Research Report

PERCEPTUAL GROUPING AND MOTION COHERENCE IN VISUAL SEARCH

A. Kingstone and W.F. Bischof
University of Alberta, Edmonton, Canada

Abstract—*Search for a conjunction of form and motion is greatly affected by manipulations of phase in the target and nontarget motion sets. To test whether this finding can be best explained by perceptual grouping, we moved a random set of dots in phase or counterphase with target or nontarget motion. Perceptual grouping was found to have a dramatic effect on search performance. We propose that this interaction between perceptual grouping and visual search is governed by three general rules. Our data also provide convincing evidence of the preattentive organization of a visual display into surfaces defined by common motion.*

In the real world, different items compete for visual attention. How does one select a target item from among the many different nontarget items? Treisman (1986; Treisman & Gelade, 1980) hypothesized that a target defined by a single feature (e.g., a red book among blue books) would “pop out” from the nontargets because single features were processed in parallel, without attention. However, a target defined by a conjunction of features (e.g., a small red book among small blue books and large red books) would require serial spatial attention in order to identify the correct conjunction of features because attention is needed to conjoin features.

Subsequent studies, however, were problematic for this feature integration theory. Investigators reported that parallel, rather than serial, search could occur for a variety of feature conjunctions (Enns & Rensink, 1990; McLeod, Driver, & Crisp, 1988; Nakayama & Silverman, 1986; Treisman, 1988; Wolfe, Cave, & Franzel, 1989). Two amendments to the feature integration theory were proposed to accommodate for these findings. One solution involved inhibition of items that shared a nontarget feature (Treisman, 1988; Treisman & Sato, 1990). For example, to find a small red book among small blue books and large red books, one might inhibit the processing of all the large books, causing the small red book to pop out from among the small blue books. We call this proposal search by *inhibitory guidance.* An alternative solution was facilitation of the items that shared a target feature (Cave & Wolfe, 1990; Wolfe et al., 1989). Thus, in this same example, one might excite the processing of all the books that were small, resulting again in the red book popping out from among the blue books. We call this proposal search by *excitatory guidance*.

In an attempt to test between these two alternatives, Driver, McLeod, and Dienes (1992) asked subjects to search for an *X* that oscillated diagonally from top left to bottom right (hereafter called the *target diagonal*). Nontargets were *O*s that oscillated on the target diagonal and *X*s that oscillated diagonally from top right to bottom left (hereafter called the *nontarget diagonal*). Search was most efficient when the movement within each diagonal was in phase (all items on a diagonal moving in the same direction at any one time), and it was least efficient when the movement on each diagonal was out of phase (a random half of the items on a diagonal moving in one direction while the other items moved in the opposite direction). When the movement was in phase on only one diagonal, search was better if the nontarget diagonal motion was in phase than if the target diagonal motion was in phase, suggesting that it may be easier to reject a coherently moving group of items from search (search by inhibitory guidance) than to select a coherently moving group for search (search by excitatory guidance). However, it is important to note that because search efficiency was by far the poorest when the movement on both diagonals was out of phase, it appeared that both inhibitory and excitatory guidance can affect search efficiency.

Recently, Duncan (1995) reevaluated this interpretation. He argued that neither excitatory guidance nor inhibitory guidance alone could provide a satisfactory account of the data because both theories were concerned with the facilitation or inhibition of individual elements in the environment. Because motion phase among elements is, by definition, concerned with the relationships between elements in the environment, theories that emphasize element-by-element search are, by definition, incomplete. In contrast, a theory of visual search that emphasizes interactions among elements within a display is uniquely poised to explain the effects of motion phase on visual search performance. This is precisely the emphasis adopted by Duncan and Humphrey’s (1989, 1992) general theory of visual search.

Thus, according to Duncan (1995), perceptual grouping by motion is crucial to understanding the results Driver et al. (1992) obtained. Items that move in phase form a single coherent perceptual group, which can then be selected for, or rejected from, subsequent target search.

There are, however, at least two serious shortcomings with Duncan’s (1995) proposal. The first difficulty is that it has not been tested. The second difficulty is that no visual search theory, including Duncan and Humphrey’s (1989, 1992) general theory of search, makes strong predictions a priori as to how perceptual grouping by motion will interact with inhibition of a nontarget perceptual group (i.e., inhibitory guidance) or facilitation of a target perceptual group (i.e., excitatory guidance).

