulus onset (McElree & Carrasco, 1999). Therefore, dividing exposure duration by the number of stimuli (apart from assumptions about sampling) will always underestimate the search rate, thus overestimating the shifting time.

Method

Participants

A total of 10 participants from the Visual Attention Laboratory paid participant panel participated in this experiment.

Apparatus and stimuli

Stimulus presentation was controlled by a PowerMacintosh G3 running MATLAB 5 and the Psychophysics Toolbox, Version 2 (Brainard, 1997; Pelli, 1997).

Stimuli were black presented on a white background. Search arrays consisted of 6 or 10 letters arranged evenly along an imaginary circle of radius 6.5°. Letters were shown in 48-point Arial font and subtended approximately 2°. Metacontrast mask rings subtended 3°. The clock hand was a black line 2.5° long and 0.11° in width. The clock hand always originated at the centre of the display. A black fixation dot subtending 0.42° was always present at the centre of the screen.

Procedure

There were two conditions, command and anarchy. in the command condition, each trial consisted of 10 frames. Each frame began with a presentation of the search display, which consisted of 10 distractor letters drawn from the alphabet excluding "J" and "P". These distractor letters were identical from frame to frame. One frame was randomly selected as the critical frame; on this frame, the distractor letter in the *n*th position (where *n* corresponds to the serial position of the frame) was replaced by one of the two target letters, "J" or "P", selected at random. After 26 ms, the search array was replaced by the metacontrast mask array for the remainder of the frame duration.

At the beginning of each trial, the clock hand appeared, pointing outward from the fixation dot towards the letter in the 12 o'clock position (Figure 6). During each frame, the clock hand moved through 36 radial degrees, so that it was always pointing towards the position to be attended. The clock hand advanced with each refresh of the monitor; the radial displacement per refresh depended on the frame rate. Participants were told that the clock hand would reliably indicate the location of the target, when it appeared, and that they should move their attention with the clock hand.

In the anarchy condition, a single frame was presented for 10 times the frame duration, followed by the metacontrast mask array. The target was continuously present in a randomly selected location during this frame. The clock hand was present and moved according to the same rules as those in the command condition, except that the initial orientation was random, and the position of the clock hand was uninformative with respect to the location of the target. Participants participated in two anarchy blocks, one with set size 10 and one with set size 6. The motion of the clock hand was the same for both set sizes.

As in previous experiments, participants first ran through each condition (command and the two anarchy set size blocks, in counterbalanced order) under the staircase procedure to estimate the frame duration producing 66.7% accuracy. The minimum possible staircase value for the command condition was two refreshes (26 ms) of the metacontrast mask, for a total of four refreshes (53 ms) per frame; the minimum value for the

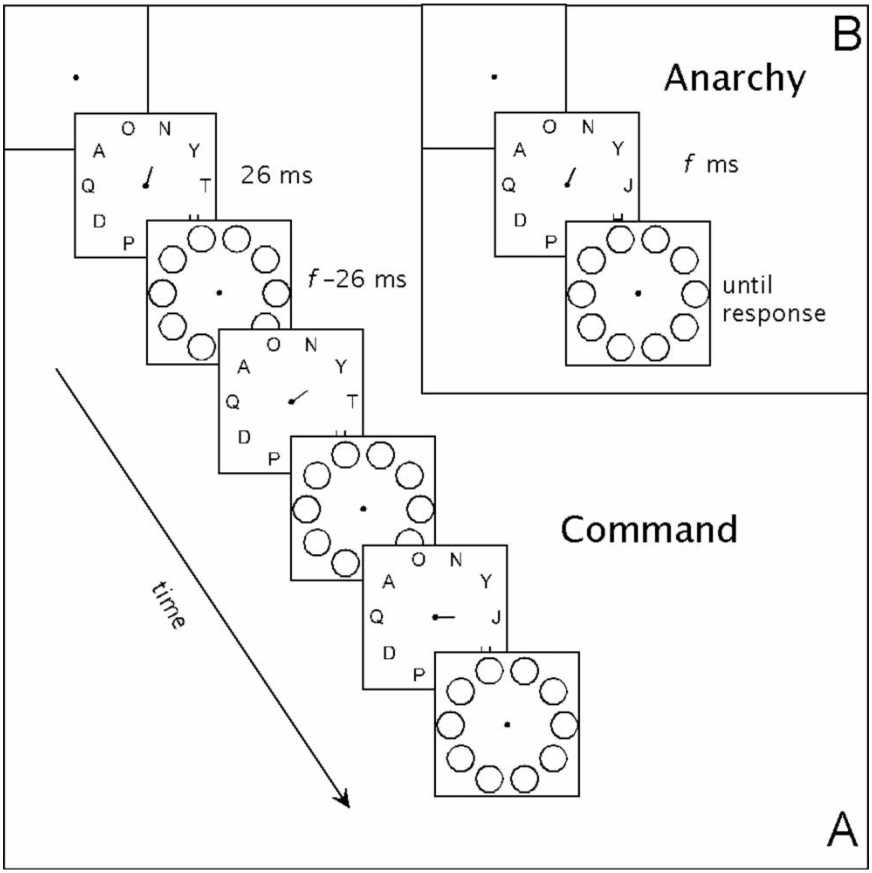

Figure 6. *Methods for Experiment 3. Panel A illustrates a series of three frames from the command condition, moving from upper left to lower right. The search array is presented for 26 ms, followed by the metacontrast mask for the remainder of the frame duration, f. Here the target "J" appears on the third frame. The clock hand always points to the potential target location. Panel B illustrates a trial from the anarchy condition, in which a single frame, always containing the target, is presented for the entire frame duration, followed by the metacontrast mask.*

Downloaded By: [Harvard College] At: 19:58 4 March 2009

anarchy condition was one refresh. We also imposed a ceiling of 80 refreshes (1,067 ms). Following the staircase version, the conditions were run again in a block where the frame duration was fixed. In this experiment, we fixed the frame duration at each participant's asymptote for the appropriate condition.

In all conditions, the participant's task was to report whether a "P" or "J" was present in the display. Participants were instructed to respond as quickly and accurately as possible; if the participant responded before the end of the trial, the display was terminated. Each block consisted of 30 practice trials, followed by 100 experimental trials.

