

Visual Search and Stimulus Similarity

John Duncan
MRC Applied Psychology Unit
Cambridge, England

Glyn W. Humphreys
Birkbeck College, University of London
London, England

A new theory of search and visual attention is presented. Results support neither a distinction between serial and parallel search nor between search for features and conjunctions. For all search materials, instead, difficulty increases with *increased* similarity of targets to nontargets and *decreased* similarity between nontargets, producing a continuum of search efficiency. A parallel stage of perceptual grouping and description is followed by competitive interaction between inputs, guiding selective access to awareness and action. An input gains weight to the extent that it matches an internal description of that information needed in current behavior (hence the effect of target-nontarget similarity). Perceptual grouping encourages input weights to change together (allowing "spreading suppression" of similar nontargets). The theory accounts for harmful effects of nontargets resembling *any possible* target, the importance of local nontarget grouping, and many other findings.

The Efficiency of Visual Selection

It is common knowledge that we can pay attention (at any one time) to only a small amount of the information present in a visual scene. Experimentally, it is easy to confirm that people can take up and report only a small amount of the information contained in a brief visual display (Helmholtz, cited in Warren & Warren, 1968). Such a limitation imposes a strong requirement for selection: Ideally, we should confine attention to that information needed to guide current behavior, and again it is easy to confirm that people can use many different selection criteria (location, color, movement, etc.) to choose which information to see in a briefly glimpsed scene (e.g., Helmholtz, cited in Warren & Warren, 1968; von Wright, 1970). This article deals with the efficiency of selection.

In visual search experiments, subjects are asked to detect particular target stimuli presented among irrelevant nontargets. Results depend on the combination of targets and nontargets used. With some combinations, the number of nontargets in a display has little if any effect. Obviously, they are rejected without access to those rate-limiting stages of processing responsible for our limited ability to pay attention to several stimuli at once. The experience is that attention is drawn directly to the target, implying an efficient prior rejection of nontargets (Duncan,

1980b, 1985; Hoffman, 1978; Shiffrin & Schneider, 1977). In other cases, increasing the number of nontargets substantially increases the time taken to find the target. The experience is that we must pay attention to several nontargets in turn before the target is "found," implying that the efficiency of nontarget rejection is reduced. Here, we seek to understand selection in general by investigating boundary conditions on efficient nontarget rejection in visual search.

Feature Integration Theory

Our point of departure is Treisman's feature integration theory (Treisman & Gelade, 1980; Treisman & Souther, 1985). According to this theory, input from a visual display is processed in two successive stages. The first stage consists of a set of spatiotopically organized "maps" of the visual field, each coding the presence of a particular, elementary stimulus attribute or "feature." Thus, one map might code where redness occurs, one where 45°-tilted lines occur, and so on. This stage works in parallel across the visual field but is limited in that it produces no useful information about the *conjunction* of elementary features. Thus, activity in separate maps might show that the field contains redness, greenness, a diagonal line, and a closed loop, but it cannot show that the line is red and the loop green. Useful conjunction information only becomes available with processing at the second stage. Attention is focused on a particular area of the field. Outputs from those maps with activity in this particular area are then combined to produce the percept of a whole object (e.g., a green loop). If features are to be accurately conjoined, attention must be focused serially on one object after another. It is this serial process that is responsible for our limited ability to see a whole scene at a glance.

The difficulty of visual search is thus determined by whether a target is unique in some elementary feature or only in its conjunction of features. As an example of feature search, the target might be a blue shape presented among a mixture of reds and

Financial support was provided by the MRC Applied Psychology Unit, where the work was carried out. The second author's research is also supported by grants from the Medical Research Council, Economic and Social Research Council, and Science and Engineering Research Council.

We are grateful to Claus Bundesen, Howard Egeth, Harold Pashler, Richard Shiffrin, and Anne Treisman, all of whom commented extensively on earlier drafts.

Correspondence concerning this article should be addressed to John Duncan, MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, England.

greens. Net activity in the blueness map is sufficient to show whether a target is present, and because activity in this map develops in parallel across the visual field, there should be little effect of the number of items present. In conjunction search, on the other hand, the target might be a red O presented among mixed blue Os and red Xs. Because display items can only be classified as targets or nontargets with focused attention, the target must be found by scanning serially through the display, and the number of items will have a large effect. Although later we consider some exceptions, results supporting this distinction between feature and conjunction search have now been reported many times, using color, form, size and other stimulus attributes (e.g., Treisman, 1982; Treisman & Gelade, 1980; Treisman, Sykes, & Gelade, 1977).

Recently, the theory has been modified to allow the possibility that feature search can be serial when targets and nontargets are closely similar (Treisman & Gormican, 1988). Suppose that targets and nontargets differ slightly in color. Then, each nontarget might have some tendency to excite the target map, making it hard to decide whether this map contains enough net activity to indicate that a target is really present. The more nontargets are present, furthermore, the smaller will be the proportional increase in activity produced by a target, and the harder will be the decision. When this happens, Treisman and Gormican (1988) suggest that attention is focused serially on one clump of items after another. The size of the clump is chosen such that, within one clump, net activity in the target map will reliably indicate whether a target is present. The more discriminable the targets and nontargets, the larger can be the clumps. The original version of the theory then emerges as a special case. With high enough discriminability, the whole display can be treated as a single clump.

Feature integration theory is consistent with a range of psychological phenomena beyond visual search (Treisman & Gelade, 1980; Treisman & Schmidt, 1982; Treisman & Souther, 1985). It is supported by physiological evidence for early analysis of different stimulus attributes in different brain areas (Maunsell & Newsome, 1987). Results of work with connectionist models of vision also suggest that serial processing may be a good solution to the problem of correctly integrating an object's different attributes (Feldman, 1985).

Overview

Like feature integration theory, the present work deals with how search efficiency is determined by the nature of relevant (target) and irrelevant (nontarget) stimulus materials. Although other variables such as practice are important in search (Schneider & Shiffrin, 1977), stimulus factors are our main concern here.

We begin with an assessment of feature integration theory—in particular, its account of letter search. A series of four experiments shows very large variations in search efficiency across stimulus materials, variations that are inconsistent with feature integration theory *whatever* the postulate concerning elementary features of simple shapes. We then present a new account, different from feature integration theory in several important respects. First, the dichotomy between serial and parallel search

has no real place in our account, which is based on a continuum of search efficiency. Second, our approach is based not on a distinction between different stimulus attributes, but more abstractly on stimulus relations (similarities) that in principle can be specified for any attribute. Thus, we argue that very similar stimulus principles control search difficulty whatever the search materials, from simple color patches to complex feature conjunctions. In particular, search efficiency decreases with (a) *increasing* similarity between targets and nontargets (which we call *T-N similarity*), and (b) *decreasing* similarity between nontargets themselves (*N-N similarity*), the two interacting to scale one another's effects. We try to show that these principles are consistent both with the body of the search literature and with the apparent contrast between feature and conjunction search itself. We go on to develop a theory of how, in search and other tasks, attention is directed to behaviorally relevant information in the visual field. This theory deals with similarities between possible targets and nontargets in search, with local effects of similarity within a display, and with a variety of other findings holding across a range of different search materials.