The goals of the present study were twofold: to assess what role, if any, perceptual grouping by motion plays in visual search and to determine what rules, if any, govern the relationship between perceptual grouping and search by excitatory and inhibitory guidance. If perceptual grouping does indeed play a crucial role in visual search, then, we hypothesized, it should be possible to fundamentally alter search efficiency by altering the perceptual grouping of target and nontarget elements.

It is likely that search items in the experiment by Driver et al. could be grouped through the *law of common fate* (Wertheimer, 1923), but it is not clear to what extent this was possible, given that there are intrinsic limitations to the grouping of moving elements (see, e.g., Qian, Andersen, & Adelson, 1994; Ullman, 1979; Watson & Humphreys, in press). In order to study the effect of perceptual grouping, we added a dense set of random dots that could move with or against a subset of

Address correspondence to Alan Kingstone, Department of Psychology, University of Alberta, Edmonton, Alberta, Canada T6G 2E9; e-mail: alan@psych.ulberta.ca.

elements, thereby enhancing or diminishing the perceptual grouping among elements.

METHOD

Subjects

Ten subjects (5 males and 5 females), ranging in age from 19 to 29 years (mean = 23.6 years), participated as paid volunteers. All reported normal or corrected-to-normal vision. Testing was conducted in two 1-hr sessions over successive days.

Stimuli and Apparatus

Stimuli were presented on a Tektronix 608 oscilloscope equipped with a P15 phosphor and controlled by a fast point plotter (Finley, 1985) connected to a PC clone. Luminous directional energy of the dots was approximately 3×10^{-9} cd s, and background luminance was approximately 5 cd/m^2.

Stimuli consisted of *X*s and *O*s that were randomly distributed on a rectangular 21 × 21 grid, subject to the constraint that every row and column of the grid contained no more than one element. The size of each element was 10' × 20' of visual angle, that of the rectangular area was 9.5° × 6.9°, and the rectangular area was rotated, on the screen, by 45° counterclockwise from the horizontal. All elements oscillated along either the target diagonal or the nontarget diagonal. Elements moved at a velocity of 5°/s, reversing their direction every 350 ms (i.e., they moved 1.75° before reversing direction).

In all conditions, the target was an *X* oscillating on the target diagonal. Nontargets consisted of an equal number of *O*s oscillating on the target diagonal and *X*s oscillating on the distractor diagonal. The target was present in 50% of the trials. On nontarget trials, the target was replaced by an *O*. The kind of display (target-present vs. target-absent display), and the number of display elements (9, 15, or 21 items), varied randomly from trial to trial.

Elements on either diagonal could oscillate in phase (i.e., all elements moving in one direction at the same time) or out of phase (i.e., half the elements moving in one direction while half moved in the opposite direction). Thus, there were four oscillation patterns: elements on the target diagonal in phase and elements on the distractor diagonal in phase (in-in), elements on the target diagonal in phase and elements on the distractor diagonal out of phase (in-out), elements on the target diagonal out of phase and elements on the distractor diagonal in phase (out-in), and elements on the target diagonal out of phase and elements on the distractor diagonal out of phase (out-out).

In addition, 300 dots were randomly distributed in the rectangular stimulus area. The dots all moved together in one of five ways: moving in phase with the target, moving (counterphase) against the target, moving in phase with the elements on the nontarget diagonal, moving (counterphase) against the elements on the nontarget diagonal, or not moving at all.[1] The static-dot condition established baseline performance for when random dots were simply displayed among elements that moved on the target and nontarget diagonals. Figure 1 illustrates a trial with the target diagonal in phase and nontarget diagonal out of phase (in-out).

1. When the elements on the distractor diagonal were out of phase, dot motion on the nontarget diagonal was by definition always in phase with half of these elements. For coding purposes, dot motion was considered to be in phase when the dots moved from bottom left to top right at the same time that elements in the target diagonal moved from bottom right to top left, and vice versa.