Results

In this experiment, the asymptotic values shown in the left-hand panel of Figure 7 provide the data to differentiate command and anarchy conditions. As before, these results show commanded deployments to be much slower than anarchic. A one-way repeated measures ANOVA revealed a significant effect of condition, $F(2, 18) = 13.9$, $p < .01$, $\eta = .61$. Post hoc comparisons indicated

Downloaded By: [Harvard College] At: 19:58 4 March 2009

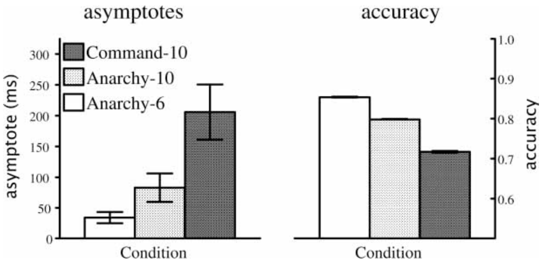

Figure 7. *Results from Experiment 3. Left-hand panel shows staircase asymptotes as a function of condition and set size, while the right-hand panel shows accuracy in the fixed-duration blocks.*

that both anarchy conditions produced significantly lower asymptotes than the command condition ($p < .0001$ for set size 6, $p < .005$ for set size 10), but there was no reliable difference between the two set sizes (Fisher's PLSD, critical value = 70.7 ms).

Accuracy in the fixed frame duration blocks are shown in the right-hand panel of Figure 7. Recall that, in this experiment, the asymptote for each staircase condition was used as the duration in the corresponding fixed condition. Thus, the fixed durations used for the anarchic conditions were much shorter than the fixed duration in the command condition. Nevertheless, an ANOVA revealed a significant effect of condition for the fixed-duration results, $F(2,\ 18) = 6.4$, $p < .01$, $\eta = .42$. Post hoc comparisons indicated that performance in both anarchy conditions was superior (marginally so for set size 10) to that in the command condition (Fisher's PLSD, critical value = 1.0%; $p < .005$ for set size 6, $p = .06$ for set size 10), but there was no difference between the two set sizes ($p > .10$).

Discussion

Command condition performance here is quite similar to that in Experiment 2, indicating that the cross-modal demands of previous experiments were not inflating the rate estimates. Of course, this experiment does not eliminate the possibility that any external pacing cue will lead to interference; that possibility is addressed in Experiment 4.

The second goal of this experiment was to get a better estimate of the anarchic search rate by reducing the staircase floor and introducing a set size manipulation. As noted earlier, the standard method from visual search experiments is to infer a search rate from the slope of the RT × Set Size function. Here we would infer a rate from the mean Asymptote × Set Size function, which was 12.3 ms/item (*SEM* 6.2 ms/item). Depending on one's position on this issue of memory in visual search, this implies an interval of 12.3 to 24.6 ms between shifts of attention (for search without or with memory, respectively). Therefore, even the longest estimated shift time for priority-driven shifts of attention is an order of magnitude faster than that for voluntary shifts of attention.

EXPERIMENT 4: STATIC DISPLAYS

In Experiment 4, we sought converging evidence from a different paradigm. In Experiments 1–3, we forced the rate of deployment of attention, and we measured whether or not participants could keep up with the experimentally imposed rate. It could be that this coercion crippled a normally swift mechanism for voluntary deployments of attention. In Experiment 4, we estimated the unforced rate of attention shifting in command and anarchic conditions. This method uses a static display in which the target was always present. We accomplished this by defining the target as the first mirror-reversed letter clockwise around a circular array and asking participants to report the target's identity. Since there were several mirror-reversed letters in the array, participants were obliged to shift attention in the proper order. We can use the function relating RT to ordinal position around the circle to compute a rate in the same way as we use the RT × Set Size function. For example, if it took an average of 800 ms to find a target in Ordinal Position 2 and 1,200 ms to find the target in Position 4, we would infer a voluntary shifting rate of 200 ms/shift. A standard visual search for a mirror-reversed letter served as the anarchy control.

Method

Participants
A total of 10 participants from the Visual Attention Laboratory's paid panel participated in this experiment.

Apparatus and stimuli
The same apparatus as that used in Experiment 3 was employed here.

Stimuli were green letters, "S" and "P", presented in 48-point Arial font on a black background, subtending ~1.8°. The fixation cross was the " + " character, also in 48-point Arial font. A white disc, 2.5° diameter, served as the cue. Letters and cue were arranged evenly around an imaginary circle of radius 7°.

Procedure
Each trial began with presentation of the fixation cross in the middle of the screen and the onset of a 100-ms 650-Hz warning beep. After 500 ms, the search array was presented until response. The search array consisted of 7, 9, or 11 letters and one cue disc. Set size was varied randomly within a block of trials. Participants responded by pressing the quote key with their left hand if the target was an "S", or the "a" key with their right hand if the target was a "P". Participants were asked to respond as quickly and accurately as possible; RT was the primary dependent variable.

There were two conditions, command and anarchy. In the command condition, the target was defined as the first mirror-reversed letter, counting clockwise from the white disc. The position of the target relative to the cue was constrained, such that on 50% of the trials the target was presented in the first four positions ("near" positions), and on the remaining trials in later positions ("far" positions). We anticipated that participants would be less likely to command their attention in an orderly fashion the further they got from the cue position. That is, after making three or four shifts of attention, we suspected that participants might "fall off course". Accordingly, we placed a high percentage of targets in the near positions to make sure there were enough data at the positions in which we expected participants to "stay on course". Letters were randomly distributed between S and P, and letters after the target position in the sequence were randomly distributed between normal and mirror-reversed letters (Figure 8).

In the anarchy condition, the target was defined as the only mirror-reversed letter in the array. Thus, no particular order of search was required, and participants could adopt any method that produced the answer.

Participants completed 446 trials in the anarchy condition (50 practice trials, and two blocks of 196 test trials) and 450 trials in the command condition (50 practice trials, and two blocks of 200 test trials). The order of conditions was counterbalanced across subjects.

Data analysis
We removed all RTs < 200 ms and > 10,000 ms and analysed only correct RTs. Trimming the RTs removed <1% of the data. Furthermore, since Positions 8–11 only occurred at the larger set sizes, we restricted analysis to RTs from

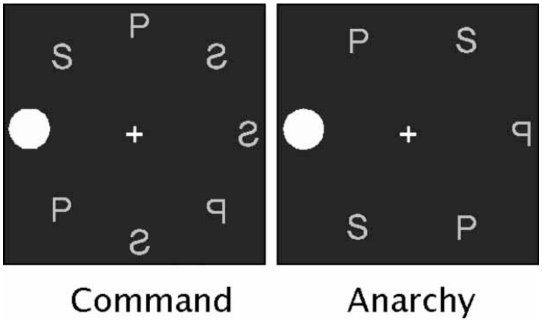

Figure 8. *Stimuli for Experiment 4. An example stimulus from the command condition is shown on the left. The target is the first mirror-reversed letter clockwise from the cue disc, in this case an "S". An example stimulus from the anarchy condition is shown on the right. Here the target is the only mirror-reversed element in the display, a "P".*

Downloaded By: [Harvard College] At: 19:58 4 March 2009

Downloaded By: [Harvard College] At: 19:58 4 March 2009

Positions 1–7. Slopes were tested against 0 by *t* test.