Feature Integration Theory and Letter Search

Prior Evidence

Feature integration theory has been applied to letter search by considering the conjunction of a shape's parts. There have been many visual search experiments using simple shapes such as letters and digits. In some tasks there is very little effect of the number of nontargets in a display—for example, search for a C among 4s (Egeth, Jonides, & Wall, 1972) or for a T or F among Os (Shiffrin & Gardner, 1972)—whereas in other tasks the effect is substantial (Kleiss & Lane, 1986). Does the difference depend on whether targets possess some unique (shape) feature? Despite some positive findings, the literature as a whole is rather puzzling.

In fact, feature integration theory has been applied to letter search in two ways. The first, called by Duncan (1987) the case of *within-object conjunctions*, deals with the spatial arrangement of strokes within a letter. According to several accounts, the elementary features of letters include lines of particular length and orientation, intersections (line crossings), line terminators, and a few other features (e.g., Bergen & Julesz, 1983; Treisman & Paterson, 1984; Treisman & Souther, 1985). Suppose then that two letters share exactly the same features, differing only in their spatial arrangement. Obvious candidates are pairs like L and T, which contain different arrangements of the same strokes. Individual feature maps will not be able to separate these letters; only when the outputs of different maps are put together with serial attention will the distinction be made. Correspondingly, Beck and Ambler (1973) reported a large effect of display size in search for an L among nontarget Ts, contrasting with a much smaller effect when the target (a tilted T) had strokes of a unique orientation. Similar results were reported by Bergen and Julesz (1983), contrasting search for a T or a + (with its unique intersection) among nontarget Ls.

The second case concerns *across-object conjunctions*. Here,

the target can be formed by recombining strokes from different nontargets (e.g., search for R among Ps and Qs); again, the target is unique only in its conjunction of strokes. Treisman and Gelade (1980) and Duncan (1979) found large effects of display size in such tasks. If the target had a unique stroke (e.g., R among Ps and Bs), on the other hand, the effect of display size was rather smaller.

Of course, the interpretation of such results in terms of feature integration theory depends on assumptions concerning what elementary letter features are coded at the first, parallel processing stage. To deal with within-object conjunctions, for example, the theory must assume that the position of strokes within a letter is not coded. We shall consider such issues later. For the moment, we may refer to these stimuli as stroke conjunctions rather than feature conjunctions.

Other results complicate the picture. Consider first the case of within-object conjunctions. Humphreys, Riddoch, and Quinlan (1985) studied search for an inverted T among upright Ts. Despite the resemblance of this to the within-object conjunction tasks of Beck and Ambler (1973) and Bergen and Julesz (1983), there was little effect of display size, search times increasing by only 3-ms/item when arrays had a regular spatial arrangement. What can we say about these apparently conflicting results? A first point to note is the difficulty of comparing effects across experiments. For various reasons, even unlimited-capacity parallel models predict some drop in performance with increasing display size (Duncan, 1980a; Eriksen & Spencer, 1969). In reaction time (RT) studies, effects up to 5- or 6-ms/item are comparable with those usually given by feature search (Treisman & Souther, 1985), at least when the target is present. Effects as great as 20- to 30-ms/item are typical of conjunction search. Beck and Ambler (1973) and Bergen and Julesz (1983), however, measured accuracy rather than RT in studies with limited exposure duration. Little is known about the scale of display size effects in such experiments. Second, Humphreys, Quinlan, and Riddoch (in press) showed that a crucial variable in these studies is letter size or, more accurately, the ratio of size to retinal eccentricity. Using the same task as before, they obtained display size effects of 14- and 2-ms/item, respectively, for size/eccentricity ratios of 1/6 and 1/3. These ratios may be compared with about 1/8 for Beck and Ambler (1973) and up to 1/9 for Bergen and Julesz (1983). It seems likely that the results of these authors' investigations were in part dependent on their use of relatively small letters.

Questions may also be raised over Treisman and Gelade's (1980) study of across-object conjunctions. The smallest effects of display size (4- and 7-ms/item, respectively, for target-present and target-absent displays) were obtained when nontargets were homogeneous (e.g., search for R among Ps). With heterogeneous nontargets, the effect was always much bigger, whether the target had a unique stroke (12- and 37-ms/item) or not (23- and 46-ms/item).¹ A study of similar tasks by Kleiss and Lane (1986) is also instructive. Only heterogeneous nontargets were used. Following Shiffrin and Gardner (1972), Kleiss and Lane (1986) measured the accuracy of target detection in displays of constant size, presented either all at once or two at a time. The technique is useful because unlimited-capacity parallel models predict no effect of presentation mode. In fact, there was a large

advantage for presentation two at a time in both feature and conjunction tasks.

Altogether, then, there are several aspects of letter search data that feature integration theory does not explain. One important variable is letter size. With large enough letters, there can be little effect of display size even when the target is unique only in its within-object conjunction of strokes. A second important variable is nontarget homogeneity. With heterogeneous nontargets, there can be large effects of display size—and large departures from unlimited-capacity parallel search—even if the target has a unique stroke. The four experiments that follow develop these puzzles for feature integration theory, and taken together, they show that the theory cannot explain the large variations in search efficiency seen across different letter search tasks.

Experiment 1

Experiment 1 was designed to investigate effects of letter size and nontarget homogeneity on both feature and conjunction search. Using an RT task, we replicated Beck and Ambler's (1973) comparison between search for Ls and tilted Ts among nontarget Ts that were either upright or rotated 90° clockwise. We used two extremes of letter size (size/eccentricity ratios of 1/12 and 1/3) and nontargets that were either homogeneous (upright in one block of trials, sideways in another) or heterogeneous (both upright and sideways mixed in each display).

Method

Tasks. Experiment 1 was run on-line on a Cambridge Electronic Design laboratory computer system, controlling a Hewlett-Packard X-Y display (1332A) with P24 phosphor. Displays were viewed from a chin rest, at a distance of about 65 cm. On each trial, the subject fixated a dot in the center of the screen, pressed a foot switch, and saw an immediate 180-ms display of 2, 4, or 6 letters. The response was to be made, as quickly as possible, by pressing a key with the right hand if a specified target letter was present or with the left hand if it was absent. An interval of 1,000 ms preceded onset of the fixation point for the next trial.

