Notices and Announcements

Relating attention to visual mechanisms

GORDON L. SHULMAN

*Washington University School of Medicine
St. Louis, Missouri*

The effect of attention on visual perception has been the subject of great controversy. Much research on attention over the last 30 years has been framed by the debate on this question between proponents of early and late selection.

Late-selection theorists (Deutsch & Deutsch, 1963; Duncan, 1980; Keele, 1973) have argued that all perceptual encoding, including recognition, and certain semantic analyses are accomplished in parallel, whereas early-selection theorists (Broadbent, 1982) have countered that only simple "physical" analyses can be conducted in parallel. More sophisticated analyses of shape that support object recognition and memory access are conducted by limited capacity systems.¹

Although the terms *perceptual* and *semantic* include a wide variety of processes, there has been an unfortunate tendency to extrapolate conclusions to the entire collection of processes based on results from a few. For example, researchers have tacitly assumed that, if any evidence for semantic processing of unattended material is found, then perceptual operations must be parallel. This inference, however, only applies to the perceptual operations relevant to the recognition of the experimental stimuli, which are generally upright block alphanumeric characters presented in the image plane. Recognition of these stimuli does not require the resolution of a vast number of problems the visual system must solve (the analysis of motion, three-dimensional space, constancies of various sorts, etc.).

More direct evidence that perceptual operations are parallel has been collected using visual search tasks (Egeth, Jonides, & Wall, 1972; Neisser, 1963) in which the experimenter measures the reaction time to decide whether a prespecified target appears in an array. If reaction time is independent of array size, the operations underlying search are considered parallel; if reaction time increases with array size in a linear fashion, the underlying operations are assumed to be serial (or parallel but capacity limited; Townsend, 1972). Visual search studies over the last several decades have detailed the conditions over which search is parallel or serial. Since early studies

indicated that search was particularly efficient when targets were defined by gross physical differences (Neisser, 1963), researchers during the 1970's attempted to demonstrate that parallel search could be based on memory-dependent qualities, such as category (Duncan, 1983; Egeth et al., 1972; Schneider & Fisk, 1984; Shiffrin & Schneider, 1977). In the 1980s, Treisman and colleagues (Treisman & Gelade, 1980; Treisman & Gormican, 1988; Treisman & Souther, 1985) attempted to define more carefully the physical features that produced parallel search. They concluded that single features could be analyzed in parallel but that conjunctions of features required attention.

Results demonstrating parallel search for single features should not be overgeneralized. Early studies focused mainly on search defined by form or color. Although recent studies have expanded the list of attributes under consideration (Nakayama & Silverman, 1986; Wolfe & Franzel, 1988), a large number of visual processes have still not been considered.

Most visual search studies yielding parallel functions have also involved fairly crude discriminations, such as S from T and X, or blue from brown and green (Treisman & Gelade, 1980). More difficult discriminations generally yield serial functions (Bundesen & Pedersen, 1983; Gardner, 1973; Treisman & Gormican, 1988; see Duncan & Humphreys, 1989, for a recent discussion of the effects of target and distractor similarity in visual search). Although visual search techniques have not been applied to most of the threshold tasks studied by psychophysicists (vernier acuity, stereoacuity, wavelength discrimination, etc.), one would expect them to show serial functions when threshold stimuli are used as targets. Given the usual interpretation of visual search functions, one might conclude that most mechanisms underlying early vision are capacity limited. However, there are difficulties in the interpretation of serial search functions.

Since the development of visual search techniques in the 1960s and 1970s (Egeth et al., 1972; Neisser, 1963), it has been widely appreciated that the logic relating display-size effects to attentional limitations is flawed (Eriksen & Spencer, 1969; Shiffrin & Gardner, 1972). As one increases display size, one also increases the probability of noise-generated false alarms. It is for this reason that the visual search technique has not been applied to the study of thresholds and fine discriminations so widely studied in psychophysics, producing an unfortunate gulf between the cognitive and psychophysical literature. (The basic problem still remains even with the crude discriminations studied in visual search tasks, prompting the development of methods that do not share this problem;

¹This work was supported by the Office of Naval Research (Contract N-0014-86-0269) and by the McDonnell Center for Higher Brain Function. Correspondence may be sent to the author at Washington University School of Medicine, Department of Neurology, 660 S. Euclid, Box 8111, St. Louis, MO 63110.

see Duncan, 1980; Eriksen & Spencer, 1969; Shiffrin & Gardner, 1972.)