Results

Figure 9 shows RT data from Experiment 4 as a function of set size (left-hand panel) and position (right-hand panel). In the left-hand panel, we see a steeply increasing RT × Set Size function for the anarchy data, with a slope (58.5 ms/item; significantly nonzero, $p < .0001$). Command RTs, on the other hand, actually decreased with set size (−44.3 ms/item, significantly nonzero, $p < .001$).

The opposite pattern can be seen in the right-hand panel, where RTs are plotted as a function of position. Here the anarchy data yielded a weak RT × Position function (14.0 ms/position, significantly nonzero, $p < .05$), while the command RT × Position function rose steeply (195.9 ms/position, significantly nonzero, $p < .00001$).

These impressions are confirmed by ANOVA. For the anarchy RTs, there was a main effect of set size, $F(2, 18) = 27.3$, $p < .0001$, $\eta = .75$, but no effect of position, $F(6, 54) = 1.4$, $p > .10$, $\eta = .14$, nor any interaction, $F(12, 108) = 1.0$, $p > .10$, $\eta = .10$. For the command RTs, ANOVA showed a significant effect of position, $F(6, 54) = 94.2$, $p < .0001$, $\eta = .91$, but no main effect of set size, $F(2, 18) < 1$, $\eta = .03$. The interaction was marginal, $F(12, 108) = 1.8$, $p = .056$, $\eta = .17$, indicating a trend towards shallower position slopes at greater set sizes. Observers may have deviated from a strictly ordered search on some trials, and this was more likely when there were more items. This pattern also explains the trend towards a negative set size slope in the command condition: A reduced position slope means there were fewer long RTs at greater set sizes than at lower set sizes. However, keep in mind that this is a small effect: Position slopes were 206.4 ms/position for set size 7, 191.0 ms/position for set size 9, and 177.7 ms/position for set size 11.

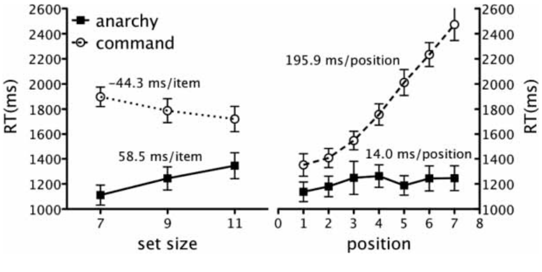

Figure 9. *Results from Experiment 4. In the left-hand panel, RTs are plotted as a function of set size for the anarchy and command conditions. Note that these data only include targets at Positions 1–7. The right-hand panel plots RTs as a function of position (clockwise from the cue). Filled squares denote anarchy condition data, open circles command condition data.*

Participants made significantly more errors, $t(9) = 4.9$, $p < .001$, in the command condition (7.7%, *SEM* 0.8%) than in the anarchy condition (3.7%, *SEM* 0.2%). Presumably, this reflects the presence of false targets (other inverted letters) in the command condition but not in the anarchic condition. Since RTs in the command condition were also slower than those in the anarchy condition, this reflects a speed–accuracy covariance.

Discussion

Participants appeared to be following directions. In Figure 9, the anarchy functions are fairly flat with respect to position, but separated by set size, while the command RTs increase linearly with position, but the set size functions overlap. In the anarchy condition, participants did not search systematically, or at least they did not begin their search at the cue and move anticlockwise. In the command condition, position was the primary predictor of RT, indicating that participants were searching in the prescribed order, though there may have been some deviations at higher set sizes.

What do the slope values tell us? In the anarchy condition, the RT × Set Size slope is our index of search rate. Depending on sampling assumptions, we obtain an estimate of 58.5–117.0 ms per shift. In the command condition, the RT × Position slope is a direct estimate of the search rate. In this experiment, the rate of 195.9 ms/item is comparable to that observed in the previous experiment and markedly longer

Downloaded By: [Harvard College] At: 19:58 4 March 2009

than the anarchic estimate. If observers were occasionally deviating at higher set sizes, then we may be slightly underestimating the command rate.

There are two potential problems with this experiment. First, observers might have simply been making eye movements, rather than shifting attention as directed. This possibility is dealt with in Experiment 5. Second, there is a confound between the command and anarchy conditions, in that the command displays included multiple mirror-reversed letters, while in the anarchy condition the target was the only mirror-reversed letter. Since the task was to report the identity of a mirror-reversed letter, observers might have experienced interference from response-incompatible mirror-reversed letters in the command condition; for example, if the target were a mirror-reversed P, the presence of mirror-reversed Ss may have slowed responses. Of course, there were (on average) an equal number of response-compatible letters in the command displays, but these effects may not be symmetrical. Thus, it is possible that RTs in the command condition are artefactually elevated relative to the anarchy RTs.

However, it is important to remember that our argument is built not on overall RT differences but on the differential effects of position and set size on RT. RTs increase strongly with position in the command condition, but not in the anarchy condition, while the opposite pattern holds for the effects of set size. The number of (potential) response-incompatible mirror-reversed distractors actually decreases with position, since all mirror-reversed letters must be between the target and the cue position, yet RT increases with position. Meanwhile, the number of response-incompatible distractors ought to increase with set size. Furthermore, flanker effects are known to increase as the spacing between the target and flankers decreases (Miller, 1991). Thus, if there were net flanker interference in this experiments, we should see an increase in RT with set size in the command condition, which we do not see. If there were net flanker facilitation, we should see a negative set size slope in the command condition. While we do see such a trend in Experiment 4, it does not replicate in Experiment 5 (see upper left panel of Figure 10), or in two additional control experiments not reported here, all of which replicate the position effect with almost precisely the same slope.

