Research Article

EYE MOVEMENTS REVEAL THE SPATIOTEMPORAL DYNAMICS OF VISUAL SEARCH

Gregory J Zelinsky,[1,2] Rajesh P N. Rao,[2] Mary M Hayhoe,[1] and Dana H. Ballard[2]

[1]*Center for Visual Science and* [2]*Computer Science Department, University of Rochester*

Abstract—*Given that attention precedes an eye movement to a target it becomes possible to use fixation sequences to probe the spatiotemporal dynamics of search Applying this method to a realistic search task, we found eye movements directed to the geometric centers of progressively smaller groups of objects rather than accurate fixations to individual objects in a display Such a binary search strategy is consistent with zoom-lens models positing an initially broad distribution of search, followed by a narrowing of this search region until only the target is selected We also interpret this oculomotor averaging behavior as evidence for an initially parallel search analysis that becomes increasingly serial as the search process converges on the target*

Visual search, the process by which one locates a target in a cluttered scene, is a common and important behavior that has defied definitive understanding despite decades of diligent research Two factors have combined to make the study of this topic exceedingly difficult First, search movements and shifts of selective visual attention are almost certainly intertwined Any complete description of search must therefore also address these difficult-to-observe underlying attention movements, as well as assume a theoretical stance among the various metaphors for attentional function (e g , spotlights, zoom lenses, filters, channels) Second, visual search is more than the time taken by an observer to detect a target and press a button It is instead a richly complex behavior having both a spatial and a temporal dynamic Most search studies, however, largely discard this spatiotemporal information by collapsing the search process into a single measure of reaction time (RT) Such reliance on a dependent measure that cannot directly resolve these search dynamics introduces unexplained variability into every search experiment and fuels the endless debates that threaten to paralyze research into the process of visual search

A newfound relationship between eye movements and directed visual attention offers a promising method of avoiding both of these difficulties Although it is certainly possible to shift visual attention without making an accompanying eye movement (Klein & Farrell, 1989, Posner, 1980, Remington, 1980, Treisman & Gormican, 1988), several recent studies have shown that the reverse dissociation may not be possible (Deubel & Schneider, 1996, Hodgson & Muller, 1995, Hoffman & Subramaniam, 1995, Irwin, 1992, Kowler, Anderson, Dosher, & Blaser, 1995, Shepherd, Findlay, & Hockey, 1986, also see Henderson, 1996, for a review of this and related topics) Specifically, a shift in visual attention to a location in space must accompany an eye movement to that same location This association is believed to be due to a shared neural substrate between covert attentional orienting and oculomotor programming (Hornak, 1992, Kustov & Robinson, 1996, Posner, Petersen, Fox, & Raichle, 1988, Rafal, Calabresi, Brennan, & Sciolto, 1989, Rizzolatti, Riggio, & Sheliga, 1994, Robinson & Kertzman, 1995, Sheliga, Riggio, & Rizzolatti, 1994, Walker & Findlay, 1996) Typically, selection of a target in a spatially organized neural map would elicit both an attentional movement and an eye saccade Purely covert orienting would occur when an eye movement cannot be executed because of the oculomotor refractory period (Carpenter, 1988) or in those unnatural cases when the saccade is being voluntarily inhibited (Klein & Farrell, 1989, Zelinsky & Sheinberg 1997) Indirect evidence supporting this relationship between eye movements and visual attention can also be found in the rapidly growing number of studies showing an alignment between saccadic inspection and visual search during free viewing (Behrmann, Watt, Black, & Barton, in press, Engel, 1977, Findlay, 1997, Gould, 1973, Jacobs, 1986, Rayner & Fisher, 1987, Williams, 1967, Zelinsky 1996, Zelinsky & Sheinberg, 1997)

EXPERIMENT 1

The current study by exploiting the relationship between eye movements and directed visual attention, introduces a new methodology for studying the spatiotemporal dynamics of visual search If an attentional shift to a location in space necessarily precedes an eye movement to that same location, then each ocular fixation provides a spatial marker or record of a display region visited by attention and search Furthermore, unlike RTs, which provide only a single temporal measure of search and no spatial measure whatsoever, the sequence of saccades and fixations accompanying search provides a more detailed picture of how the search process evolves over time Note, however, that this methodology is not without its limitations It cannot track any attentional shifts occurring in addition to those accompanying refixation, nor can it say with great certainty which items in a display have been processed by attention What it does provide is a rough indication of the spatial and temporal attentional allocation to items in a search display