With the modification noted later, letters appeared on the perimeter of an imaginary circle of radius 2° 24', centered on fixation. Starting at 12 o'clock, there were eight possible letter positions, evenly spaced round the circle. A (randomly selected) arc of adjacent positions was used for each display, equating the distance between adjacent characters across display sizes.

Three factors varied between blocks. The target was either an upright L or a T tilted 45° clockwise. Nontargets were either homogeneous—in which case they were either all upright Ts or all Ts rotated 90° clockwise—or heterogeneous—in which case each display contained, as nearly as possible, an equal number of Ts in these two orientations, randomly arranged. The two strokes of each letter were equal in length. They measured either 12' arc or 48' arc, giving letter size/eccentricity ratios of 1/12 or 1/3.

A possible difficulty with regular nontarget displays is that supraletter cues might show whether a target is present. Consider the case of search for an L among Ts. Suppose, for example, that the horizontal lines of

¹ Values have been estimated from Treisman and Gelade's (1980) Figure 6. Estimates are based only on the comparison of display sizes 1 and 15, the values available for all conditions.

Table 1
Experiment 1: Reaction Times (in Milliseconds) as a Function of Display Size

Condition	Small letters (1/12)				Large letters (1/3)			
	2	4	6	Slope (ms/item)	2	4	6	Slope (ms/item)
Target = L								
Homogeneous								
Present	432	432	444	3	388	383	402	4
Absent	448	460	464	4	413	413	410	-1
Heterogeneous								
Present	420	416	430	3	383	395	397	4
Absent	448	452	475	7	414	425	416	1
Target = tilted T								
Homogeneous								
Present	446	461	470	6	400	392	398	-1
Absent	470	478	490	5	404	409	404	0
Heterogeneous								
Present	492	505	523	8	424	427	445	5
Absent	530	530	547	4	444	464	450	2

all letters could be grouped together, and the shape of the resulting group could be determined. In a homogeneous display with letters arranged around the perimeter of a circle, this shape would be a smooth arc when the target was absent, but distorted when the target was present. To eliminate such cues, each nontarget was shifted slightly so that one of its strokes, horizontal or vertical (randomly selected), fell in the position that the corresponding stroke of an L would occupy in the same display location. The result was a display of rather irregular appearance, in which, when the target was an L, any target stroke fell in a position possible for a nontarget stroke.

Design. Each subject served in six sessions of about 1 hr each, on different days. Type of nontargets, homogeneous or heterogeneous, was fixed for any one session and alternated between sessions, with the order counterbalanced across subjects.

Each session was divided into four blocks, one for each combination of letter size and target. There were always two blocks at one letter size followed by two at the other, with the same order of targets in each pair. With these constraints, the order of blocks was counterbalanced across subjects, although fixed for any one. Each block was further divided into two sub-blocks, each of 24 practice followed by 96 experimental trials. When nontargets were heterogeneous, the two sub-blocks were identical, but when nontargets were homogeneous, one sub-block was devoted to the upright and one to the sideways T.

Within each experimental run of 96 trials, there were equal numbers of trials with and without a target at each display size. Otherwise, the order of trials was random, as was the arc of letter positions chosen for each display and the position of the target (if present) in this arc. At the end of every 24 practice and 96 experimental trials, the subject was shown mean reaction time and error rate for the run.

Subjects. All of the experiments in this series used subjects from the paid panel of the Applied Psychology Unit. Here, they were 4 women, between 28 and 35 years of age.

Results and Discussion

Table 1 shows mean RTs in each condition, as well as slopes of best-fitting linear functions relating RT to display size. Data are from experimental trials on the last 2 days of practice. Trials with RTs greater than 1,500 ms have been excluded.

There were four important results. First, slopes were all in the range normally taken to suggest parallel search (Treisman &

Souther, 1985), with a maximum of 6-ms/item (averaged across present and absent trials). Second, slopes were very similar for the two targets, unlike the results of Beck and Ambler (1973). Third, slopes were slightly greater with small than with large letters, although the effect was much smaller than the one reported by Humphreys et al. (in press). Fourth, slopes were little affected by nontarget homogeneity, although homogeneous nontargets gave slightly quicker responding overall.

Analysis of variance (ANOVA) confirmed these conclusions. There was a significant main effect of display size, $F(2, 6) = 27.8$, $p < .001$, which interacted with letter size, $F(2, 6) = 9.5$, $p < .02$, but not target type, $F(2, 6) = 0.2$, or nontarget homogeneity, $F(2, 6) = 2.4$. There was also a significant but small three-way interaction between display size, letter size, and target presence, $F(2, 6) = 7.0$, $p < .05$, which we neglect. Finally, there were significant main effects of nontarget homogeneity, $F(1, 3) = 10.9$, $p < .05$, and letter size, $F(1, 3) = 91.5$, $p < .005$, and a four-way interaction, probably spurious, between nontarget

Table 2
Experiment 1: Error Proportions

Condition	Small letters (1/12)			Large letters (1/3)		
	2	4	6	2	4	6
Target = L						
Homogeneous						
Present	.023	.047	.086	.008	.031	.078
Absent	.031	.016	.039	.008	.008	.016
Heterogeneous						
Present	.031	.047	.094	.024	.000	.024
Absent	.024	.024	.016	.039	.008	.016
Target = tilted T						
Homogeneous						
Present	.039	.047	.094	.023	.023	.023
Absent	.016	.024	.055	.016	.024	.000
Heterogeneous						
Present	.039	.125	.149	.024	.024	.031
Absent	.031	.032	.063	.016	.047	.000

Table 3
Experiment 2a: Reaction Times (in Milliseconds) as a Function of Display Size

Condition/session	Small letters (1/6)				Large letters (1/3)			
	2	4	6	Slope (ms/item)	2	4	6	Slope (ms/item)
Target = L								
Present								
1	523	565	556	8	499	468	521	6
2	488	509	519	8	485	483	513	7
3	487	505	513	7	491	499	502	3
<i>M</i>	500	526	529	7	492	483	512	5
Absent								
1	537	540	560	6	483	478	490	2
2	515	513	514	0	477	484	491	4
3	496	497	497	0	489	476	485	-1
<i>M</i>	516	517	524	2	483	479	489	2
Target = tilted T								
Present								
1	500	531	539	10	458	455	463	1
2	504	508	525	5	457	464	461	1
3	470	491	497	7	446	453	463	4
<i>M</i>	491	510	520	7	454	457	462	2
Absent								
1	495	499	511	4	473	463	477	1
2	511	507	519	2	465	474	458	-2
3	476	488	477	0	467	464	443	-6
<i>M</i>	494	498	502	2	468	467	459	-2

homogeneity, target type, letter size, and target presence, $F(1, 3) = 26.0$, $p < .02$, for which MS_e was more than 10 times smaller than any other in the analysis.