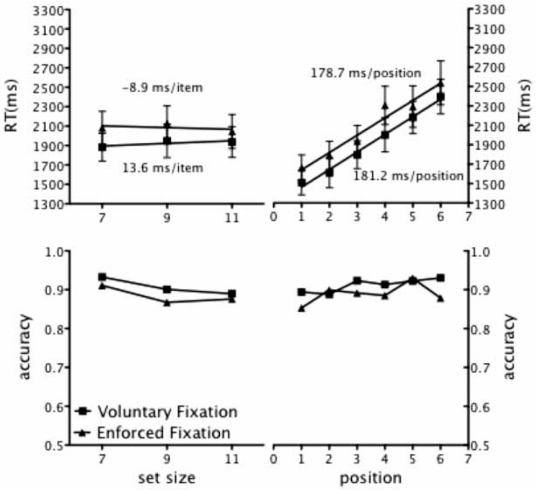

Figure 10. *Results from Experiment 5. Here we have included only data from trials where the target appeared in Positions 1–6. The upper panels show RTs, the bottom panels accuracy. Data are plotted as a function of set size on the left and of position on the right. Voluntary fixation condition data are shown as squares, and enforced fixation data are shown as triangles. RT slopes are indicated near the appropriate curves.*

EXPERIMENT 5: EYE MOVEMENT CONTROLS

The rate at which participants can voluntarily shift attention, as measured in Experiments 1–4, is similar to the rate of saccadic eye movements (4–5 Hz; Carpenter, 1977). This may have theoretical implications, an issue that we address in the General Discussion. However, the correspondence leads to the suspicion that we may have simply been measuring eye movements rather than shifts of attention. Attention and eye movements are tightly linked (T. Moore, Armstrong, & Fallah, 2003). It is all but impossible to move the eyes without deploying attention (Kowler, Anderson, Dosher, & Blaser, 1995). Perhaps attention

Downloaded By: [Harvard College] At: 19:58 4 March 2009

cannot be commanded without also commanding the eyes. In the previous experiments, participants were instructed to fixate, but we did not monitor eye movements. Participants may have simply moved their eyes along the circle. Since it is the command condition in which we suspect eye movements might be occurring, we replicated the command condition of Experiment 4 while monitoring eye movements. In the voluntary fixation condition we simply instructed participants to fixate, so we could measure the rate of eye movements that participants normally engaged in during this task. In the enforced fixation condition, an aversive tone played whenever the eye tracker detected that eye position was outside of the bounding box of the fixation cross; we expected that this condition would minimize saccades.

Method

Participants

A total of 13 participants were recruited from the volunteer pool of the Computational Visual Cognition Laboratory at the Massachusetts Institute of Technology.

Apparatus and stimuli

Stimuli were presented on Sony Trinitron CRT monitor set to 1,024 × 768 pixel resolution and a 75-Hz refresh rate. Stimulus presentation was controlled by a Dell Optiplex Gx270 computer running Matlab 7.1 and the Psychophysics Toolbox, Version 3. Stimuli were the same as those in Experiment 4. The pixel size of stimuli were adjusted to match the visual angles of Experiment 4, given that participants were sitting 75 cm from the screen in Experiment 5. Eye position data were recorded by an ISCAN RK-464 eye tracker sampling at 240 Hz.

Procedure

Procedure was as described in Experiment 4, with the following exceptions. First, participants only ran in the command condition. Second, the direction in which participants were asked to shift attention around the circle, clockwise or anticlockwise, was varied in blocks. Third, and most importantly, we monitored eye position and manipulated whether or not participants could move their eyes.

The eye tracker was calibrated for each participant using a 5-point calibration screen. Participants were seated at a distance of 75 cm from the monitor and used a chin rest to maintain head position.

There were two fixation conditions. In both conditions, the computer instructed participants to fixate on the central cross before initiating each trial. In the enforced fixation condition, the computer emitted a 1,000-Hz tone for 100 ms if gaze position wandered outside of the 1.3° bounding rectangle of the fixation cross. In the voluntary fixation condition, we simply monitored eye movements.

There were four blocks of trials, one for each combination of fixation condition and shift direction. Condition order was counterbalanced across participants. Each block began with 5 practice trials, followed by four groups of 50 experimental trials, or 205 trials per block, for a total of 820 trials per participant. There were six deviations from this protocol. The first participant ran 405 trials per block, except for the clockwise enforced fixation block, in which he ran 210 trials. Another participant completed only 110 trials in the clockwise enforced fixation block. The initial block of trials for 3 participants was discarded because of technical issues with the eye tracker or the presentation computer; these participants then provided full datasets. One participant ran 15 practice trials for her first block.

Eye position analyses

Saccades were identified using an acceleration-based algorithm derived from Araujo, Kowler, and Pavel (2001). A saccade onset was identified when the velocity difference between two successive data points exceeded 6°/s. Saccade termination was identified when the acceleration threshold was crossed in the other direction. The origin of the saccade was defined as the eye position at saccade onset. The endpoint of the saccade was the eye position during the saccade that was most distant from the origin. Saccades

Downloaded By: [Harvard College] At: 19:58 4 March 2009

within the fixation box were considered to be fixational eye movements (Martinez-Conde, Macknik, & Hubel, 2004).

Results

A total of 2 participants were excluded for excessive error rates. RTs < 200 ms or $> 10{,}000$ ms were removed: a total of 62 trials or 0.67% of the data. Since the target could not occur in Positions 7–10 for set size 7, only data from Positions 1–6 were included in the figures and analyses. Accuracy was arc-sine transformed before analysis.

The upper left panel of Figure 10 plots mean RT as a function of set size and fixation condition. RTs were not influenced by set size, $F(2, 20) = 2.5$, $p = .10$, $\eta = .20$, or condition, $F(1, 10) = 2.3$, $p = .16$, $\eta = .19$, nor was there an interaction, $F(2, 20) < 1.0$, $\eta = .08$. In contrast, the upper right panel of Figure 10 shows a strong linear dependence on position, $F(5, 50) = 77.5$, $p < .001$, $\eta = .89$, with again no effect of condition, $F(1, 10) = 2.5$, $p = .15$, $\eta = .20$, nor an interaction, $F(5, 50) = 1.2$, $p = .34$, $\eta = .10$.

Accuracy data are plotted in the lower panels. The lower left panel illustrates that accuracy declined with set size, $F(2, 20) = 14.0$, $p < .001$, $\eta = .58$, but did not differ between the fixation conditions, $F(1, 10) = 2.6$, $p = .14$, $\eta = .21$, nor was there an interaction, $F(2, 20) < 1.0$, $\eta = .04$. The lower right panel shows that accuracy declined with position, $F(5, 50) = 4.9$, $p = .001$, $\eta = .33$, but again did not differ between fixation conditions, $F(1, 10) = 2.5$, $p = .15$, $\eta = .20$. There was a subtle interaction, $F(5, 50) = 2.5$, $p = .04$, $\eta = .20$, such that accuracy improved over the first three positions for the voluntary condition, but was more constant for the enforced condition (except for Position 5).