Procedure

The four phase conditions and five dot conditions yielded 20 different phase-dot combinations. Each of these 20 conditions was run in a separate block, with each block composed of 96 trials. Within each block, a rest break was provided after 48 trials. Subjects received 11 practice trials at the beginning of each block. The order of the 20 conditions was counterbalanced across subjects.

On each trial, the stimulus display was presented until the subject pressed one of the two response buttons, indicating whether the target was present or absent. After a delay of 800 ms, the next trial was initiated. Subjects were instructed to respond as quickly as possible while maintaining a high level of accuracy. Feedback was provided by a beep that sounded after each incorrect response. Subjects were told their response accuracy at the end of each block. Subjects rested after each block.

RESULTS

Because the task instructions emphasized speed, the main data analyses were of correct response time (RT). There were also systematic differences in accuracy, but in no condition was the reported RT effect contradicted by the accuracy data.

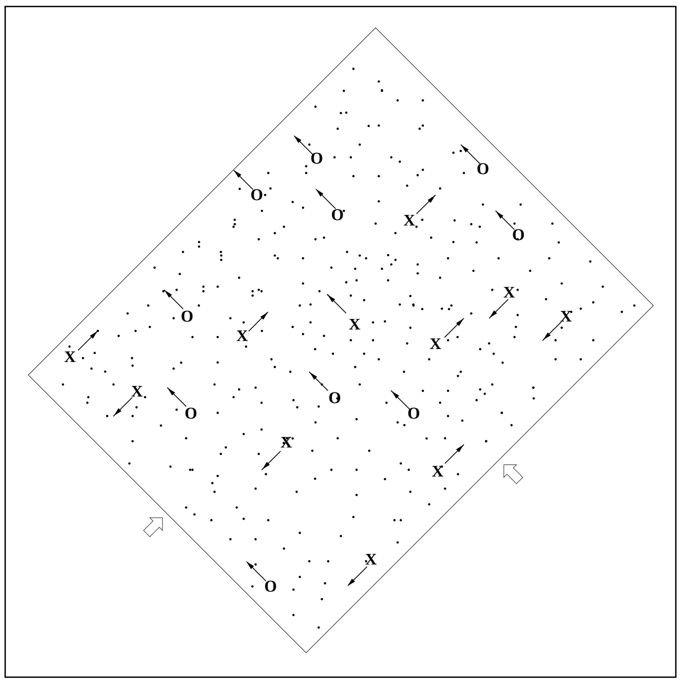

Fig. 1. Example of a target-present display at a set size of 21 in the static-dot condition when the elements on the target diagonal (top left to bottom right) are moving in phase and the elements on the nontarget diagonal (top right to bottom left) are moving out of phase. The solid arrows depict the current motion direction of each element. The target is the letter *X* in the center of the display (as indicated by the two arrows outside the frame) moving with the distractor letter *O*s. The distractor *X*s are moving on the nontarget diagonal. Elements moved at a velocity of 5°/s, reversing their direction every 350 ms (i.e., 1.75°).

An analysis of variance (ANOVA) was conducted for the target-present and -absent response data. Each analysis had 5 (dot condition) × 4 (phase condition) × 3 (set size) levels. Because all main effects and interactions were significant ($p < .002$), separate ANOVAs were conducted for each dot condition with phase and set size as factors. Fisher's (protected t) tests ($p < .05$) were used to test specifically for differences in these factors. RT patterns were the same for the target-present and -absent responses, although overall, RTs were significantly slower, and search slopes significantly steeper, for the target-absent responses. Table 1 gives the mean RTs and error rates for target-present and target-absent responses in each of the dot, phase, and set-size conditions, as well as the mean regression slopes of RT across set size.

Figure 2 shows the mean RTs for target-present responses in the static- and moving-dot conditions as a function of phase and set size.

Dots Static

The static-dot condition established baseline performance patterns. Phase, set size, and their interaction were significant ($p < .0001$). RT and search slopes differed significantly between all phase conditions, with two exceptions. First, as suggested in Figure 2, performance was the same when the items on the target and nontarget diagonals moved in phase (in-in) and when only the nontarget-diagonal items moved in phase (out-in). Second, the steep search slope that was observed when the items on both diagonals moved out of phase (out-out) was the same as the slope when only the items on the nontarget diagonal moved out of phase (in-out).