Method

We tracked the eye movements of 6 participants (4 naive) as they searched for objects in three pseudorealistic scenes (a crib, dining table, and workbench, Fig 1) Dominating each scene was either one, three, or five contextually appropriate objects electronically arranged on a background surface (i e , toys in the crib, tools on the workbench and food-related objects on the dining table) Using the midpoint of a 2 5° imaginary box enclosing each object as a reference, we constrained the positions of objects in each search display to six locations of equal eccentricity (either 22 5°, 45°, 67 5°, 112 5°, 135°, or 157 5° at an eccentricity of 7°) along an arc centered on the observer's initial fixation The composite scene (background surface and objects) subtended 16° horizontally and 12° vertically, filling the entire 640 × 480

Address correspondence to Gregory Zelinsky Beckman Institute, University of Illinois at Urbana-Champaign, 405 North Mathews Ave , Urbana, IL 61801 e-mail gzelinsk@casper beckman uiuc edu

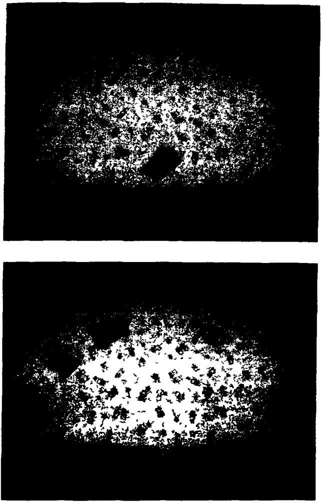

Fig. 1. Images from 1 of the 360 trials used in this experiment, converted to gray scale. The first image (top) instructed observers to search for the toy car in the following image (bottom). Note that in this case the correct response would be that the target is absent from the search display.

pixel screen. The 360 trials per observer (each trial described a unique configuration of objects and positions on a surface) were evenly divided into randomly interleaved target-present or target-absent conditions and three set sizes, leaving 60 trials per cell of the experimental design. In the case of target-absent trials the target was replaced by a distractor from the same object category rather than leaving the position unoccupied.

A trial consisted of the sequential presentation of two 8-bit indexed-color images. The first image was visible for 1 s and designated the target of the search task by showing a single object at a bottom-center location on the surface (Fig. 1, top panel). The observer's task was to indicate the presence or absence of the target object in the following search scene (Fig. 1, bottom panel) by making a speeded key-press response. An SRI Generation-V dual-Purkinje-image eyetracker was used to sample eye position every other millisecond during the presentation of the search display. Eye position was also monitored during a fixation display preceding the search scene. This additional measure made the search presentation contingent upon accurate (±0.25°) fixation of the cross, thereby ensuring a constant distance between the eye's starting position and the search items in the following display.

Results and Discussion

An analysis of manual responses revealed that search times increased with the number of objects appearing in the scene, $F(1, 5) = 69.21$, $p < .001$, and the rate of this increase in the target-absent trials (41 ms/item) was significantly greater than the target-present slope (28 ms/item), $F(2, 10) = 8.75$, $p = .006$, by two-way repeated measures analysis of variance. Proponents of several popular models of search might argue, on the basis of this analysis, that attention was serially directed from object to object in these scenes until the target was detected or the displays were exhaustively inspected (Posner, Snyder, & Davidson, 1980; Treisman, 1988; Treisman & Gelade, 1980; Wolfe, 1994). According to these "spotlight" models, it is this serial, item-by-item movement of a focused region of attentional processing that accounts for the longer RTs with larger set sizes.

Although these data cannot rule out a serial spotlight model, an analysis of the eye movements accompanying this search task does suggest a different spatiotemporal dynamic. Surprisingly, most initial saccades (Fig. 2, top panel) were directed toward the center of the scenes even though no objects ever appeared there, a fixation pattern reminiscent of center-of-gravity averaging observed for simple stimuli in early oculomotor studies (Findlay, 1982, 1987; P. He & Kowler, 1989; Richards & Kaufman, 1969). The scatterplot of landing positions for second saccades (middle panel) shows gaze moving closer to the search objects, but notice that these eye movements were still fairly inaccurate, forming an undifferentiated band of endpoints along each side of the display. It was typically not until after the third saccades (bottom panel) that individual objects in the scene were fixated accurately.