Participants made few saccades outside of the fixation box in this experiment. The left panel of Figure 11 shows the proportion of correct trials without any saccades (not counting fixational eye movements, see Method) for the two fixation conditions. The enforced fixation condition was somewhat better on this measure, as expected, though the difference was not significant, $F(1, 10) = 3.1$, $p = .11$, $\eta = .23$. The middle and right panels in Figure 11 replot the RT data, now collapsed across fixation condition, as a function of whether or not a saccade was detected. Trials with saccades have much longer RTs than nonsaccade trials: set size analysis, $F(1, 10) = 63.5$, $p < .001$, $\eta = .864$; position analysis, $F(1, 10) = 49.9$, $p < .001$, $\eta = .83$. Neither type of trial shows a set size effect: main effect, $F(2, 20) = 1.8$, $p = .19$, $\eta = .15$; interaction, $F(2, 20) = 1.5$, $p = .25$, $\eta = .13$. Most importantly, there is an RT × Position effect, $F(5, 50) = 59.6$, $p < .001$, $\eta = .86$, for both no-saccade and saccade trials—interaction, $F(5, 50) = 1.1$, $p = .37$, $\eta = .10$—indicating that

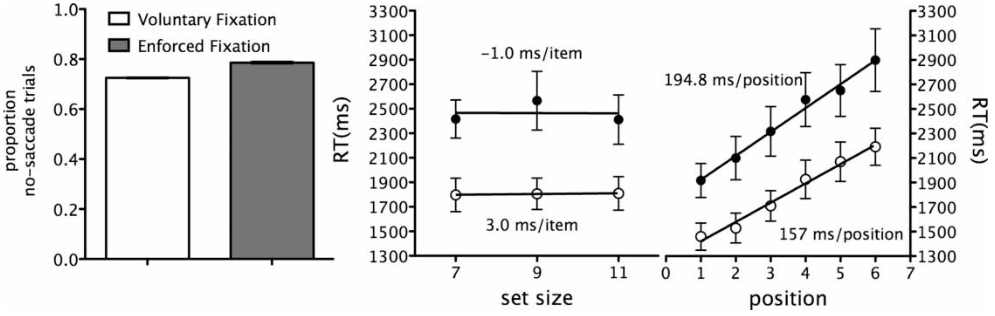

Figure 11. *Experiment 5 data filtered on eye movements. Left panel shows the proportion of trials without any nonfixational eye movements in the voluntary (white bar) and enforced (grey bar) fixation conditions. The other two panels show RTs collapsed across fixation condition for trials without eye movements (open circles) and with eye movements (filled circles). Middle panel shows RT as a function of set size; right panel shows RT as a function of position.*

Downloaded By: [Harvard College] At: 19:58 4 March 2009

the position effect does not depend on participants making eye movements. However, there is some contribution of eye movements: While the slope for the saccade trials (194.8 ms/position) is comparable to that obtained in Experiment 4, when we look at only trials with no saccades, the slope is shallower (157 ms/position), an observation confirmed by the interaction of the linear contrast on position with the saccade variable, $F(1, 10) = 93.2$, $p < .001$, $\eta = .90$.

Discussion

A straightforward explanation for the differences in search rate between command and anarchy conditions is that when instructed to search stimuli in a particular order (e.g., around the circle), observers move their eyes from item to item. This hypothesis is clearly false. Even when eye movements were allowed, observers fixated on a large majority of trials. There was some contamination of the search rate by eye movements, since the position slope was shallower, and RTs substantially faster, on trials without eye movements. This is consistent with data from Zelinsky and Sheinberg (1997), who demonstrated that eliminating eye movements (in what we would term anarchic search) yielded faster RTs and somewhat shallower slopes, but did not change the overall pattern of behaviour. Also note that the negative trend in set size slopes observed in Experiment 4 did not recur in this experiment.

Combined, these results argue that the observed rate of attentional shifting in the command condition (here and in Experiment 4, see also Experiment 3 of Kane, Poole, Tuholski, & Engle, 2006) is not due to systematic eye movements. Instead, it seems likely that voluntary shifts of attention are slow.

GENERAL DISCUSSION

Voluntary attention

When participants were asked to make multiple sequential attention shifts in a known, prespecified direction, they require 150–300 ms per shift. This estimate is consistent across the variety of search tasks and methods used in the five experiments presented here, though Experiment 5 suggests that when the influence of eye movements are removed, the true value is toward the lower end of this range. In comparison to the cueing and attentional gating experiments described in the introduction, we minimized the degree to which participants had to interpret a cue in order to determine when or how to shift attention. A pacing cue was available in Experiments 1–3. However, participants always knew that they would be required to shift attention clockwise, beginning from a known start position. The very first interval provided sufficient information for participants to determine the timing of attention shifts. Nevertheless, one might object that participants had difficulty adjusting to a rate that varied from trial to trial and thus relied on the pacing cue. Processing the cue might have taken some capacity and slowed shifts of attention. There is little support for this hypothesis in the fixed-duration blocks. In the fixed conditions of Experiment 3, for example, the duration was constant for 100 trials. Even so, participants performed at 70%, marginally above the 66.7% level obtained with variable pacing in the staircase portion of the experiment. Moreover, the timing of shifts is similar in Experiment 4, where participants were free to set any pace they chose for voluntary deployments of attention.

Recall that, in standard visual search, the time between successive shifts of attention seems to be less than the time required to identify an object. Perhaps, in commanded search, participants do not shift until they know whether or not they are attending to the target at the current focus of attention. Metaphorically, these participants would be allowing only one car at a time into the car wash. On this assumption, estimates of the volitional shift time would be contaminated by the time required to identify a letter. We think this is unlikely. First, when others have monitored eye movements in visual search tasks, they have found that, participants tend to make one or two fixations after fixating the target. This suggests that they did not wait for stimulus identification to complete before commencing the next saccade (Korner & Gilchrist, 2008). There is no reason to suspect that our attention shifts would be any different. Second, in a different set of studies, we had participants shifting attention among empty place holders, waiting for the flash of a target letter (Horowitz, Holcombe, Wolfe, Arsenio, & DiMase, 2004). In these experiments, there is no letter to identify on most shifts of attention. Nevertheless, this version of the commanded condition produces, if anything, somewhat longer estimates of the time needed to shift attention. Therefore, we are confident that we have obtained a relatively pure measure of the time needed to voluntarily shift attention.

As noted in the introduction, cueing and attentional gating studies also estimate the time to shift attention. Those studies agree on a 300–400-ms interval between shifts of attention. This is somewhat slower than our ≈200 ms though it is in the same ballpark. We suggest that some of the discrepancy is due to the time necessary to decipher the cue in these paradigms. Indeed, both Peterson and Juola (2000) and Shih and Sperling (2002) have made similar suggestions with respect to the attentional gating results, using very different arguments. As far as we know, cue interpretation time has not been studied with the traditional cueing paradigms. On the other hand, as noted in the discussion of Experiment 1, our method might be slightly underestimating volitional shift times.