Error data appear in Table 2. They suggest only one modification to our conclusions. When letters were small, error rates increased with increasing display size. This suggests that RT results may underestimate the true interaction between display size and letter size.

Experiment 2

Experiment 1 left us with two questions. First, we confirmed the finding of Humphreys et al. (in press) that smaller letters produce a greater effect of display size. Perhaps because of the brief exposure of stimulus displays, however, the result was reflected partly in error rates rather than RTs. Experiment 2 examined the effect further, using displays that remained visible until the response. This also allowed an examination of RT effects early in practice—data we have not presented for Experiment 1 because high error rates with small letters made RT data uninterpretable.

Experiment 1 also showed no effect of target type, in disagreement with the results of Beck and Ambler (1973). Instead, the effect of display size for both types of target depended on letter size. These results are also reexamined in Experiment 2.

Method

Tasks. There were the following changes from Experiment 1. Nontargets were always homogeneous upright Ts. Letter sizes were 12' and 24'; eccentricity (with the same small jitter as before) was 1°12' in Experiment 2a, giving size/eccentricity ratios of 1/6 and 1/3, and 2°24' in Experiment 2b, giving ratios of 1/12 and 1/6. Displays remained visible until the response. A right-hand response was required when all display items were the same (target absent), a left-hand response when one item was different (target present).

Design. Each subject served in three similar sessions. The four blocks of each session, one for each combination of letter size and target, were counterbalanced as before. Each block had a run of 24 practice trials and then two runs each of 72 experimental trials.

Subjects. Each of Experiments 2a and 2b had 4 subjects, between 19 and 32 years of age. There were 7 women and 1 man.

Results

Reaction times from Experiment 2a are shown in Table 3. The table shows separate results for the three sessions of practice, as well as the mean across sessions. As before, all of the slopes were in the range normally taken to suggest parallel search. An ANOVA showed a significant effect of display size, $F(2, 6) = 10.4$, $p < .02$, which interacted with session, $F(4, 12) = 3.4$, $p < .05$, but not with target type, $F(2, 6) = 1.4$, or letter size, $F(2, 6) = 3.4$. The only other significant effects were

Table 4
Experiment 2b: Reaction Times (in Milliseconds) as a Function of Display Size

Condition/session	Small letters (1/12)				Large letters (1/6)			
	2	4	6	Slope (ms/item)	2	4	6	Slope (ms/item)
Target = L								
Present								
1	476	528	527	13	427	458	473	12
2	437	459	482	11	388	407	430	11
3	399	420	431	8	382	393	409	7
<i>M</i>	437	469	480	11	399	419	437	10
Absent								
1	539	594	687	37	473	514	525	13
2	496	528	565	17	447	464	478	8
3	480	507	521	10	452	457	441	-3
<i>M</i>	505	543	591	22	457	478	481	6
Target = tilted T								
Present								
1	478	497	525	12	412	431	443	8
2	431	447	451	5	393	372	394	0
3	417	435	426	2	374	377	380	2
<i>M</i>	442	460	467	6	393	393	405	3
Absent								
1	546	583	636	23	479	485	499	5
2	477	510	534	14	443	448	442	0
3	467	481	490	6	413	426	412	0
<i>M</i>	497	525	553	14	445	453	451	2

letter size, $F(1, 3) = 25.7$, $p < .02$, and a three-way interaction between target type, letter size, and target presence, $F(1, 3) = 11.6$, $p < .05$, which we shall neglect.

Reaction times from Experiment 2b are shown in Table 4. This time slopes showed a large effect of letter size, especially early in practice. For the first time, as well, slopes were lower when the target was a tilted T than when it was an L—although both showed an effect of letter size. An ANOVA showed a significant effect of display size, $F(2, 6) = 11.8$, $p < .01$, which interacted significantly with letter size, $F(2, 6) = 10.2$, $p < .02$, with target type, $F(2, 6) = 27.2$, $p = .001$, and with session, $F(4, 12) = 4.6$, $p < .02$. The only other significant effects were letter size, $F(1, 3) = 12.8$, $p < .05$, and session, $F(2, 6) = 10.6$, $p < .02$.

Error data are shown in Tables 5 and 6. Values are means across sessions. The data suggest no reason to alter our conclusions. Even when the size/eccentricity ratio was 1/12, there was no great increase in error rate with increasing display size.

Discussion

The major point made by Experiments 1 and 2 is that although slopes may sometimes differ in feature (tilted T) and conjunction (L) search, the overall pattern of results is the same. In both cases, slopes are close to zero when the ratio of letter

size to eccentricity is large (1/3). As the ratio decreases, so slopes increase, especially across the range 1/6 to 1/12.

Note that Humphreys et al. (in press), whose task was search for an inverted T among upright Ts, found a large increase in slopes across the range of size/eccentricity ratios 1/3 to 1/6. A plausible suggestion is that letter search slopes generally increase with decreasing size/eccentricity ratio, but that the exact shape of the function depends on the particular targets and non-targets used.

Obviously, these results do not suggest that search for Ls was serial, whereas search for tilted Ts was parallel. More generally, many results show that in both visual search and related tasks,

Table 5
Experiment 2a: Error Proportions

Condition	Small letters (1/6)			Large letters (1/3)		
	2	4	6	2	4	6
Target = L						
Present	.031	.035	.038	.031	.042	.031
Absent	.031	.038	.042	.045	.031	.014
Target = tilted T						
Present	.021	.014	.045	.045	.004	.024
Absent	.049	.028	.038	.049	.021	.011

Table 6
Experiment 2b: Error Proportions

Condition	Small letters (1/12)			Large letters (1/6)		
	2	4	6	2	4	6
Target = L						
Present	.028	.028	.052	.007	.034	.042
Absent	.042	.045	.049	.059	.024	.045
Target = tilted T						
Present	.024	.032	.042	.021	.021	.038
Absent	.066	.056	.049	.035	.056	.018

the effect of adding irrelevant (nontarget) information to a display varies *continuously*, from little or none (or even a beneficial effect; see Polich, 1986; Sagi & Julesz, 1987), to at least 100-ms/item (e.g., Bundesen, Shibuya, & Larsen, 1985; Duncan, 1983, 1987; Quinlan & Humphreys, 1987; Treisman & Gelade, 1980; Treisman & Souther, 1985). Variations in slope across letter size, practice, and task in Experiments 1 and 2 make the same point. Of course, such continuous variations in search efficiency are consistent with a variety of theoretical positions, including some based on varying mixtures of serial and parallel processing (Fisher, Duffy, Young, & Pollatsek, 1988; Schneider & Shiffrin, 1977; Treisman & Gormican, 1988). We may note these, nevertheless, because the account we present is based on continuous variables.