In a modification of the above analysis, Treisman and Gormican (1988) have proposed that attention determines the spatial region over which evidence for the presence of a feature is pooled. With gross discriminations, pooling can occur over a wide area because distractors and targets are sufficiently separate that distractor noise will not activate target detectors. However, with fine discriminations, noise can more easily activate target detectors. To prevent false alarms, evidence from restricted regions must be considered serially.

Other explanations of the serial functions found under conditions involving fine discriminations are possible. Perhaps, for example, the grain or resolution of the representation that is computed is affected by attention. Additionally, it is misleading to talk about a single representation, since most perceptual properties are derived through a complex series of computations. Consider motion perception. Some models posit an initial filtering stage in which the input is analyzed by orientation, direction, and spatial frequency-specific mechanisms, followed by an integration stage in which motion along different contours is integrated to yield an object motion (Movshon, Adelson, Gizzi, & Newsome, 1985). Eye movements also need to be taken into account, as well as motion of surrounding objects and context (Johansson, 1975). The point is that the extraction of any perceptual property requires a series of operations that are not separated in the standard visual search task. The serial function found with difficult discriminations could be imposed at a variety of computations.

Finally, the difficulty may lie in the fact that subjects are required to make a direct report about the presence or absence of a target. As a number of researchers have pointed out, direct reports require a series of processes from image encoding through response selection (Allport, 1977; Duncan, 1980). Effects of attention can occur at any point in this sequence. In a visual search task, the target stimulus must first be analyzed and then passed, by selections mechanisms, from parallel to limited-capacity systems for decision and response. The performance increases caused by consistent mapping training (Shiffrin & Schneider, 1977), for example, may be due to the more efficient selection of stimulus representations rather than the initial analysis of the stimulus. The limitation implied by a serial function may therefore be imposed at the stage at which information is transferred to decision and response processes.

In summary, visual search studies do not force the conclusion that perceptual operations are automatic. Visual search techniques have been applied to a limited range of perceptual operations, and, within this range, serial functions are found if the discrimination is made difficult. These serial functions have multiple interpretations; the limitations they reflect could arise from a variety of mechanisms.

The problem of specifying the process on which attention acts is also apparent in other recent studies of the ef-

fect of attention on perception. For example, cuing has been shown to aid various types of feature detection (Downing, 1988; Prinzmetal, Presti, & Posner, 1986). The difficulty again is that standard cuing tasks do not distinguish different levels of perceptual representations and require direct reports. To the extent that a stimulus is more efficiently transferred—faster and/or with less noise added—when it is spatially attended, one will find beneficial effects of spatial cuing in virtually all tasks requiring direct report. An effect of cuing on vernier acuity, for example, would not imply that attention affects early visual mechanisms.²

One solution to this problem might involve making better use of the techniques developed by perception researchers to isolate perceptual processes for careful study. For example, a general strategy in perception has been to study various misperceptions or illusions that reflect the application, misapplication, or absence of specific computational routines. By embedding attentional manipulations (such as cuing or search) within these paradigms, it may be possible to localize the effects of attention to particular computations.

Gogel and Tietz (1976) asked observers to judge the direction of motion of a single dot, moving up and down, that could be stereoscopically presented at either a far or near distance (Figure 1). Two pairs of horizontally moving context dots were also displayed: one pair on a TV screen placed at the near distance, one on a second TV screen at the far distance. Each pair consisted of one dot whose vertical position corresponded to the "top" of the test motion, and one corresponding to the "bottom." The pair moved horizontally in a coupled fashion, as the test dot moved vertically between them. However, the relative phase of the near and far pair was reversed (one pair moved left, while the other was moving right, and vice versa). As a result, if each pair had been presented in isolation, the test dot would have been perceived to move obliquely in different directions. Gogel and Tietz determined the effect of a particular context pair by measuring the perceived direction of the test dot as a function of its depth separation from the pair and whether the pair was attended. Attention was manipulated through instruction. Subjects were told to attend to one pair and ignore the other. Gogel and Tietz found that a given context pair was more effective in producing the appropriate tilted motion path for the test dot if that pair was attended.

In a second study, Gogel and MacCracken (1979) presented a vertically moving test dot at different depths with respect to a single horizontally moving context pair. Subjects were told to judge the direction of motion of the test dot and to either attend or ignore the context dots. Gogel and MacCracken again found a greater effect of the context pair on test-dot direction (i.e., it was seen as more oblique) when it was attended.