Attention and eye movements

It would not be surprising if the mechanisms controlling volitional deployments of the eyes were closely related to those producing volitional movements of the eyes. The inferred rate of voluntary attention is close to the rate of eye movements (Carpenter, 1977), and it is widely agreed that the neural mechanisms used to plan shifts of attention overlap substantially with those used to generate saccadic eye movements (Corbetta, 1998; Khurana & Kowler, 1987; Kowler et al., 1995; Nobre, Gitelman, Dias, & Mesulam, 2000; Schall, 2004). We can describe the attentional shifts elicited by the command condition as "attentional saccades" (Horowitz, Holcombe, Wolfe, Arsenio, & DiMase, 2004).

There also seem to be two modes of oculomotor control. It is known that eye movements can be either reflexive or voluntary (Machado & Rafal, 2000). In many circumstances, it can be shown that eye movements are driven by salience (Itti & Koch, 2001) or by a priori scene information (Oliva, Torralba, Castelhano, & Henderson, 2003; Peterson & Kramer, 2001). Such eye movements can be modelled as random among salient loci (Brockmann & Geisel, 2000), or at least fairly unsystematic (Gilchrist & Harvey, 2000). However, we can certainly direct our eyes voluntarily, by an act of will, to locations that may not be salient at all; the ability of psychophysical observers to maintain fixation is the most extreme example of this. We can make ordered sequential eye movements in order to read (Fischer, 1999). Even in visual search, there is evidence that observers can shift the eyes systematically (Gilchrist & Harvey, 2006; Peterson, Kramer, Wang, Irwin, & McCarley, 2001). Theeuwes and colleagues (Theeuwes, Kramer, Hahn, & Irwin, 1998) argued that there are separate neural mechanisms responsible for programming voluntary and stimulus-driven saccades.

However, it is important to keep in mind that shifts of attention and eye movements, while related, are separable behaviours. In some cases, attention precedes an eye movement to a location (Fischer, 1999; Peterson, Kramer, & Irwin, 2004). In other cases, attention and the eyes can move in different directions (Murthy, Thompson, & Schall, 2001). In our work, Experiment 5 shows that, even if they are drawing on similar underlying mechanisms, oculomotor saccades can be suppressed while observers make "attentional saccades".

Attentional blink and attentional dwell times

As noted in the introduction, there has been a tendency to use "attentional dwell time" to mean the

Downloaded By: [Harvard College] At: 19:58 4 March 2009

Downloaded By: [Harvard College] At: 19:58 4 March 2009

same thing as "shift time". We think that this is unwise. The idea of a long attentional dwell time gains its strongest support from "attentional blink" studies (Raymond, Shapiro, & Arnell, 1992) using the Rapid Serial Visual Presentation (RSVP) procedure. In RSVP, stimuli are presented sequentially at fixation. If observers are asked to report two prespecified targets from an RSVP stream (e.g., two digits in a stream of letters), accuracy in reporting the second target (T2) is usually impaired for 300–700 ms after the first target (T1; Broadbent & Broadbent, 1987; Chun & Potter, 1995; Kristjansson & Nakayama, 2002). A full stream is not required; Ward and his colleagues (Ward et al., 1996) have replicated the basic attentional blink effects with a skeletal RSVP display consisting of just T1 and T2. Duncan et al. (1994) used a modified skeletal RSVP paradigm in which T1 and T2 were presented sequentially at different locations and observed a similar, long-lasting interference on T2 report. They have used this and similar findings (Ward et al., 1996; Ward, Duncan, & Shapiro, 1997) to claim that attention shifts are quite slow, at least twice as slow as the times we report here. Moore and colleagues (C. M. Moore et al., 1996) have disputed these estimates, arguing that the effects of masking exaggerate the dwell times.

However such "attentional dwell time" studies may not be measuring the time course of attention shifts at all. If 300–700 ms were required to shift from one stimulus to the next, then it is hard to see how any task could be accomplished at RSVP rates of one item every 100 ms or so. Accordingly, recent theoretical accounts explain the T2 impairment as a consequence of the attentional system's effort to segment episodes in time. The ST^2 architecture of Bowman and Wyble (2007) and the "boost and bounce" theory of Olivers and Meeter (2008) differ in many of their particulars. However, in both theories, the observed attentional blink/dwell time results reflect the combined effect of excitatory and inhibitory "microdynamics" operating at time courses of less than 100 ms. The appearance of the target triggers a transient attentional gate or window. A distractor entering during this window triggers an inhibitory reaction (bounce), which impairs processing of T2. These models explain the attentional dwell time results without assuming slow attentional switching. Indeed, the Olivers and Meeter model explicitly assumes voluntary ("endogenous") switching times of around 200 ms and priority-driven ("exogenous") switch times of around 75 ms (Olivers & Meeter, 2008, p. 843).

There have been other studies than can be construed as measures of the time required to shift attention. For example, Theeuwes, Godijn, and Pratt (2004) developed an improved dwell time paradigm, which avoids the problem of encoding multiple targets. In their experiment, an arrow directed attention to an initial position. Then, two potential target stimuli appeared simultaneously, along with two more arrows. One of the arrows, presented at the initial position, indicated which target had to be reported. At this point, the observer had to initiate a second attentional shift in order to discriminate the target. Simple RT to an onset probe was used to measure attentional allocation. Probes at the second position were responded to more quickly than probes at the initial position between 200 and 300 ms after the onset of the display with the second arrow, so Theeuwes et al. concluded that the dwell time must be around 250 ms. This measurement is in general agreement with our findings here, though slightly longer. However, time to interpret the second arrow cue might be factored into the estimate in this case.

Priority-driven attention

We can confidently claim that priority-driven shifts of attention are substantially faster than voluntary shifts. In four experiments, across a variety of search tasks and stimulus factors, anarchic search proved noticeably faster than commanded search. Estimating the actual speed of priority-driven shifts is more difficult. The asymptotes observed in Experiments 1 and 2 were probably overestimates, due to the floor in our staircase procedure. The values observed in Experiments 3

Downloaded By: [Harvard College] At: 19:58 4 March 2009

and 4, which did not suffer from the floor limitation, differ by a factor of five. This may reflect differential difficulty of the search tasks; discriminating "P" or "J" from the remainder of the alphabet is relatively simple compared to discriminating a letter from its mirror reversal. The discrepancy may also reflect the fact that we were measuring the rate to achieve 66.7% accuracy in Experiment 3 (or 79.5–84.6%, based on data from the fixed-duration blocks), while in Experiment 4, we measured the rate that produced 96.1% accuracy.