How might feature integration theory account for the finding that, if letters are sufficiently large, search time is almost independent of display size even though targets and nontargets differ only in their conjunction of strokes? The obvious possibility is that some elementary feature other than (or in addition to) the stroke is detected at the first, parallel stage of processing. Of course, an L has attributes not possessed by a T—a particular sort of junction, for one. In Experiment 3, we used nontargets that were identical to the target across a 90° rotation.

Experiment 3

Method

Tasks. The target was an upright L, and nontargets were Ls rotated 90° clockwise or counterclockwise from the target position. Nontarget homogeneity varied between blocks, as in Experiment 1. Letter size was

24' or 48'. Eccentricity (with the usual modification) was 2°24', giving size/eccentricity ratios of 1/6 and 1/3. Displays remained visible until the response. The right hand was used to indicate target present, the left to indicate absent.

Design. Each subject served in three similar sessions. The four blocks per session, one for each combination of letter size and nontarget homogeneity, were counterbalanced as before. Each block was divided into two sub-blocks, each of 24 practice and 96 experimental trials. When nontargets were homogeneous, one sub-block was devoted to each nontarget orientation.

Subjects. The 4 subjects, 2 of each sex, were between 20 and 26 years of age.

Results

Mean RTs for the third day of practice are shown in Table 7. The results were extremely striking. When nontargets were homogeneous, the effect of display size was comparatively small. At least with the larger letters, slopes were in the range normally taken to indicate parallel search. Search through heterogeneous nontargets, however, was extraordinarily difficult. There were very large effects of display size, especially when the target was absent. For the first time, these results with heterogeneous nontargets resembled those reported by Treisman and Gelade (1980) for conjunction search. Sample displays are shown in Figure 1.

An ANOVA showed a significant effect of display size, $F(2, 6) = 25.3, p < .005$, which interacted with nontarget homogeneity, $F(2, 6) = 16.3, p < .005$. There were also significant mean effects of nontarget homogeneity, $F(1, 3) = 23.4, p < .02$, and target presence, $F(1, 3) = 25.8, p < .02$, and a significant interaction between them, $F(1, 3) = 48.1, p < .01$. The last significant effect was the four-way interaction between letter size, nontarget homogeneity, target presence, and display size, $F(2, 6) = 23.5, p < .005$, reflecting the fact that the largest interaction between display size and target presence occurred with large letters and heterogeneous nontargets.

Error data are shown in Table 8. They suggest no change in conclusions.

Discussion

Experiment 3 adds to the difficulties of feature integration theory by showing that search time can be almost independent of display size even when homogeneous nontargets are rotations

Table 7
Experiment 3: Reaction Times (in Milliseconds) as a Function of Display Size

Condition	Small letters (1/6)				Large letters (1/3)			
	2	4	6	Slope (ms/item)	2	4	6	Slope (ms/item)
Homogeneous								
Present	457	488	511	14	445	479	478	8
Absent	479	502	525	12	470	473	474	1
Heterogeneous								
Present	498	568	678	45	480	561	632	38
Absent	601	711	824	56	564	716	846	71

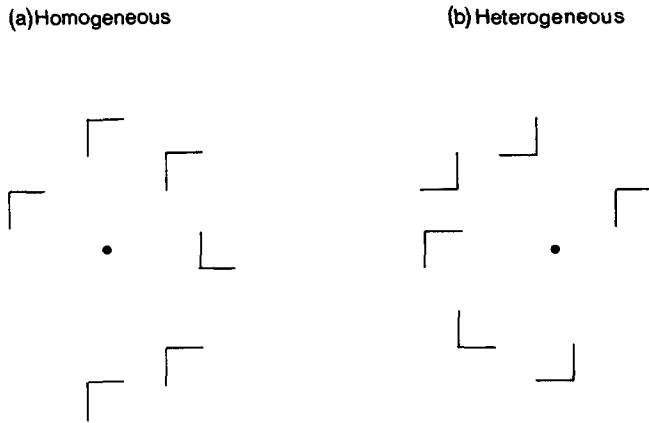


Figure 1. Sample displays from Experiment 3.
(Size/eccentricity ratio = 1/3.)

of the target shape, preserving the orientation of individual strokes. We shall consider possible elementary features distinguishing such rotations after presenting the next experiment.

Experiment 3 also shows that, when nontargets are rotations of the target shape, the task is made very difficult by mixing different nontarget rotations in the display. Humphreys et al. (in press) reported a similar result: Search for an inverted T is very much easier among homogeneous upright Ts than among a mixture of both upright and sideways Ts. This very large effect of nontarget homogeneity in Experiment 3 contrasts with the negligible effect obtained using the same target (L) in Experiment 1.

Experiment 4

Consider the relation between an upright L and the four rotations of a T shown in Figure 2. Letting the length of a single stroke be s , Ts at 180° and 270° differ from the L only in displacement of a single stroke by $s/2$, whereas Ts at 0° and 90° differ from the L in this same displacement, plus an additional displacement of the other stroke by s . Given the constraint that their strokes should be horizontal and vertical, then, the nontarget Ts used in Experiment 1 (0° and 90°) may have been those most easily discriminable from the target. To confirm this, we asked 8 subjects to classify singly presented letters as Ls or Ts,

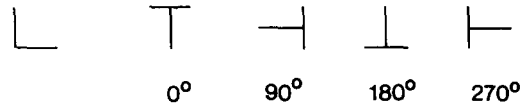


Figure 2. An upright L compared with a T at four rotations.

indicating their choice with a speeded keypress. Letters were in the same display positions, and at the smaller size, to be used in Experiment 4. Ls were always upright, whereas Ts were either at 180° and 270° (difficult condition) or at 0° and 90° (easy condition). Reaction times were 22 ms longer in the difficult condition, $F(1, 7) = 5.9, p < .05$. With this in mind, Experiment 4 investigated search for an upright L among the more difficult 180° and 270° rotated Ts.

One reason for examining these alternative nontarget rotations derives from the work of Bergen and Julesz (1983). Their task was search for a T among Ls, with all letters randomly oriented. Randomizing letter orientation requires that each nontarget be discriminated from a target in any possible orientation, including those in which it would most resemble the actual stimulus. It seems possible that this may in part have been responsible for the large effect of display size that was obtained.