These data replicate the key findings of Driver et al. (1992): Search was relatively easy when items on the target and nontarget diagonals were in phase, and extraordinarily hard when items on both diagonals were out of phase. When the elements were in phase on one diagonal and out of phase on the other, search was far easier if the items were in phase on the nontarget diagonal than in phase on the target diagonal. This latter difference suggests that it is far easier to use motion phase to reject items from search (inhibitory guidance) than it is to use motion to select items for search (excitatory guidance). However, because search performance was worse when all items were out of phase, the data indicate that both excitatory and inhibitory guidance can be used to control search.

Table 1. *Response time (in seconds) and accuracy (% correct) in detecting the presence or absence of a target as a function of phase, set size, and dot condition*

	Phase condition											
	All elements in phase			Target diagonal in phase			Nontarget diagonal in phase			All elements out of phase		
Set size	Yes	No	% correct	Yes	No	% correct	Yes	No	% correct	Yes	No	% correct
Dots with target												
9	0.877	1.139	98	0.918	1.315	96	0.834	1.293	97	0.927	1.363	96
15	1.004	1.473	99	1.093	1.698	94	0.943	1.525	99	1.085	1.729	97
21	1.184	1.726	94	1.200	1.897	95	1.080	1.761	94	1.192	1.915	94
Slope[a]	26	49		24	49		21	39		22	46	
Dots against target												
9	1.033	1.460	98	1.486	2.471	95	1.073	1.528	96	1.599	2.480	92
15	1.274	1.786	94	2.071	3.307	90	1.182	1.809	95	2.002	3.236	92
21	1.401	2.106	93	2.494	4.082	88	1.325	2.137	93	2.555	3.941	90
Slope[a]	31	54		84	134		21	51		80	122	
Dots with nontargets												
9	0.798	0.958	96	1.404	1.837	92	0.816	0.999	96	1.658	1.982	91
15	0.885	1.130	96	1.769	2.582	93	0.919	1.255	94	1.911	2.670	87
21	0.909	1.408	95	1.883	2.887	87	0.981	1.473	94	2.200	3.272	86
Slope[a]	9	38		40	88		14	40		45	108	
Dots against nontargets												
9	1.334	1.677	96	1.353	1.702	96	1.566	1.889	97	1.626	2.093	95
15	1.474	2.330	94	1.539	2.246	93	1.966	2.671	91	2.168	2.877	89
21	1.811	2.764	93	1.659	2.722	91	2.085	3.339	89	2.405	3.483	88
Slope[a]	40	91		26	85		43	121		65	116	
Dots static												
9	1.308	1.728	95	1.452	2.199	92	1.416	1.797	96	1.803	2.586	93
15	1.684	2.350	93	1.872	3.092	85	1.685	2.420	88	2.340	3.744	89
21	1.851	2.962	90	2.506	3.748	85	1.945	2.826	89	3.006	4.535	83
Slope[a]	45	103		88	129		44	86		100	162	

[a]In ms/item.

Dots Moving With Target Phase

The moving dots formed a coherent perceptual group. As shown in Figure 2, when this group was moved in phase with the target, search was very fast and easy. Indeed, moving the dots in phase with the target abolished all the effects of phase observed in the static-dot condition and by Driver et al. (1992), with only set size affecting performance ($p < .0001$). These data indicate that subjects found it easy and efficient to simply search for the target among the moving dots.

Dots Moving Against Target Phase

When the dots moved against the target, the effects of phase and set size and their interaction were significant ($p < .0002$). Search was slow, and the slopes steep, when all the elements moved out of phase or only the elements on the target diagonal were out of phase. These two phase conditions did not differ. Similarly, search was relatively fast, and the slopes shallow, when all the elements moved in phase or only the elements on the nontarget diagonal were in phase. These two phase conditions did not differ.

The key finding is that when the dots moved against the target, it did not matter whether the elements on the target diagonal moved in phase or out of phase. Thus, the motion of the dots neutralized search by excitatory guidance. Nevertheless, search was relatively easy when the elements on the nontarget diagonal moved in phase, suggesting that subjects were able to reject these items from search (search by inhibitory guidance).