Furthermore, slope data must be interpreted according to a theory of sampling in search. If the display is sampled with replacement, then the slope directly estimates the search rate. If sampling is without replacement, the search rate would be twice the observed slope. Both of these estimates would need to be modulated by some account of the effect of errors. On the basis of our previous work (Horowitz & Wolfe, 1998, 2001, 2003), we believe that the sampling regime in this type of search task is closer to with replacement than without, implying search rates at the faster end of the range (~50 ms/item). This is faster than the time course estimated from the peak of the SOA × Accuracy functions measured in cueing studies (Cheal & Lyon, 1991; Nakayama & Mackeben, 1989). As we previously noted, cue interpretation time has not been studied in those paradigms. Even a reflexive cue must be processed before it can be effective, so it is likely that a significant portion of the 100-ms delay between cue onset and peak accuracy reflects the time necessary for the signal to reach orienting mechanisms in parietal cortex. The actual shifting time is probably less than 100 ms. In visual search experiments such as those used in Experiments 3 and 4, all stimuli are simultaneously present, and the salience mechanism can be computing the next attentional destination in parallel with the current shift of attention, so that the next shift can be executed with little delay. Some support for this conclusion can be derived from the work of Danziger and Kingstone (1999), who demonstrated that attention can be reoriented rapidly (i.e., within 50 ms) from a cued location to a location likely to contain a target.

CONCLUSIONS

Volitional changes in the focus of attention take quite a long time. It is substantially faster to "delegate authority". If you tell yourself to find the letter "P" or red verticals or your coffee mug, selective attention will shift around the visual world at a rate at least four times faster, in our estimation, than it would if you insisted on commanding each deployment of attention with an individual act of will. This general conclusion might have been anticipated from the endogenous versus exogenous cueing literature. However, in studies of that sort, it is hard to tell whether the slower speed of endogenous cueing has something to do with cue interpretation or with a fundamental limit on a class of attentional shift. In our experiments, observers made multiple shifts of attention in a predictable pattern, which did not require interpretation of a cue. Therefore, the difference between priority-driven and voluntary shifts of attention cannot be attributed to difficulty in deciding where to shift attention.

Instead, we propose that there are two modes of directing spatial attention. In the priority-driven mode, attention is driven by a competitive network, which encodes information about bottom-up stimulus salience (Itti & Koch, 2001; O'Grady & Muller, 2000) and top-down information about target identity, as well as a priori information about where targets are likely to be in a scene (Chun & Jiang, 1998, 2003; Graboi & Lisman, 2003; Torralba, Oliva, Castelhano, & Henderson, 2006). Activation in this network represents the visual system's "best guess" about which loci are likely to contain the target at any given moment. An autonomous module can deploy attention to the most "active" location or object every 25–50 ms (Wolfe, 1994, 2006).

It is possible to overrule this autonomous agent and select the next object of attention. This volitional deployment is much slower than the autonomous, priority-driven mode. We propose that the slower rate reflects the "clock speed of free will". It might also reflect the clock speed of perceptual experience. Even if we can search through a display at 20–40 items/second, we do not experience 20–40 discrete selection events. We experience the search and its outcome but the rapid autonomous deployments of attention that are revealed by experiments are not available to consciousness.

This dual-control scheme has implications for our claim that visual search has a very limited memory for prior deployments of attention (Horowitz & Wolfe, 1998, 2001, 2003). The amnesic nature of brief, laboratory search tasks suggests that the brain is stupid. Why not search in some systematic manner that would assure that we search without replacement? These data provide an answer. Search without replacement requires, on average, half as many deployments of attention as completely amnesic search. Searches with partial memory lie between these extremes. Commanding yourself to search without replacement will slow each deployment at least fourfold. It is a bad investment to accept a 4× slowing of attentional deployments in order to produce a 2× improvement in the number of deployments. Systematic search may not be worth it unless an eye movement or an act of will can be used to make a more substantial contribution (e.g., I am not going to search the bedroom because I know my keys are in the kitchen).

Downloaded By: [Harvard College] At: 19:58 4 March 2009

Original manuscript received 10 June 2006
Accepted revision received 17 October 2008
First published online day month year

REFERENCES

Araujo, C., Kowler, E., & Pavel, M. (2001). Eye movements during visual search: The costs of choosing the optimal path. *Vision Research*, *41*, 3613–3625.

Bowman, H., & Wyble, B. (2007). The simultaneous type, serial token model of temporal attention and working memory. *Psychological Review*, *114*, 38–70.

Boyer, J., & Ro, T. (2007). Attention attenuates metacontrast masking. *Cognition*, *104*, 135–149.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 443–446.

Broadbent, D. E., & Broadbent, M. H. (1987). From detection to identification: Response to multiple targets in rapid serial visual presentation. *Perception & Psychophysics*, *42*, 105–113.

Brockmann, D., & Geisel, T. (2000). The ecology of gaze shifts. *Neurocomputing*, *32*, 643–650.

Carpenter, R. H. S. (1977). *Movements of the eyes*. London: Pion.

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

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*, 28–71.

Chun, M. M., & Jiang, Y. (2003). Implicit, long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 224–234.

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 109–127.

Cole, G., Gellatly, A., & Blurton, A. (2001). Effect of object onset on the distribution of visual attention. *Journal of Experimental Psychology: Human Perception & Performance*, *27*, 1356–1368.

Corbetta, M. (1998). Frontoparietal cortical networks for directing attention and the eye to visual locations: Identical, independent, or overlapping neural systems? *Proceedings of the National Academy of Sciences of the United States of America*, *95*, 831–838.

Danziger, S., & Kingstone, A. (1999). Unmasking the inhibition of return phenomenon. *Perception & Psychophysics*, *61*, 1024–1037.

Duncan, J., & Humphreys, G. (1992). Beyond the search surface: Visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception & Performance*, *18*, 578–588.

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

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*, 425–451.

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, *4*, 345–352.

Fischer, M. H. (1999). An investigation of attention allocation during sequential eye movement tasks. *Quarterly Journal of Experimental Psychology A*, *52*, 649–677.

Gibson, B. S., Li, L., Skow, E., Salvagni, K., & Cooke, L. (2000). Memory-based tagging of targets during visual search for one versus two identical targets. *Psychological Science*, *11*, 324–328.

Gilchrist, I. D., & Harvey, M. (2000). Refixation frequency and memory mechanisms in visual search. *Current Biology*, *10*, 1209–1212.

Gilchrist, I. D., & Harvey, M. (2006). Evidence for a systematic component within scanpaths in visual search. *Visual Cognition*, *14*, 704–715.