Method

Except that nontargets were Ts rotated 180° and 270° clockwise from the upright, Experiment 4 was similar in all respects to Experiment 3. The 4 subjects, 2 of each sex, were between 19 and 27 years of age.

Results

Mean RTs from the third day of practice are shown in Table 9. The obvious difference from Experiment 1 was that slopes showed a fairly large effect of nontarget homogeneity, especially when letters were small. Even with homogeneous nontargets, slopes were somewhat greater than in the most comparable conditions of Experiments 1 and 2. Even in this case, therefore, increasing the similarity between target and nontargets may have a small effect. The effect is much bigger, however, when nontargets are heterogeneous.

An ANOVA showed a significant effect of display size, $F(2, 6) = 46.5, p < .001$, a marginal effect of nontarget homogeneity, $F(1, 3) = 9.6, p < .06$, and a significant interaction between them, $F(2, 6) = 9.7, p < .02$. The only other significant effects were letter size, $F(1, 3) = 13.2, p < .05$, and its interaction with target presence, $F(1, 3) = 35.2, p < .01$.

A direct comparison of search for Ls in Experiments 1 and 4 is weak because the two experiments differed in several respects. An ANOVA on RTs showed, however, a main effect of experiment, $F(1, 6) = 11.9, p < .02$, and an interaction between experiment and display size, $F(2, 12) = 15.8, p < .001$, modulated by an interaction of experiment, display size, and nontarget homogeneity, $F(2, 12) = 5.6, p < .02$. To the extent that these results can be believed, they support our conclusion that slopes may be affected by nontarget homogeneity only when targets and nontargets are relatively similar.

Table 8
Experiment 3: Error Proportions

Condition	Small letters (1/6)			Large letters (1/3)		
	2	4	6	2	4	6
Homogeneous						
Present	.016	.031	.039	.031	.039	.008
Absent	.008	.008	.008	.008	.008	.016
Heterogeneous						
Present	.016	.031	.063	.016	.063	.072
Absent	.031	.033	.008	.047	.016	.016

Table 9
Experiment 4: Reaction Times (in Milliseconds) as a Function of Display Size

Condition	Small letters (1/6)				Large letters (1/3)			
	2	4	6	Slope (ms/item)	2	4	6	Slope (ms/item)
Homogeneous								
Present	462	465	493	8	443	467	470	7
Absent	486	499	516	8	459	459	482	6
Heterogeneous								
Present	505	527	568	16	448	481	490	11
Absent	544	592	640	24	480	506	528	12

Error data from Experiment 4 appear in Table 10. They suggest no change in conclusions.

Discussion

The results of Experiment 4 confirm that, as conditions approach those studied by Bergen and Julesz (1983), search for an L among Ts shows an increasing effect of display size. Highly similar targets and nontargets, small letters, and heterogeneous nontarget orientations all contribute to the effect. Although the results of Bergen and Julesz (1983) are replicated under these conditions, their conclusion that serial processing is always needed to determine the arrangement of strokes within a letter is evidently too general.

Comparing the results of Experiments 1 and 4 suggests the important conclusion that, in our previous terms, T-N similarity and N-N similarity may interact. Increasing the similarity of targets to nontargets increases the importance of nontarget heterogeneity. It seems likely that this accounts, at least in part, for the very large effect of heterogeneity when nontargets are simple rotations of the target (Experiment 3).

In Experiment 4 as in Experiment 1, even heterogeneous nontargets had something in common (they were all Ts) that distinguished them from targets. Evidently, this alone is insufficient to nullify effects of heterogeneity. Other experiments we have run confirm this conclusion.

Assessment of Feature Integration Theory

It is clear that feature integration theory cannot account for the results of Experiments 1 to 4 if the elementary letter feature

coded at the first, parallel processing stage is taken to be the stroke. All four experiments show that, under some circumstances, search for particular stroke conjunctions is little affected by display size. Here, we consider various alternative assumptions.

The first possibility is that the corner of an L is a feature distinct from the junction of a T. This proposal is sufficient to account for the results of Experiments 1 and 2, but not Experiments 3 and 4. Experiment 3 shows that, when nontargets are homogeneous, a corner-junction distinction between targets and nontargets is not a necessary condition for insensitivity of RT to display size. Experiment 4 shows that, when nontargets are heterogeneous and closely similar to targets, it is not a sufficient condition either.

A second proposal revises the theory more radically—the elementary feature may be some part of a letter (stroke, corner, etc.) linked to its position within the letter or to some other spatial attribute (e.g., direction of pointing). We consider three versions of the idea. All can account for some aspects of the data, but all can be eliminated in the same way: by showing that none is a sufficient condition for display size insensitivity when T-N similarity and N-N similarity are appropriately manipulated. First, the elementary feature could be the stroke coupled to its within-letter location. Results reported by Duncan (1987) eliminate this possibility. The target was an upright L, whereas nontargets were a mixture of rotated Ls (Γ or \perp) and arrows combining the major diagonal of the letter with either the horizontal or the vertical stroke of the target (\lrcorner or \llcorner). In some conditions the target had a stroke that was unique in terms of its position within the letter (e.g., nontargets Γ and \lrcorner), but in others it did not (e.g., Γ and \llcorner). There was no difference between conditions; both gave a large effect of display size (22-ms/item on the third day of practice). Second, the elementary feature might be the corner coupled to either its location within the letter or its direction of pointing. The present Experiment 3 eliminates this possibility because display size insensitivity is predicted even with heterogeneous nontargets. Third, the elementary feature might be the free line terminator coupled to its direction of pointing. An L, for example, has free terminators pointing upwards and to the right. Eliminating this possibility, Humphreys et al. (in press) obtained large effects of display size (50-ms/item) in search for an inverted T among a mixture of upright Ts and Is (formed by the union of upright and inverted Ts). Similarly, in pilot work, we have replicated Experiment 3

Table 10
Experiment 4: Error Proportions

Condition	Small letters (1/6)			Large letters (1/12)		
	2	4	6	2	4	6
Homogeneous						
Present	.031	.024	.078	.031	.039	.047
Absent	.024	.016	.024	.023	.031	.031
Heterogeneous						
Present	.016	.110	.063	.008	.031	.070
Absent	.000	.016	.024	.039	.016	.016

using an upright \vee (60° angle between strokes) as a target and \vees rotated by 120° clockwise (∇) or counterclockwise (∇) as nontargets. The same very strong interaction between display size and nontarget homogeneity was observed, as before, despite the fact that upward pointing terminators were unique to the target even with heterogeneous nontargets.²

The final, most radical proposal is that the elementary feature is the whole letter. All conditions showing large effects of display size are inconsistent with this possibility.