The main question in these studies is not whether attention improves the accuracy of report (the dot is actually moving vertically), but how it affects an aspect of that report that can be related to a particular computation. Gogel and Tietz examined a paradigm that has been

Figure 1. Context dots are placed at the near and far depth with test dot at the far depth. (From "Adjacency and Attention as Determiners of Perceived Motion" by W. C. Gogel and J. Tietz, 1976, *Vision Research*, 16, p. 840. Copyright 1976 by Pergamon Press. Adapted by permission.)

well investigated in the perception literature and for which computational rationales have been developed. This paradigm was originated by Johansson (1975) in his studies of vector decomposition. As Rock (1983) has noted, this decomposition is functionally reasonable. If we are traveling on a train and waving goodbye with our hands, it is the relative motion of the hands, not the absolute motion, that is of significance. Waving goodbye carries the same meaning whether one is on a moving train or a stationary platform.

Gogel's work suggests that the observer has control over which framework is used in the vector decomposition. When two frameworks are available (Gogel & Tietz, 1976), the observer can choose which one to use. Similarly, the results of Gogel and MacCracken (1979) suggest that an observer can choose to judge the absolute motion of a stimulus rather than its relative motion. Selection apparently allows us to decompose the same display in different ways. By studying attention within the framework of an established perceptual task, Gogel was able to assign an attentional effect to a particular computational stage. This, in turn, allows one to rationalize why that stage is attentionally coupled.

Gogel and Tietz (1976) write, "The large effect of voluntary attention upon the direction of the apparent path of motion of the test point in this experiment provides the clearest visual example known to the authors of the effect of attention on perception" (p. 843). One might question several aspects of the methodology: attention was controlled through instruction, while several reviewers of this paper have objected that oculomotor variables, such as convergence and accommodation, were not adequately

controlled. However, the more important point is that this work apparently has not been considered part of the domain of studies relevant to the issue: How does attention affect stimulus processing? Although this study was published over 10 years ago, it has not generally been cited in reviews of the early-selection/late-selection controversy.

There is a diverse literature on the effects of attention on perception that to a greater or lesser degree embodies the approach suggested in this paper. This literature includes studies of size and shape perception (Epstein & Broota, 1986; Epstein & Lovitts, 1985; Gogel, 1967), brightness and achromatic color perception (Brussell, 1973; Brussell & Festinger, 1973; Coren, 1969; Festinger, Coren, & Rivers, 1970), motion perception (Dick, Ullman, & Sagi, 1987), illusory contours (Pritchard & Warm, 1983), perception of length (Coren, 1986; Coren & Girgus, 1972; Coren & Porac, 1983; Girgus & Coren, 1982; Jordan & Schiano, 1986; Pressey & Murray, 1976; Riddoch & Humphreys, 1983; Tsal, 1984), and perception of ambiguous figures (Hochberg & Peterson, 1987; Peterson, 1986; Tsal & Kolbet, 1985).

These studies vary in the success with which they isolate a particular computation, and some suffer from poor control over attention or oculomotor variables. These studies also do not necessarily bear directly on whether or not recognition and memory access occur in parallel. For example, Biederman's (1987) theory of object recognition implies that effects of attention on quantitative aspects of shape, such as length, are irrelevant to recognition. However, the more important point is to deemphasize the distinction between early and late selection.

Treisman's work on conjunctions (e.g., Treisman & Gelade, 1980), which sought to tie attentional effects to a particular computation has already shifted research away from the early/late dichotomy. However, Treisman's distinction between single features and conjunctions does not adequately summarize the range of computations involved in perception. Researchers should perhaps be skeptical of such dichotomies as early/late or single feature/conjunction.

It is increasingly clear that researchers in attention have only been partly successful in specifying the locus of attentional effects. One strategy for solving this problem is to embed attentional manipulations within perceptual paradigms that isolate particular computations.