Graboi, D., & Lisman, J. (2003). Recognition by top-down and bottom-up processing in cortex: The control of selective attention. *Journal of Neurophysiology*, *90*, 798–810.

Horowitz, T. S., Holcombe, A. O., Wolfe, J. M., Arsenio, H. C., & DiMase, J. S. (2004). Attention pursuit is faster than attentional saccades. *Journal of Vision*, *4*, 583–603.

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature*, *394*, 575–577.

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

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

Horowitz, T. S., & Wolfe, J. M. (2005). Visual search: The role of memory for rejected distractors. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Neurobiology of attention* (pp. 264–268). San Diego, CA: Elsevier.

Hunt, S. (1994). MacProbe: A Macintosh based experimenters workstation for the cognitive sciences. *Behavior Research Methods, Instruments & Computers*, *26*, 345–351.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*, 194–203.

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: Lawrence Erlbaum Associates.

Kane, M. J., Poole, B. J., Tuholski, S. W., & Engle, R. W. (2006). Working memory capacity and the top-down control of visual search: Exploring the boundaries of "executive attention". *Journal of Experimental Psychology: Learning Memory and Cognition*, *32*, 749–777.

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

Korner, C., & Gilchrist, I. D. (2008). Memory processes in multiple-target visual search. *Psychological Research*, *72*, 99–105.

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

Kristjansson, A., & Nakayama, K. (2002). The attentional blink in space and time. *Vision Research*, *42*, 2039.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*, 119–159.

Luck, S. J., & Thomas, S. J. (1999). What variety of attention is automatically captured by peripheral cues? *Perception & Psychophysics*, *61*, 1424–1435.

Ludwig, C. J., & Gilchrist, I. D. (2002). Stimulus-driven and goal-driven control over visual selection. *Journal of Experimental Psychology: Human Perception & Performance*, *28*, 902–912.

Machado, L., & Rafal, R. D. (2000). Strategic control over saccadic eye movements: Studies of the fixation offset effect. *Perception & Psychophysics*, *62*, 1236–1242.

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition*, *22*, 657–672.

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience*, *5*, 229–240.

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*, 1517–1539.

Miller, J. (1991). The flanker compatibility effect as a function of visual angle, attentional focus, visual transients, and perceptual load: A search for boundary conditions. *Perception & Psychophysics*, *49*, 270–288.

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

Downloaded By: [Harvard College] At: 19:58 4 March 2009

Downloaded By: [Harvard College] At: 19:58 4 March 2009

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, UK: Oxford University Press.

Moore, T., Armstrong, K. M., & Fallah, M. (2003). Visuomotor origins of covert spatial attention. *Neuron*, *40*, 671–683.

Müller, H. J., & Rabbitt, P. M. (1989). Spatial cueing and the relation between the accuracy of "where" and "what" decisions in visual search. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *41A*, 747–773.

Murthy, A., Thompson, K. G., & Schall, J. D. (2001). Dynamic dissociation of visual selection from saccade programming in frontal eye field. *Journal of Neurophysiology*, *86*, 2634–2637.

Nakayama, K., & Joseph, J. S. (1998). Attention, pattern recognition, and pop-out visual search. In R. Parasuraman (Ed.), *The attentive brain* (pp. 279–298). Cambridge, MA: The MIT Press.

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

Nobre, A. C., Gitelman, D. R., Dias, E. C., & Mesulam, M. M. (2000). Covert visual spatial orienting and saccades: Overlapping neural systems. *NeuroImage*, *11*, 210–216.

Nothdurft, H. C. (1991). Texture segmentation and pop-out from orientation contrast. *Vision Research*, *31*, 1073–1078.

O'Grady, R. B., & Muller, H. J. (2000). Object-based selection operates on a grouped array of locations. *Perception & Psychophysics*, *62*, 1655–1667.

Oliva, A., Torralba, A., Castelhano, M. S., & Henderson, J. M. (2003). Top-down control of visual attention in real world scenes [Abstract]. *Journal of Vision*, *3*, 3a.

Olivers, C. N., & Meeter, M. (2008). A boost and bounce theory of temporal attention. *Psychological Review*, *115*, 836–863.

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, *40*, 1227–1268.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

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

Peterson, M. S., & Kramer, A. F. (2001). Attentional guidance of the eyes by contextual information and abrupt onsets. *Perception & Psychophysics*, *63*, 1239–1249.

Peterson, M. S., Kramer, A. F., & Irwin, D. E. (2004). Covert shifts of attention precede involuntary eye movements. *Perception & Psychophysics*, *66*, 398–405.

Peterson, M. S., Kramer, A. F., Wang, R. F., Irwin, D. E., & McCarley, J. S. (2001). Visual search has memory. *Psychological Science*, *12*, 287–292.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3–25.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception & Performance*, *18*, 849–860.

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

Schall, J. D. (2004). On the role of frontal eye field in guiding attention and saccades. *Vision Research*, *44*, 1453–1467.

Serences, J. T., & Yantis, S. (2006). Selective visual attention and perceptual coherence. *Trends in Cognitive Sciences*, *10*, 38–45.

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

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

Theeuwes, J., & Godijn, R. (2002). Irrelevant singletons capture attention: Evidence from inhibition of return. *Perception & Psychophysics*, *64*, 764–770.

Theeuwes, J., Godijn, R., & Pratt, J. (2004). A new estimation of the duration of attentional dwell time. *Psychonomic Bulletin & Review*, *11*, 60–64.

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.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*, 520–522.

Torralba, A., Oliva, A., Castelhano, M., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*, 766–786.

Downloaded By: [Harvard College] At: 19:58 4 March 2009

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

VanRullen, R., & Thorpe, S. J. (2001). The time course of visual processing: From early perception to decision-making. *Journal of Cognitive Neuroscience*, *13*, 454–461.

von Mühlenen, A., Müller, H. J., & Müller, D. (2003). Sit-and-wait strategies in dynamic visual search. *Psychological Science*, *14*, 309–314.

Ward, R., Duncan, J., & Shapiro, K. (1996). The slow time-course of visual attention. *Cognitive Psychology*, *30*, 79–109.

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. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in Cognitive Sciences*, *7*, 70–76.

Wolfe, J. M. (2006). Guided Search 4.0: Current progress with a model of visual search. In W. D. Gray (Ed.), *Integrated models of cognitive systems*. New York: Oxford.

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

Yantis, S. (1993). Stimulus-driven attentional capture and attentional control settings. *Journal of Experimental Psychology: Human Perception & Performance*, *19*, 676–681.

Zelinsky, G. J., & Sheinberg, D. L. (1997). Eye movements during parallel-serial visual search. *Journal of Experimental Psychology: Human Perception & Performance*, *23*, 244–262.