Notes and Comment

Target and nontarget grouping in visual search

JOHN DUNCAN

MRC Applied Psychology Unit, Cambridge, England

Results recently reported by Driver, McLeod, and Dienes (1992) are used to contrast three accounts of visual search—in particular, their mechanism for easy conjunction search. In the Driver et al. study, the target was defined by a conjunction of form and movement; the key manipulation was phase in both target and nontarget motion sets. Mechanisms working separately on each display element (inhibition from nontarget features, facilitation from target features) are unable to explain large effects of phase, since this is defined only by relationships between one element and another. As implemented in the guided search model of Cave and Wolfe (1990), local suppression between similar elements is also unable to account for the results. More promising is an approach based on perceptual grouping: Elements moving in phase can be selected (target motion) or rejected (nontarget motion) as a group. Rather than a bias against elements that are similar to or grouped with their neighbors, there is a bias to treat grouped elements together.

A common aspect of many theories of visual attention is that objects or elements in the visual field *compete* with one another for visual analysis and/or control of behavior (Bundesen, 1990; Cave & Wolfe, 1990; Duncan & Humphreys, 1989; Treisman, 1988; see also Broadbent, 1958; Neisser, 1967). Subjectively, this competition is reflected in limited attentional capacity. A test bed for alternative theories of competition has become visual search. In the search task, the subject detects or identifies a single target object presented in an array of multiple nontargets. Performance reflects how easily competition among the multiple display elements is resolved in favor of the target (Bundesen, 1990; Cave & Wolfe, 1990; Duncan, 1980; Treisman, 1988).

In this note, I shall discuss a recent search experiment described by Driver, McLeod, and Dienes (1992). Though I shall be agreeing with their main conclusions, I shall take a different view of the compatibility between these conclusions and several recent, more general theories of search (Cave & Wolfe, 1990; Duncan & Humphreys, 1989; Treisman, 1988). In fact, I shall suggest that their

results provide severe difficulties for at least two of these theories.

The Experiment

In the Driver et al. experiment, subjects searched for an X oscillating along the minor display diagonal (top left to bottom right) among Os oscillating along this same diagonal and Xs oscillating along the major diagonal (top right to bottom left). All characters moved at 2.1°/sec and reversed direction simultaneously every 100 msec. Thus, the task was a form of conjunction search (Treisman & Gelade, 1980); it is now well known that conjunction search in general varies from very difficult (Treisman & Gelade, 1980) to very easy (McLeod, Driver, & Crisp, 1988; Nakayama & Silverman, 1986), and it is widely accepted that these are important results for different search theories to explain.

In fact, in their experiment, Driver et al. were able to manipulate difficulty very widely. Their manipulation was one of phase. Elements moving on the minor diagonal (target motion) oscillated either in phase (i.e., all moving in the same direction at any one time) or out of phase (a random half of the elements moving to the top left when the other half moved to the bottom right, and vice versa). So too did elements moving on the major diagonal (nontarget motion). Independent manipulation of these two phases produced four conditions, with minor diagonal only, major diagonal only, both, or neither in phase. Search was best when both minor- and major-diagonal sets were in phase, and it deteriorated only slightly when one set or the other was out of phase. With both sets out of phase, however, the task became catastrophically difficult. In this condition, the slope of the function relating search time to the number of display elements was 78 msec/item for target-present trials and 200 msec/item for target-absent trials.

Feature Integration Theory

Among various possible explanations for these results, the first that Driver et al. considered was a mechanism originally proposed by Treisman (1988) to account for cases of easy conjunction search. According to Treisman's (1988) feature integration theory, the first stage of visual analysis is parallel registration of elementary features and their locations in separate spatiotopic maps of the visual field. Each element in an array also activates the corresponding location in a spatial "master map." In this theory, "attention" is a subsequent serial process, operating in tasks like these on one element after another. To explain instances of easy conjunction search, Treisman (1988) proposed that the master map might be

The work was supported in part by a grant from the Human Frontier Science Program. I am grateful to Kyle Cave, Jon Driver, and Jeremy Wolfe for helpful comments on earlier drafts of the paper. Correspondence should be addressed to J. Duncan, MRC Applied Psychology Unit, 15 Chaucer Rd., Cambridge CB2 2EF, England.

used to determine which element in the field is attended first. Inhibitory connections might be set up between nontarget feature maps (in Driver et al.'s experiment, O maps and major-diagonal motion maps) and the master map, so that each nontarget element tended to send some inhibition to its corresponding location in the master map. If the process were sufficiently reliable, then directing attention to the most active location in the master map would ensure that the target was the element processed first.