Dots Moving With the Nontarget-Diagonal Phase

When the dots moved in phase with the elements on the nontarget diagonal, the effects of phase and set size and their interaction were significant ($p < .0001$). RT and search slopes differed significantly between all phase conditions, save for the two exceptions previously observed in the static-dot condition. And as was the case in the static-dot condition, these data indicate that both excitatory and inhibitory guidance can be used to control search, although they suggest inhibitory guidance is easier. Indeed, dots moving in phase with the items on the nontarget diagonal greatly accentuated the advantage of inhibitory guidance.

Dots Moving Against the Nontarget-Diagonal Phase

The advantage for inhibitory guidance can also be reversed by dot motion. When the dots moved against the nontarget-diagonal phase, the effects of phase and set size and their interaction were significant ($p < .0001$). RT and search slopes differed significantly between all phase conditions, with one exception. Search was fastest, and the slopes the most shallow, when the items on the target and nontarget diagonals moved in phase and when only the items on the target diagonal moved in phase. These two phase conditions did not differ. Note that rejecting the in-phase items on the nontarget diagonal from search (inhibitory guidance) produced slower and steeper search than selecting the in-phase items on the target diagonal for search (excitatory guidance). There was no difference between the two conditions with the target diagonal in phase.

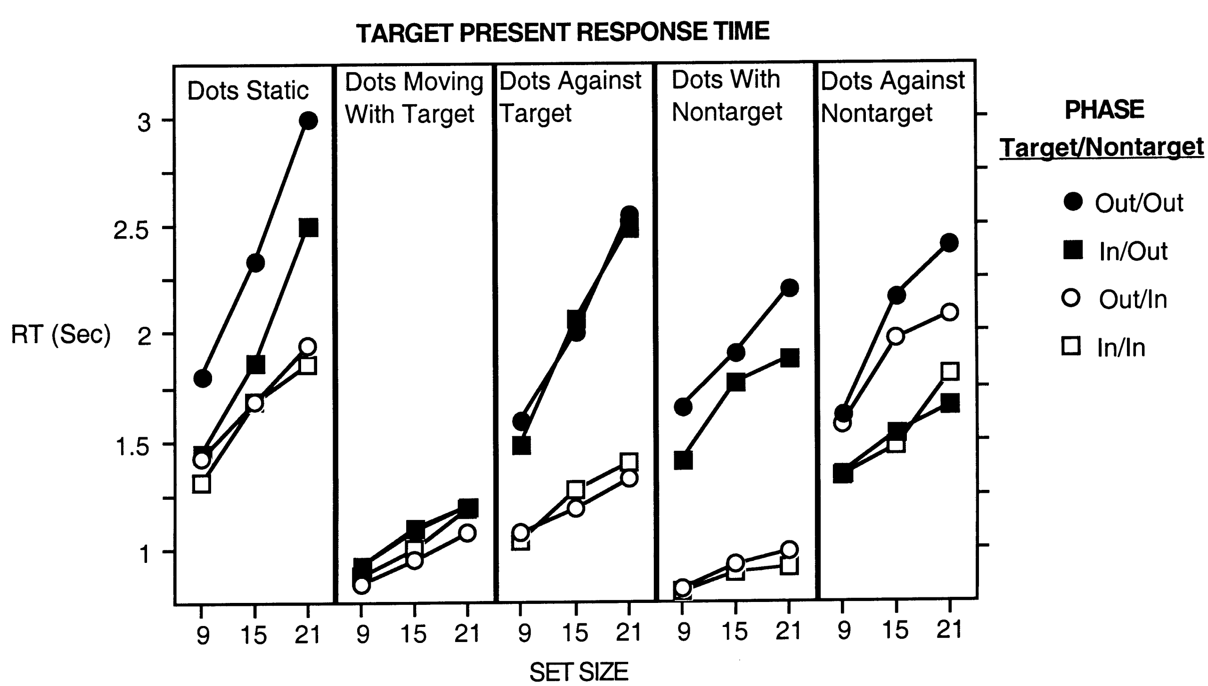

Fig. 2. Response time (RT) to detect the presence of a target stimulus as a function of phase, set size, and dot condition.