The general point here is that no account of the results in terms of feature integration theory seems internally consistent. Let us analyze the reason. According to this theory, the key consideration is the target-nontarget relation and, in particular, that set of distinguishing properties possessed by the target but by no nontarget. If a discriminable feature is included in this set, then search will be parallel; otherwise, it will be serial. We can find no description of letter properties, however, that reveals a consistent association between sets of distinguishing properties and good or poor performance. When this happens we suspect that the theory is neglecting some further, important variable. (If a model has too few degrees of freedom, it is impossible to obtain an internally consistent fit to the data.) This further variable, we suggest, is the relation between one nontarget and another.

Stimulus Factors Determining the Efficiency of Target Selection

Principles

By the standards of the literature, Experiments 1-4 have shown large variations in search efficiency across stimulus materials. Given the inadequacy of feature integration theory to account for these results, we turn now to an alternative characterization of how stimulus factors determine search difficulty.

Our results suggest four general principles:

1. Search efficiency (as indexed by the effect of display size) varies continuously across tasks and conditions. There is no clear implication of a dichotomy between serial and parallel search modes.
2. Search efficiency decreases with *increasing* T-N similarity. The point is most clearly made by comparing Experiment 4 with Experiments 1 and 2.
3. Search efficiency decreases with *decreasing* N-N similarity. The point is most clearly made in Experiments 3 and 4.
4. The preceding two factors interact to scale one another's effects. Increasing T-N similarity has relatively little effect when N-N similarity is high (compare homogeneous nontarget conditions in Experiments 1 and 4). Decreasing N-N similarity has little effect if T-N similarity is low (Experiment 1). Alterations in search efficiency can only be understood by considering both variables together.

Figure 3 summarizes our proposals graphically. T-N similarity and N-N similarity are plotted, respectively, on the x-axis and the z-axis, with a reversal of scale between them. As a joint function of these two, the effect of display size is plotted on the y-axis. Axes are not marked numerically because the figure is intended only as a qualitative summary. It reflects the following

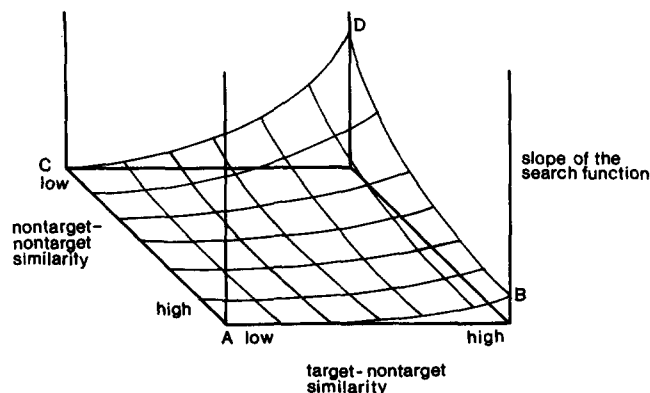


Figure 3. The search surface.

hypotheses. When T-N similarity is low, search is always highly efficient, irrespective of N-N similarity (line AC). This hypothesis is consistent with the independence of display size and nontarget heterogeneity observed in Experiment 1. When N-N similarity is maximal (nontargets identical but for spatial translation), T-N similarity has a relatively small effect (line AB). Such a small effect is suggested by comparing the homogeneous nontarget conditions of Experiments 1 and 4. When N-N similarity is reduced, however, T-N similarity becomes very important (line CD). Comparing the heterogeneous nontarget conditions of Experiments 1 and 4 supports this suggestion. Similarly, when T-N similarity is high, N-N similarity has a very substantial effect (line BD and Experiment 4). Performance is worst at point D: T-N similarity is high and N-N similarity is low or, in other words, nontargets each have much in common with the target but rather less in common with one another. This relates very plausibly to the heterogeneous condition of Experiment 3, in which the two nontargets were simple rotations of the target, each sharing one stroke with the target but no strokes with one another. Interpolation between these boundaries produces a continuous surface of search efficiency. Its exact shape is speculative, but we intend its general form (implied by the preceding hypotheses) to be correct.

Experiments 1 to 4 were planned as tests of feature integration theory. Although the difficulties they raise for that theory are quite clearcut, the results provide only preliminary support for the alternative account summarized in Figure 3. Findings are limited to a small domain of stimulus materials, and we have used only rough manipulations of letter similarity. We noted too that comparisons across experiments were questionable because experiments were not all run under the same conditions. To bolster our conclusions, we show next that the search surface is consistent with a wide range of published results, in

² The same result eliminates a more elaborate proposal suggested by a journal reviewer. Suppose that an angle's direction of pointing is coded in terms of elementary features up, down, left, and right. In Experiment 3, the target's angle (down and left) would have a unique feature only when nontargets were homogeneous. In the same experiment with Vs, however, the target's downward pointing angle would always be unique.

search for letters, shapes, color patches, single lines, and so on. Indeed, it accounts for a great deal of the variation in search difficulty observed across different stimulus materials.

To speak of stimulus similarity is fairly unambiguous when stimuli differ only along simple dimensions like size and color. Here, there are simple physical scales for measuring such differences, and many psychophysical studies confirming that proximity on such physical scales is closely associated with discriminability for the human observer. For more complex stimuli, such as letters differing in shape, the situation is more complex. Because such stimuli have multiple attributes, they can be categorized in different ways. Thus, a pair of letters can be similar in one respect and different in another; and the relative importance of different stimulus attributes may depend upon the psychological context. As one example, Beck (1966) showed that the perceived boundary between an area containing upright Ts and an area containing tilted Ts is stronger than the corresponding boundary between areas of upright Ts and Ls. These results suggest the relative importance of line orientation. In direct similarity ratings, however, it is the upright and tilted Ts that are judged to go together, suggesting reliance on other attributes. We shall return to the contrast between perceptual grouping and single element classification later on. For the moment, we may note that, even for shapes, there remain *some* physical constraints on plausible notions of similarity, backed up by psychophysical work. As we might expect, the confusability of two letters is strongly correlated with the amount of contour that they share (Townsend, 1971), a constraint consistent with our earlier argument concerning similarity of an L to Ts at different rotations (Figure 2). As an extreme case, identical letters are the same in all their attributes (cf. our manipulations of N-N similarity in Experiments 1-4). Such constraints allow us to speak of similarity between letters and shapes without complete circularity. Each time we do so, however, we shall try to make clear the particular attributes in respect of which similarity is specified, either by definition or by examples.

Simultaneous-Successive Comparisons

We begin with a small group of experiments that measure the accuracy of target detection or recognition in displays of constant size, presented either all at once (simultaneous) or two at a time (successive). As we have already mentioned, the technique is useful because it holds constant the total amount of material presented but varies the amount to be processed at any one time. Unlimited-capacity parallel models predict no difference between simultaneous and successive conditions (Eriksen & Spencer, 1969), contrasting with their inexact prediction of some unknown effect when the total amount of material changes (as in RT tasks).