This same oculomotor averaging behavior is shown in Figure 3 for two representative trials. The scanpath illustrated in the top panel shows an initial eye movement to the centroid of a group of three objects, followed by a second saccade to an intermediate location between two of these items and a third saccade to the target. What is absent from this scanpath is evidence for a serial process directing search to individual objects in the display. This process by which gaze gradually converges on the target suggests a binary search strategy, rather than a sequential item-by-item search. The analysis occurring after the initial center-of-gravity saccade effectively divides the display in two, isolating the hemifield in which the target is located. The second saccade then brings gaze to the center of an object configuration on the selected side of the display, after which another binary decision is made (in the case of the top panel in Fig. 3, this third oculomotor decision was to shift gaze upward toward the butter target rather than downward toward the napkin and silverware). Notice that when only a single object appeared on the surface (Fig. 3, bottom panel) the initial saccade was weighted heavily toward the target but the center-of-gravity averaging tendency did not disappear entirely. This evidence for averaging behavior even at a set size of one may be due to an initial interpretation of the entire surface as being relevant to the search task and to its being partially weighted in the saccade computation (Z. He & Nakayama, 1992).

This oculomotor convergence toward an object was quantified in Figure 4 by plotting how close the first three eye movements brought

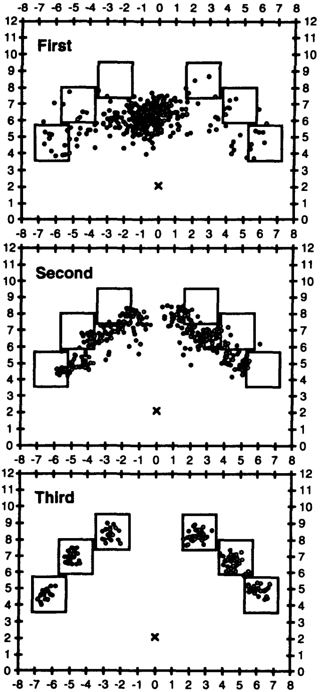

Fig 2. Endpoints of first (top), second (middle), and third (bottom) saccades from 1 naive observer. Axes are in degrees of visual angle, with values along the abscissa indicating distance from the fixation cross. Data from all three set sizes and scene types are shown. The fixation cross indicates the starting eye position, and the black boxes correspond to where objects appeared in the scenes. Note that there are fewer data points in the middle and bottom panels than in the top panel because of the observers' occasional failure to make a second or third saccade in a trial.

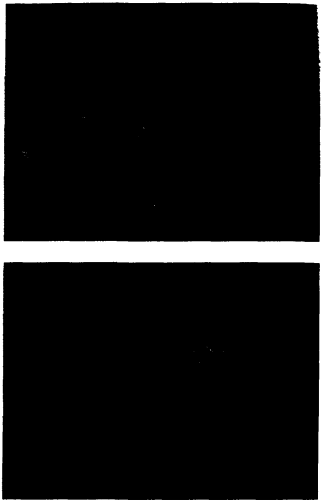

Fig 3. Eye movements occurring on two representative trials, superimposed over the scenes that were being viewed. Individual fixations are indicated by the white circles, with the diameter of each circle representing the relative fixation duration. The sequences of fixations and saccades shown are for a three-item trial (top) and a single-item trial (bottom).

gaze to the target as a function of set size. Initial saccades landed approximately 5° from the target in the three- and five-item displays. This measure of initial endpoint error decreased to 3.1° in the single-item trials, yielding a significant main effect of set size, $F(2, 10) = 111.44$, $p < .001$, by one-way repeated measures analysis of variance. Targeting accuracy improved markedly for the second saccades, $F(1, 5) = 141.33$, $p < .001$, with endpoint errors rising linearly from 0.8° at the smallest set size to 2.7° at the largest set size, $F(2, 10) = 73.89$, $p < .001$. This influence of neighboring distractors on fixation accuracy, although still significant ($F[2, 10] = 6.83$, $p = .013$), was largely attenuated by the third saccades. Note also that the difference in first-saccade endpoint error between the one-object and the three- and five-object displays suggests that the initial eye movements were not simply anticipatory or preprogrammed behaviors. Because set size was randomly interleaved throughout the experiment, observers had no way of anticipating the number of objects that would appear in the

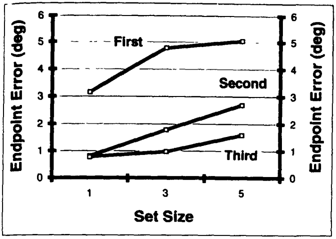

Fig. 4. Mean endpoint errors as a function of set size and saccade (first, second, and third) for all 6 participants. Endpoint error refers to the distance between the landing position of a saccade and the target in degrees of visual angle. Because endpoint error is undefined when a target does not appear in the display, results for target-present trials only are shown.

following search scene. Rather than reflecting a routinized oculomotor response, this difference in endpoint error therefore suggests that the center-of-gravity fixations observed were sensitive to the stimulus properties of each display.