DISCUSSION

Our data demonstrate that perceptual grouping by motion phase plays a crucial role in search for a conjunction of motion and form. One can abolish all effects of target and nontarget motion phase by simply moving a coherent set of random dots in phase with the target motion. Or one can selectively eliminate the effect of target motion phase by moving the dots counterphase with the target motion. Similarly, one can accentuate the fact that rejecting items on the nontarget diagonal from search (inhibitory guidance) is generally easier than selecting items for search on the target diagonal (excitatory guidance) simply by moving dots in phase with the items on the nontarget diagonal. And one can even reverse the advantage that inhibitory guidance typically enjoys over excitatory guidance by moving the dots counterphase with the items on the nontarget diagonal.

The second goal of the present study was to determine what rules govern the relationship between perceptual grouping and search for a conjunction of motion and form. A reasonable explanation of the present data can be captured by adopting three hierarchically organized general rules:

- *Rule 1: Search is easier when there are fewer motion groups.* Search was easier as the number of different motion groups decreased. For example, in the static-dot condition, there were four motion goups when all the items moved out of phase, and two motion groups when all the items moved in phase.
- *Rule 2: Inhibitory guidance is easier than excitatory guidance.* Search was easier if the items on the nontarget diagonal all moved together and could be rejected from further consideration than if the elements on the target diagonal moved together and could be selected for further processing.
- *Rule 3: Search gets easier as the dots move with more elements.* There were four possible motion groups: two on the target diagonal and two on the nontarget diagonal. The moving dots could move in phase with all, half, or none of the elements on either diagonal. Search was easiest when all the elements on a diagonal moved with the dots, harder when half the items on a diagonal moved with the dots, and hardest when none of the items moved with the dots. Of course, if the target moved in phase with the dots, search was always easy.

These three rules serve as good predictors of search performance, as is illustrated in Figure 3. The left-hand panel plots search slope on target-present trials as a function of predicted rank ordering of search difficulty. The right-hand panel plots search slope for a weighted average of target-absent and target-present trials as a function of predicted rank ordering of search difficulty. It is important to note that although Figure 3 emphasizes search slope, the effects were occasionally in the search intercept (see Fig. 2).

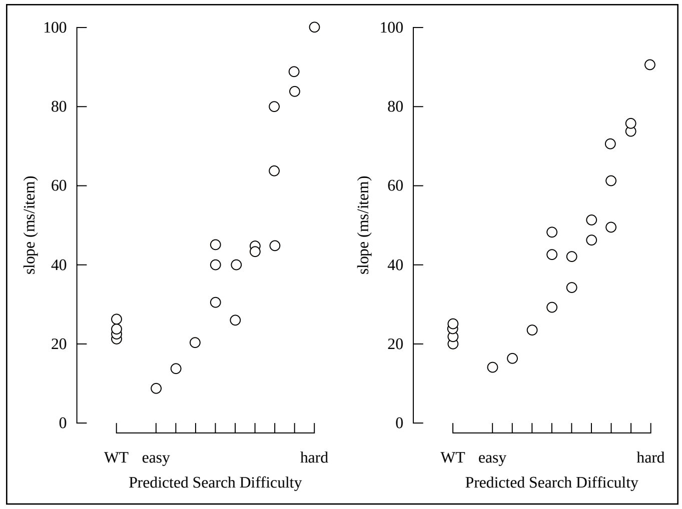

Fig. 3. Performance as a function of the predicted rank ordering of search difficulty. Rank ordering is based on the hierarchical application of the three rules discussed in the text. Each tick on the *x*-axis represents a predicted increase in search difficulty (e.g., an increase in the number of motion groups). The left panel shows search slopes for target-present responses. The right panel shows search slopes obtained from target-present and target-absent slopes using the following weighted average: slope(average) = 0.25*(2*slope(target present) + slope(target absent)). These slopes thus indicate approximately the average extra time spent per distractor. WT refers to the condition when the dots moved in phase with the target.