Driver et al. concluded that their results might be partially explained by such a mechanism. In particular, they thought the mechanism was well suited to explaining the effects of phase in the major-diagonal (nontarget) motion set. To develop their explanation, they assumed that inhibition of the master map from the major-diagonal motion feature would be strongest when all movements along this diagonal were in phase. In this case, search could at least be restricted to elements with minor-diagonal motion.

In contrast, I would suggest that, at least as it stands, Treisman's (1988) proposal offers no satisfactory way to account for the results. Phase is defined by relationships between *one element and another within the display*; however, Treisman's mechanism works element by element. In Driver et al.'s experiment, the effective strategy would have been to set up inhibition between two motion maps (top left to bottom right, bottom right to top left) and the master map. Given this strategy, it simply should not have mattered whether the motions of different nontargets were in or out of phase. In either case, each individual nontarget's moment-by-moment inhibition of the master map would have been the same.

The general point suggested by the importance of phase is that search depends critically on interactions between one display element and another. The same point is made by other demonstrations of the importance of perceptual grouping between nontargets (Banks & Prinzmetal, 1976; Bundesen & Pedersen, 1983; Duncan & Humphreys, 1989; Farmer & Taylor, 1980). Of course, this is not to argue that feature integration theory is wrong in other respects, or indeed that it cannot be supplemented by grouping mechanisms (see, e.g., Treisman, 1982). The particular mechanism proposed to account for easy conjunction search, however, cannot deal satisfactorily with Driver et al.'s results.

Guided Search

The guided search theory (Cave & Wolfe, 1990) was developed as a revision of feature integration theory and shares many of the same mechanisms. In particular, there are the same ideas of initial parallel feature maps and use of the master map to guide attention serially from one display element to another. Rather than nontarget inhibition, however, Cave and Wolfe (1990) use excitation from target feature maps to ensure that master map activation is greatest for the target. In the Driver et al. experiment, excitation from minor-diagonal

motion and X maps could (if sufficiently reliable) ensure that the target was processed first.

By the same reasoning as before, Driver et al. thought this mechanism was well suited to explain the importance of phase in the minor-diagonal (target) motion set. They assumed that excitation from the target-motion feature would be strongest when all elements sharing this motion were in phase. For the same reason as before, I disagree with this conclusion. Again, the results suggest the importance of interactions within the display, while the proposed mechanism works element by element.

In fact, the guided search theory does incorporate an additional mechanism of interaction within the display; but this mechanism only compounds its difficulties. According to the theory, activation in the master map derives from two sources. First, each location gains activity to the extent that the element in that location has target features, as already considered. Second, each location gains activity to the extent that the element in that location differs from other elements in the array (cf. Sagi & Julesz, 1984; Ullman, 1984). The intention is to bias attention toward unusual or mismatching elements in the field. The mechanism allows the theory to deal with well-known effects of nontarget homogeneity (e.g., Gordon, 1968).

To apply this theory to the Driver et al. experiment, two things must be borne in mind. First, increasing or decreasing the net similarity between each element and its neighbors has little effect (Cave & Wolfe, 1990, p. 255). The order of attending to the different items in an array is determined by *relative* activation; it is little affected by increasing or decreasing all activations equally. Second, comparisons between array elements are carried out independently for different features, such as form and motion. In all conditions of the Driver et al. experiment, each element shared form with about half of the others. This feature can accordingly be ignored; relative activations of different display elements could have been influenced only by differential motion matches.