EXPERIMENT 2

Before discussing what these oculomotor results mean for the allocation of visual attention, we must rule out a competing explanation for the center-of-gravity fixations observed in Experiment 1. Namely, if participants were unable to peripherally resolve the search objects at 7° eccentricity, then they may have purposefully directed initial fixation to a centroid location in an attempt to see the objects better and perform the task more accurately. It follows from this hypothesis that if participants were prevented from making eye movements, then these same visual acuity limitations would translate into higher manual error rates. However, if these saccades were part of the normal spatiotemporal search process, then no meaningful differences in the pattern of errors would be expected between the tasks.

Method

Four new observers searched the identical stimuli as in Experiment 1 without moving their eyes during the presentation of the search displays. To help participants follow this "no eye movement" instruction, we added a small (0.25°) fixation cross to each of the background scenes at a 7° eccentric location corresponding to the starting eye position in the previous experiment. Eye data were analyzed off-line, and trials in which gaze deviated by more than 0.5° from the cross were discarded. Observers were instructed to respond as accurately as possible without regard for time. To better equate the available information in the two tasks, we used RTs from Experiment 1 to set the durations of the search displays in this control experiment. Specifically, averaged RTs were computed for the three set sizes (collapsed across target-present and -absent trials), and these values were decremented by a constant 200 ms to help correct for motor latencies that may inflate actual search decision times. The resulting display durations were 444 ms, 497 ms, and 583 ms for set sizes of one, three, and five items. Following the appropriate fixed interval, each search display was replaced by a blank screen, which remained until the observer indicated whether the target was present or absent.

Results and Discussion

Table 1 shows the percentages of misses and false alarms in Experiment 1 and in the fixed-eye control task. The percentages of trials discarded because of the detection of a saccade are also shown for Experiment 2. Despite a reported strong subjective impression of moving their eyes, the observers rarely (less than 1.5% of the trials) initiated a saccade during the search displays. Tiny changes in eye position (~0.1°) at the practical limits of the tracking device were occasionally observed, but we did not attempt to correlate these microsaccades with the direction of the search target. Just as fixation was maintained with high accuracy, manual responses were also very accurate. Mean button-press errors in both the free-eye and the fixed-eye experiments were uniformly low across all search conditions, with the percentage of errors in the control task being generally smaller. Furthermore, the few manual errors occurring in the control experiment were reported to be simple motor confusions (pressing one button when the other was intended) rather than guessing attributable to poor visual acuity. The exceptional accuracy observed in this fixed-eye control task therefore suggests that observers were able to peripherally resolve the search objects at 7° eccentricity, and that participants elected to make eye movements in the free-eye search task even though they were quite capable of performing the task without changing fixation.

GENERAL DISCUSSION

Although eye movements are clearly incidental to accurate performance in many search tasks, when an eye movement does accompany search, it communicates a wealth of spatial and temporal information about the search process and the allocation of directed visual attention. Given that search is attentionally mediated, and the assumption that attention visits the target of a saccadic eye movement, two conclusions can be drawn from the eye data presented in Figures 2 through 4. First, search in this task began very broadly distributed over the scenes but then spatially collapsed to surround only the target object. The evidence for this search dynamic exists in the path followed by the eye to the target. Contrary to common intuitions about search (that the eye moves accurately from one item to another in a scene), gaze was directed first to the centroid of the global display configuration and then to the centers of recursively smaller groups of objects until the target was acquired. Although this binary search pattern should not be interpreted as strict evidence against item-by-item spotlight theories of search, such a pattern is more consistent with models likening the search process to the global-to-local operation of a zoom lens (Downing & Pinker, 1985, Eriksen & St. James, 1986, Eriksen & Yeh, 1985, Laberge, 1983).[1]

---

[1] Given the documented relationship between eye movements and attention, it is still possible to salvage a purely serial account of these data by relaxing the spatial coupling between attentional locus and ocular