REFERENCES

- ALLPORT, A. (1977). On knowing the meaning of words we are unable to report: The effects of visual matching. In S. Dornic (Ed.), *Attention and performance VI*. Hillsdale, NJ: Erlbaum.
- BIEDERMAN, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, **94**, 115-147.
- BROADBENT, D. (1982). Task combination and selective intake of information. *Acta Psychologica*, **50**, 253-290.
- BRUSSELL, E. (1973). Attention, brightness contrast, and assimilation: The influence of relative area. *Perception & Psychophysics*, **14**, 325-333.
- BRUSSELL, E., & FESTINGER, L. (1973). The Gelb effect: Brightness contrast plus attention. *American Journal of Psychology*, **86**, 225-235.
- BUNDESEN, C., & PEDERSEN, L. F. (1983). Color segregation and visual search. *Perception & Psychophysics*, **33**, 487-493.
- COREN, S. (1969). Brightness contrast as a function of figure-ground relations. *Journal of Experimental Psychology*, **80**, 517-524.
- COREN, S. (1986). An efferent component in the visual perception of direction and extent. *Psychological Review*, **93**, 391-410.
- COREN, S., & GIRGUS, J. S. (1972). Differentiation and decrement in the Mueller-Lyer illusion. *Perception & Psychophysics*, **12**, 466-470.
- COREN, S., & PORAC, C. (1983). The creation and reversal of the Mueller-Lyer illusion through attentional manipulation. *Perception*, **12**, 49-54.
- DEUTSCH, J., & DEUTSCH, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, **70**, 80-90.
- DICK, M., ULLMAN, S., & SAGI, D. (1987). Parallel and serial processes in motion detection. *Science*, **237**, 400-402.
- DOWNING, C. (1988). Expectancy and visual-spatial attention: Effects on perceptual quality. *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 188-202.
- DUNCAN, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review*, **87**, 272-300.
- DUNCAN, J. (1983). Category effects in visual search: A failure to replicate the "oh-zero" phenomenon. *Perception & Psychophysics*, **34**, 221-232.
- DUNCAN, J., & HUMPHREYS, G. (1989). Visual search and stimulus similarity. *Psychological Review*, **96**, 433-458.
- EGETH, H., JONIDES, J., & WALL, S. (1972). Parallel processing of multi-element displays. *Cognitive Psychology*, **3**, 674-698.
- EPSTEIN, W., & BROOTA, K. D. (1986). Automatic and attentional components in perception of size-at-a-distance. *Perception & Psychophysics*, **40**, 256-262.
- EPSTEIN, W., & LEVITTS, B. (1985). Automatic and attention components in perception of shape-at-a-slant. *Journal of Experimental Psychology: Human Perception & Performance*, **11**, 355-366.
- ERIKSEN, C., & SPENCER, T. (1969). Rate of information processing in visual perception: Some results and methodological considerations. *Journal of Experimental Psychology Monographs*, **79**(2, Pt. 2).
- FESTINGER, L., COREN, S., & RIVERS, G. (1970). The effect of attention on brightness contrast and assimilation. *American Journal of Psychology*, **83**, 189-207.
- GARDNER, G. (1973). Evidence for independent parallel channels in tachistoscopic perception. *Cognitive Psychology*, **4**, 130-155.
- GIRGUS, J. S., & COREN, S. (1982). Assimilation and contrast illusions: Differences in plasticity. *Perception & Psychophysics*, **32**, 555-561.
- GOGEL, W. C. (1967). Cue enhancement as a function of task set. *Perception & Psychophysics*, **2**, 455-458.
- GOGEL, W., & MACCRACKEN, P. (1979). Depth adjacency and induced motion. *Perception & Motor Skills*, **48**, 343-350.
- GOGEL, W., & TIETZ, J. (1976). Adjacency and attention as determiners of perceived motion. *Vision Research*, **16**, 839-845.
- HOCHBERG, J., & PETERSON, M. (1987). Piecemeal organization and cognitive components in object perception: Perceptually coupled responses to moving objects. *Journal of Experimental Psychology: General*, **116**, 370-380.
- JOHANSSON, G. (1975). Visual motion perception. In R. Held & W. Richards (Eds.), *Recent progress in perception* (pp. 67-75). San Francisco: W. H. Freeman.
- JORDAN, K., & SCHIANO, D. J. (1986). Serial processing and the parallel-lines illusion: Length contrast through relative spatial separation of contours. *Perception & Psychophysics*, **40**, 384-391.
- KAHNEMAN, D. (1973). *Attention and effort*. Englewood, NJ: Prentice-Hall.
- KEELE, S. (1973). *Attention and human performance*. Pacific Palisades, CA: Goodyear.
- MOVSHON, J., ADELSON, E., GIZZI, M., & NEWSOME, W. (1985). The analysis of moving visual patterns. In C. Chagas, R. Gattass, & C. Gross (Eds.), *Pattern recognition mechanisms* (pp. 117-151). New York: Springer-Verlag.
- MÜLLER, H. J., & FINDLAY, J. M. (1987). Sensitivity and criterion effects in the spatial cuing of visual attention. *Perception & Psychophysics*, **42**, 383-399.
- NAKAYAMA, K., & SILVERMAN, G. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, **320**, 264-265.
- NEISSER, U. (1963). Decision time without reaction time: Experiments in visual scanning. *American Journal of Psychology*, **76**, 376-385.
- PETERSON, M. (1986). Illusory concomitant motion in ambiguous stereograms: Evidence for nonstimulus contributions to perceptual organization. *Journal of Experimental Psychology: Human Perception & Performance*, **12**, 50-60.
- PRESSEY, A. W., & MURRAY, R. (1976). Further developments in the assimilation theory of geometric illusions: The adjacency principle. *Perception & Psychophysics*, **19**, 536-544.
- PRINZMETAL, W., PRESTI, D., & POSNER, M. I. (1986). Does attention affect visual feature integration. *Journal of Experimental Psychology: Human Perception & Performance*, **12**, 361-370.
- PRITCHARD, W., & WARM, J. (1983). Attentional processing and the subjective contour illusion. *Journal of Experimental Psychology: General*, **112**, 145-175.
- RIDDOCH, J., & HUMPHREYS, G. (1983). The effect of cueing on unilateral neglect. *Neuropsychologia*, **21**, 589-599.
- ROCK, I. (1983). *The logic of perception*. Cambridge, MA: MIT Press.
- SCHNEIDER, W., & FISK, A. (1984). Automatic category search and its transfer. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 1-15.
- SHIFFRIN, R. M., & GARDNER, G. T. (1972). Visual processing capacity and attentional control. *Journal of Experimental Psychology*, **93**, 72-83.
- SHIFFRIN, R., & SCHNEIDER, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, **84**, 127-190.
- SHULMAN, G. L., & POSNER, M. I. (1988). *Relating sensitivity and criterion effects to the internal mechanisms of visual spatial attention* (Tech. Rep. No. 88-2). Washington, DC: Office of Naval Research.
- TOWNSEND, J. T. (1972). Some results on the identifiability of parallel and serial processes. *British Journal of Mathematical & Statistical Psychology*, **25**, 168-199.
- TREISMAN, A. M., & 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.