Predictions may now be derived as follows. First, performance should have been much the same whether both motion sets were in phase or both were out of phase. If both sets were in phase, then every element in the array shared motion with half of the remaining elements. If both sets were out of phase, then every element in the array shared motion with a quarter of the remaining elements. In neither case was any element favored over another in terms of net inhibition from its neighbors. Performance should have been worst when the minor-diagonal (target-motion) set was in phase while the major-diagonal (nontarget-motion) set was out of phase. Because they received inhibition from fewer same-direction neighbors, elements in the major-diagonal set should have gained a competitive advantage. Complementarily, performance should have been best when the minor-diagonal set was out of phase while the major-diagonal set was in phase. At least as it is implemented in this theory, a simple bias against elements that are similar to their neighbors does not explain the effects of phase manipulation.

A possible line for guided search theory would be to argue that manipulations of phase allowed different patterns of excitation between feature and master maps.¹ If elements in the target-motion set were in phase, for example, it might have been possible to set up excitatory connections to the master map from top-left-to-bottom-right motion just for the first 100 msec of the display, from bottom-right-to-top-left motion just for the next 100 msec, from top-left-to-bottom-right again for the next 100 msec, and so on. Clearly, this strategy would not have been possible with the target-motion set out of phase. A complementary line might be taken for Treisman's (1988) mechanism of nontarget inhibition. It seems extremely unlikely, however, that control settings for search can be changed so quickly—for example, that subjects can search for an object moving left (ignoring objects moving right) for 100 msec, for objects moving right (ignoring objects moving left) for the next 100 msec, and so on. Many estimates suggest that it takes hundreds of milliseconds to establish a new endogenous attentional bias—for example, a bias to elements in a certain region of space (e.g., Müller & Rabbitt, 1989). Doubtless the same is true for bias toward one direction of motion rather than another. Much more likely is that Driver et al.'s results reflect some kind of interaction between similar elements in a display.

It would be hard to argue that no possible account based on bias toward mismatching elements could explain Driver et al.'s results. An attractive alternative, however, derives from a consideration of perceptual grouping. Indeed, as I shall describe next, this is consistent both with subjects' reports in such tasks and with the more detailed account of the findings that Driver et al. themselves presented.

Perceptual Grouping and Weight Linkage

A common report in cases of easy conjunction search is that subjects segment the display into two interleaved perceptual groups, subjectively only "searching" among the target group (e.g., McLeod, Driver, Dienes, & Crisp, 1991; Nakayama & Silverman, 1986). For example, if the task is to find a moving X among moving Os and static Xs, the impression is that the moving items can be perceived as a single coherent group. Confining attention to this group reduces the task to simple search for an X among Os.

In line with such reports, Driver et al. offered the following account of their results. When display elements moved together in phase, they were perceived as a single coherent group. When they moved out of phase, grouping broke down. Two strategies were then available to allow easy search. If elements moving along the minor diagonal were in phase, then it was possible first to restrict attention to this group, then to search within it for an X. If elements moving along the major diagonal were in phase, then it was possible first to eliminate this group from consideration, then to search for an X among the remainder. Performance became catastrophically poor, however, when neither strategy was possible.

Two-stage search strategies of this general sort have been considered by various authors (e.g., Egeth, Virzi, & Garbart, 1984; Treisman & Sato, 1990; Wolfe, 1994). Whether or not a strict two-stage process is assumed (Treisman & Sato, 1990), the results strongly support the importance of perceptual grouping. A generalized suppression of all elements with major-diagonal motion can be achieved, providing they form a strong perceptual group. A generalized facilitation of minor-diagonal elements can similarly be achieved with strong grouping. Neither process is possible, however, when grouping is weak.

This account depends on perceptual grouping—that is, on interactions within the display rather than element-by-element processing. Such interactions, however, do not take the form of a simple bias against grouped or similar elements. Instead, perceptual grouping makes a set of elements easy either to select or to reject together, depending on their relevance to the task.