Table 1 *Mean percentage of errors as a function of set size and search task*

	Free-eye task (Experiment 1)		Fixed-eye task (Experiment 2)		
Set size	Misses	False alarms	Misses	False alarms	Saccades detected
One	1.1	2.8	0.8	0.8	1.0
Three	1.1	1.4	2.1	0.4	1.3
Five	3.6	2.2	2.9	0.0	1.3

The second conclusion is closely tied to the first and involves the contribution of parallel and serial processes to search performance in this task. More specifically, parallel processing is implicated to the extent that center-of-gravity averaging appears in the data. This assertion follows from the fact that the eye, and presumably attention, was initially directed to configurations of objects rather than individual items. Because configurations consist of multiple component objects, the simultaneous influence of each of these objects on search (i.e., averaging) necessarily implies a parallel computation. It also follows that the spatial extent of this parallel process is indicated by the distribution of these component objects. The demonstration of global center-of-gravity averaging by the first-saccade endpoints suggests an attentional process that initially encompasses the entire search display. Likewise, the recruitment of a smaller group of objects with each additional eye movement, and the eventual accurate fixation of the target, also suggests movement to a more serial analysis over time. The current oculomotor data therefore suggest that there is a gradual progression from parallel to serial processing in the same search task, rather than that these two processes are dichotomous.

Recent work in our lab has shown how a simple color- and spatial-filtering computation also unfolding over time can implement such a search dynamic and parsimoniously account for these eye data (Rao, Zelinsky, Hayhoe, & Ballard, 1996). Early in the search computation, many points in a realistic scene may correlate highly with an iconic representation of the target. Because the strength and spatial distribution of these target-icon correlations are thought to describe the moment-by-moment deployment of visual attention in this task, search according to this view would be an initially parallel process spread over much of the display. An eye movement programmed at this stage of the search process will be directed to the weighted average of these iconic matches, thereby giving rise to the centroid fixation patterns observed for the initial saccades. As higher spatial frequency information becomes available, points in the display corresponding to the icon will become even more correlated, and less likely candidates will drop out of the computation, resulting in a narrowing search region and a movement of gaze toward the target. It is this alignment between saccade endpoints and the centroid of a collapsing search region that enables ocular fixations to act as spatial indicators of how the search process evolves over time.

Acknowledgments—The first author was funded by a James S. McDonnell Foundation postdoctoral training grant (No. 93-7) under the supervision of Dana Ballard in the Department of Computer Studies, University of Rochester. This work was also supported by National Science Foundation Research Grant CDA-8822724, National Institutes of Health Research Grants 1-R24-RR06853, EY-05729, and 1-P41-RR09283, and a grant from the Human Science Frontiers Program. We would like to thank Jeremy Wolfe and Michael Land for many valuable comments on an earlier version of this manuscript.

REFERENCES

Behrmann, M. Watt, S. Black, S., & Barton, J. (in press). Impaired visual search in patients with unilateral neglect: An oculographic analysis. *Neuropsychologia*.

Carpenter, R. (1988). *Movements of the eyes* (2nd ed.). London: Pion.

Deubel, H., & Schneider, W. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36*, 1827–1837.

Downing, C., & Pinker, S. (1985). The spatial structure of visual attention. In M.I. Posner & O.S. Marin (Eds.), *Attention and performance XI* (pp. 171–187). Hillsdale, NJ: Erlbaum.

Engel, F. (1977). Visual conspicuity, visual search and fixation tendencies of the eye. *Vision Research, 17*, 95–108.

Eriksen, C., & St. James, J. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics, 40*, 225–240.

Eriksen, C., & Yeh, Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 583–597.

Findlay, J. (1982). Global visual processing for saccadic eye movements. *Vision Research, 22*, 1033–1045.

Findlay, J. (1987). Visual computation and saccadic eye movements: A theoretical perspective. *Spatial Vision, 2*, 175–189.

Findlay, J. (1997). Saccade target selection during visual search. *Vision Research, 37*, 617–631.

Gould, J. (1973). Eye movements during visual search and memory search. *Journal of Experimental Psychology, 98*, 184–195.

He, P., & Kowler, E. (1989). The role of location probability in the programming of saccades: Implications for center-of-gravity tendencies. *Vision Research, 29*, 1165–1181.