- TREISMAN, A., & SOUTHER, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, **113**, 285-310.
- TSAL, Y. (1984). A Mueller-Lyer illusion induced by selective attention. *Quarterly Journal of Experimental Psychology*, **36A**, 319-333.
- TSAL, Y., & KOLBET, L. (1985). Disambiguating ambiguous figures by selective attention. *Quarterly Journal of Experimental Psychology*, **37A**, 25-37.
- VAN DER HEIJDEN, A. (1981). *Short-term visual information forgetting*. London: Routledge and Kegan Paul.
- WOLFE, J. M., & FRANZEL, S. L. (1988). Binocularity and visual search. *Perception & Psychophysics*, **44**, 81-93.

NOTES

1. Van der Heijden (1981) has noted that this controversy has collapsed two different issues: (1) what features can be analyzed in an initial stage of parallel analysis and (2) what features are used to select or transfer information to later capacity-limited stages. Although theorists initially assumed that the answers to both questions were identical, recent theories have proposed that only certain features analyzed in the parallel stage can be used for efficient selection. Physical fea-

tures may be more efficient for purposes of selection, even though both physical and semantic properties can be analyzed in parallel. The question of which stimulus attributes are analyzed in parallel is therefore partly separate from the question of which stimulus attributes are used as a basis for selection.

2. Some researchers have attempted to specify which processes are affected by attention by measuring whether d' or beta are affected in different tasks by attentional manipulations (Müller & Findlay, 1987). Attentional effects on d' are assumed to be "perceptual" in nature, whereas effects on beta reflect later decision processes. This dichotomy is not very helpful. A large number of mechanisms can produce changes in d' . Also, the assumed relationship between signal detection parameters and internal mechanisms may be more complex than the usual dichotomy implies. Although changes in beta are often ascribed to a single detection mechanism, the evidence for this is not compelling (Shulman & Posner, 1988). As Kahneman (1973) has noted, criterion shifts can have important perceptual consequences.

(Manuscript received September 16, 1988;
revision accepted for publication September 8, 1989.)