In fact, this is the approach to perceptual grouping taken in a third general theory of search, the attentional engagement theory of Duncan and Humphreys (1989, 1992). In this theory, too, display elements compete for limited processing capacity. Each element is assigned a weight indicating how strongly it is activated or competes; perceptual grouping operates by a mechanism of "weight linkage," such that the activations of strongly grouped elements tend to rise or fall together. Competition is not biased toward or against grouped elements; instead it is biased to treat grouped elements together. This mechanism deals with the well-known result that search is facilitated by nontarget grouping (Bundesen & Pedersen, 1983; Farmer & Taylor, 1980). Equally, however, it deals with the beneficial effects of target grouping in the more general case of displays containing multiple targets as well as multiple nontargets (e.g., Kahneman & Henik, 1977). A weight linkage mechanism of this sort could explain why search can be directed either by suppressing all elements in a nontarget-motion group or by facilitating all elements in a target-motion group.

Again, this is not to suggest that attentional engagement theory is right in all respects. For example, if a two-stage search strategy is assumed, the theory has little to say about the transition from initial facilitation of all elements in the target-motion set to selection of the X in this set at the expense of the Os. The data do support the general idea, however, that grouping makes a set of elements easy either to reject or to select together.

Summary

The results of Driver et al. rule out strictly element-by-element approaches to the problem of easy conjunction search. Neither can they be explained by a simple bias toward mismatching display elements, at least as it is implemented in guided search. Such a bias is the only means of within-display interaction in many current models (Sagi & Julesz, 1984; Ullman, 1984). Instead, the results suggest that perceptual grouping brings a bias

toward common fate. For nontarget groups, this fate is rejection from further consideration; for target groups, complementarily, it is selection for further processing.

REFERENCES

- BANKS, W. P., & PRINZMETAL, W. (1976). Configurational effects in visual information processing. *Perception & Psychophysics*, **19**, 361-367.
- BROADBENT, D. E. (1958). *Perception and communication*. London: Pergamon.
- BUNDESEN, C. (1990). A theory of visual attention. *Psychological Review*, **97**, 523-547.
- BUNDESEN, C., & PEDERSEN, L. F. (1983). Color segregation and visual search. *Perception & Psychophysics*, **33**, 487-493.
- CAVE, K. R., & WOLFE, J. M. (1990). Modeling the role of parallel processing in visual search. *Cognitive Psychology*, **22**, 225-271.
- DRIVER, J., MCLEOD, P., & DIENES, Z. (1992). Motion coherence and conjunction search: Implications for guided search theory. *Perception & Psychophysics*, **51**, 79-85.
- DUNCAN, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review*, **87**, 272-300.
- 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 & Performance*, **18**, 578-588.
- EGETH, H. E., VIRZI, R. A., & GARBART, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 32-39.
- FARMER, E. W., & TAYLOR, R. M. (1980). Visual search through color displays: Effects of target-background similarity and background uniformity. *Perception & Psychophysics*, **27**, 267-272.
- GORDON, I. E. (1968). Interactions between items in visual search. *Journal of Experimental Psychology*, **76**, 348-355.
- KAHNEMAN, D., & HENIK, A. (1977). Effects of visual grouping on immediate recall and selective attention. In S. Dornic (Ed.), *Attention and performance VI* (pp. 307-332). Hillsdale, NJ: Erlbaum.
- MCLEOD, P., DRIVER, J., & CRISP, J. (1988). Visual search for a conjunction of movement and form is parallel. *Nature*, **332**, 154-155.
- MCLEOD, P., DRIVER, J., DIENES, Z., & CRISP, J. (1991). Filtering by movement in visual search. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 55-64.
- MÜLLER, H. J., & RABBITT, P. M. A. (1989). Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 315-330.
- NAKAYAMA, K., & SILVERMAN, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, **320**, 264-265.
- NEISSER, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.
- SAGI, D., & JULESZ, B. (1984). Detection versus discrimination of visual orientation. *Perception*, **13**, 619-628.
- TREISMAN, A. M. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 194-214.
- TREISMAN, A. M. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology*, **40A**, 201-237.
- TREISMAN, A. M., & GELADE, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, **12**, 97-136.
- TREISMAN, A. M., & SATO, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 459-478.
- ULLMAN, S. (1984). Visual routines. *Cognition*, **18**, 97-159.
- WOLFE, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, **1**, 202-238.

NOTE

1. I thank Jon Driver for pointing out this possibility.

(Manuscript received April 13, 1994;
revision accepted for publication July 26, 1994.)