He, Z., & Nakayama, K. (1992). Surfaces versus features in visual search. *Nature, 359*, 231–233.

Henderson, J. (1996). Visual attention and the attention action interface. In K. Akins (Ed.), *Perception* (pp. 290–316). New York: Oxford University Press.

Hodgson, T., & Müller, H. (1995). Evidence relating to premotor theories of visuospatial attention. In J. Findlay, R. Walker, & R. Kentridge (Eds.), *Eye movement research: Mechanisms, processes, and applications* (pp. 305–316). Amsterdam: North Holland.

Hoffman, J., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics, 57*, 787–795.

Hornak, J. (1992). Ocular exploration in the dark by patients with visual neglect. *Neuropsychologia, 30*, 547–552.

Irwin, D. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 307–317.

---

fixation. Saccades might be directed to the centroid of an object configuration, but, because of this loose coupling, attention might actually be allocated to nearby neighboring items in a serial fashion. Although this possibility cannot explain the binary search pattern observed in the eye data, nor is it likely given the sometimes large distance between the endpoint of the initial saccade and the nearest object, in future work we will attempt to describe this spatial coupling with greater precision.

cobs A (1986) Eye-movement control in visual search How direct is visual span control? *Perception & Psychophysics 39* 47–58

lein R & Farrell M (1989) Search performance without eye movements *Perception & Psychophysics, 46* 476–482

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

ustov A & Robinson, D L (1996) Shared neural control of attentional shifts and eye movements *Nature 384* 74–77

Laberge D (1983) Spatial extent of attention to letters in words *Journal of Experimental Psychology Human Perception and Performance 9* 371–379

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

Posner M Petersen S Fox P & Raichle M (1988) Localization of cognitive functions in the human brain *Science 240* 1627–1631

Posner, M Snyder C & Davidson B (1980) Attention and the detection of signals *Journal of Experimental Psychology General 109* 160–174

Rafal R Calabresi P Brennan C & Sciolto T (1989) Saccade preparation inhibits reorienting to recently attended locations *Journal of Experimental Psychology Human Perception and Performance 15* 673–685

Rao R Zelinsky G Hayhoe M & Ballard D (1996) Modeling saccadic targeting in visual search In D Touretzky M Mozer & M Hasselmo (Eds) *Advances in neural information processing systems 8* (pp 830–836) Cambridge MA MIT Press

Rayner K & Fisher D (1987) Letter processing during eye fixations in visual search *Perception & Psychophysics 42* 87–100

Remington R (1980) Attention and saccadic eye movements *Journal of Experimental Psychology Human Perception and Performance 6* 726–744

Richards W & Kaufman L (1969) Center-of-gravity tendencies for fixations and flow patterns *Perception & Psychophysics 5* 81–84

Rizzolatti G Riggio L & Sheliga, B (1994) Space and selective attention In C Umilta & M Moscovitch (Eds) *Attention and performance XV* (pp 231–265) Hillsdale, NJ Erlbaum

Robinson D L & Kertzman, C (1995) Covert orienting of attention in macaques III Contributions of the superior colliculus *Journal of Neurophysiology 74* 713–721

Sheliga, B Riggio L & Rizzolatti G (1994) Orienting of attention and eye movements *Experimental Brain Research, 98* 507–522

Shepherd, M Findlay J & Hockey R (1986) The relationship between eye movements and spatial attention *Quarterly Journal of Experimental Psychology 38A* 475–491

Treisman A (1988) Features and objects The Fourteenth Bartlett Memorial Lecture *Quarterly Journal of Experimental Psychology 40A,* 201–237

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

Treisman A & Gormican S (1988) Feature analysis in early vision Evidence from search asymmetries *Psychological Review 95* 15–48

Walker R & Findlay J (1996) Saccadic eye movement programming in unilateral neglect *Neuropsychologia 34* 493–508

Williams L (1967) The effects of target specification on objects fixated during visual search *Acta Psychologica, 27* 355–360

Wolfe J (1994) Guided search 2 0 A revised model of visual search *Psychonomic Bulletin & Review 1* 202–238

Zelinsky G (1996) Using eye saccades to assess the selectivity of search movements *Vision Research, 36* 2177–2187

Zelinsky G & Sheinberg D (1997) Eye movements during parallel serial visual search tasks *Journal of Experimental Psychology Human Perception and Performance 23* 244–262

(RECEIVED 8/22/96 REVISION ACCEPTED 5/9/97)