SUMMARY

Duncan (1995) argued that the results of Driver et al. (1992) ruled out strictly element-by-element approaches to the problem of conjunction search, and hypothesized that approaches emphasizing perceptual grouping would prove more promising. We tested this hypothesis and discovered that perceptual grouping plays a crucial role in search for a conjunction of movement and form. As the number of different perceptual groups decreases, search performance becomes easier. Two other factors were also important predictors of search performance. First, it is easier to reject all the elements of a nontarget group (search by inhibitory guidance) than it is to select all the elements of a target group for further processing (search by excitatory guidance). Second, the more elements captured by a perceptual group, the easier the search.

Finally, it is important to note that the results of the present study provide convincing evidence of the preattentive organization of a visual display into surfaces defined by common motion, with an upper limit of two to three surfaces (Andersen, 1989).[2] That is, grouping by common motion occurs prior to guided visual search and as such is similar to preattentive organization on the basis of element clusters (Trick & Enns, 1997), texture properties (He & Nakayama, 1994), texture gradients (Aks & Enns, 1996), stereodepth (Nakayama & Silverman, 1986), shape from shading (Ramachandran, 1988), and direction of lighting (Enns & Rensink, 1990). In other words, the representations on which visual search operates are in many respects based on extended, intrinsic object properties rather than on localized imaged properties (see Watson & Humphreys, in press, for a similar conclusion when grouping is based on linear and rotational motion).

2. We thank Jim Enns for bringing this point to our attention.

Acknowledgments—This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) to both authors and by a grant from the Alberta Heritage Foundation for Medical Research to the first author. We are grateful to Michelle Kincade and Hannah Pazderka-Robinson for their invaluable help with the running of the study. We would also like to thank John Duncan, James T. Enns, Glyn Humphreys, and Raymond Klein for feedback on a previous version of this article.

REFERENCES

Aks, D.J., & Enns, J.T. (1996). Visual search for size is influenced by a background texture gradient. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1467–1481.

Andersen, G.J. (1989). Perception of three-dimensional structure from optic flow without locally smooth velocity. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 363–371.

Cave, K.R., & Wolfe, J.M. (1990). Modelling the role of parallel processing in visual search. *Cognitive Psychology, 22,* 255–271.

Driver, J., McLeod, P., & Dienes, Z. (1992). Motion coherence and conjunction search: Implications for guided search theory. *Perception & Psychophysics, 51,* 79–85.

Duncan, J. (1995). Target and nontarget grouping in visual search. *Perception & Psychophysics, 57,* 117–120.

Duncan, J., & Humphreys, G.W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458.

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

Enns, J.T., & Rensink, R.A. (1990). Influence of scene-based properties on visual search. *Science, 247,* 721–723.

Finley, G. (1985). A high-speed point plotter for vision research. *Vision Research, 25,* 1993–1997.

He, Z.J., & Nakayama, K. (1994). Perceiving textures: Beyond filtering. *Vision Research, 34,* 151–162.

McLeod, P., Driver, J., & Crisp, J. (1988). Visual search for a conjunction of movement and form is parallel. *Nature, 332,* 154–155.

Nakayama, K., & Silverman, G. (1986). Serial and parallel processing of visual feature conjunctions. *Nature, 320,* 264–265.

Qian, N., Andersen, R.A., & Adelson, E.H. (1994). Transparent motion perception as detection of unbalanced motion signals. I. *Psychophysics, The Journal of Neuroscience, 14,* 7357–7366.

Ramachandran, V.S. (1988). Perceiving shape from shading. *Scientific American, 259,* 76–83.

Treisman, A. (1986). Features and objects in visual processing. *Scientific American, 255,* 106–115.

Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *The 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., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 459–478.

Trick, L.M., & Enns, J.T. (1997). Clusters precede shapes in perceptual organization. *Psychological Science, 8,* 124–129.

Ullman, S. (1979). *The interpretation of visual motion.* Cambridge, MA: MIT Press.

Watson, D.G., & Humphreys, G.W. (in press). Segmentation on the basis of linear and local rotational motion: Motion grouping in visual search. *Journal of Experimental Psychology: Human Perception and Performance.*

Wertheimer, M. (1923). Untersuchungen zur lehre von der Gestalt, II. *Psychologische Forschung, 4,* 301–350.

Wolfe, J.M., Cave, K.R., & Franzel, S.L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 419–433.

(RECEIVED 2/13/98; ACCEPTED 4/30/98)