The range of results obtained with this technique extends from identical performance in simultaneous and successive conditions (Shiffrin & Gardner, 1972) to successive advantages of 0.8 or more in d' (Kleiss & Lane, 1986). On the whole, the results are consistent with the view that there will be a large successive advantage only when T-N similarity is high and N-N similarity is low. In the classic experiment, Shiffrin and Gardner (1972) manipulated T-N similarity but always maintained

maximal N-N similarity (identical nontargets). Their target was T or F, whereas nontargets were either all T-F hybrids or all Os. In either case, performance in simultaneous and successive conditions was identical. At the opposite extreme, Kleiss and Lane (1986) obtained large successive advantages with high T-N similarity (shared contour criterion) and heterogeneous nontargets (e.g., Experiment 4: mean successive advantage of 0.8 in d' for detecting R among mixed Ps and Qs, and 0.6 for detecting R among mixed Ps and Bs). Duncan (1980b) and Kleiss and Lane (1986) obtained intermediate results using the targets 2 to 9 and nontargets A, E, G, J, M, P, T, X. Further experiments with this technique would be very useful, especially using stimuli whose similarities were more fully understood. For example, when targets are digits and nontargets letters, should we be considering similarity of visual or categorical descriptions? At least, however, the contrasting results of Shiffrin and Gardner (1972) and Kleiss and Lane (1986) do suggest a major effect of N-N similarity.

Reaction-Time Tasks

Reaction-time tasks have been used much more widely. As we have noted, effects of display size vary from little or none to over 100-ms/item. Here, there is information on a much wider range of stimulus materials and direct study of both T-N and N-N similarity.

The results leave little doubt that across a wide range of stimulus materials, both forms of similarity exert strong effects on search difficulty. As for T-N similarity, Neisser (1963) first showed that a letter target is much harder to find when it is physically similar to nontargets (e.g., both target and nontargets are angular or both are curved). Subsequent experiments have confirmed this result using letters (Corcoran & Jackson, 1977), color patches (Carter, 1982; Farmer & Taylor, 1980), lines varying in curvature (Treisman & Gormican, 1988), contours varying in the size of a gap (Treisman & Souther, 1985), and so on. Fewer experiments have dealt directly with N-N similarity, although it may be noted that many well-known reports of "parallel search" with simple shape stimuli used identical nontargets (e.g., Beck & Ambler, 1973; Donderi & Case, 1970; Donderi & Zelnicker, 1969; Egeth et al., 1972, Experiments 1 and 2). Where direct comparisons have been made, they confirm that search is facilitated by increasing N-N similarity, whether varying the number of different nontarget letters (Gordon, 1968; McIntyre, Fox, & Neale, 1970), or using simple stimuli such as color patches (Farmer & Taylor, 1980). The data of Eriksen (1953) also suggest that search is impaired by varying nontarget color.

What of the detailed interaction between T-N and N-N similarity? In Figure 3, line AC reflects the hypothesis that, if targets are sufficiently unlike nontargets, search time is always independent of display size, irrespective of nontarget heterogeneity. In part, this hypothesis reflects the results of the present Experiment 1 and, in part, the following theoretical consideration. In the normal visual environment, selection of behaviorally relevant information must generally operate in the presence of large but irrelevant contrasts throughout the visual field. If selection is ever to be efficient, it must be possible to disregard such irrele-

vant contrasts at least when conditions are otherwise favorable, that is, when nothing irrelevant in the field is closely similar to the information sought. A direct test of this hypothesis is described by Duncan (in press, Experiment 1). Stimuli were color patches chosen, within the constraints of a color monitor, to be as dissimilar as possible (red, yellow, green, blue-green, blue, purple). Across days, each color served as the target equally often; displays contained from 2 to 5 patches, with no color repeated within a display. Under these conditions of minimal (and equal) T-N and N-N similarity, search times were almost independent of display size. Studies of search for single letters and digits also show that, when T-N similarity is minimized (on the basis of such considerations as straight vs. curved shapes), search times can be almost independent of display size even with heterogeneous nontargets (see data and reviews in Duncan, 1983; Schneider & Shiffrin, 1977).

Line AB in Figure 3 reflects the proposal that, when N-N similarity is maximal (identical nontargets), increasing T-N similarity has a relatively small effect. In letter search, the experiments of Humphreys et al. (in press) are the clearest demonstration that, even when targets and nontargets have a great deal in common (search for an inverted T among upright Ts), search time can be almost independent of display size if nontargets are all identical. That there is some effect of increasing T-N similarity even with identical nontargets, however, has been confirmed in search for simple targets defined both by length and by brightness (Treisman & Gormican, 1988).

Two final experiments have investigated T-N similarity, N-N similarity, and their interaction in detailed analysis of search for color patches. The first was reported by Farmer and Taylor (1980). Targets were greys varying in brightness, whereas nontargets were colors of middle brightness, chosen all to be equally distinct from the grey. T-N similarity was determined by target brightness, being highest for the target whose brightness matched nontargets. N-N similarity was determined by the set of nontarget hues chosen from the middle-brightness color circle, either maximally dissimilar (e.g., red, yellow, green, blue, purple) or highly similar (e.g., five hues between green and blue). At the fixed display size of 15 items, search times showed the predicted effects of T-N similarity, N-N similarity, and their interaction. The second experiment was described by Duncan (in press, Experiment 2). To understand the method, consider the list of four colors: green (1), yellow (2), orange (3), red (4). Both physically and psychologically, the adjacent pairs (1-2, 2-3, 3-4) in this list are most similar, whereas the extreme colors (1-4) are most dissimilar. (This reflects the standard geometrical representation of color space as a circle, red-orange-yellow-green-blue-purple-red.) By dividing the list into two pairs, 1-4 and 2-3, we may contrast two tasks, in one of which (middle nontargets) 1-4 are targets and 2-3 are nontargets, whereas in the other (end nontargets) 2-3 are targets and 1-4 are nontargets. In terms of T-N similarity, the two tasks are exactly matched; in terms of N-N similarity, however, they are very different. In terms of Figure 3, we remain at a fixed point on the *x*-axis and examine a large move along the *z*-axis. The experiment may then be repeated at different points along the *x*-axis by choosing sets of colors 1 to 4 more widely or more narrowly spaced. Using this method and a variety of color sets,

Duncan (in press) obtained results in good agreement with the search surface. With wide color spacing (low T-N similarity), search time was almost independent of display size in both middle-nontarget and end-nontarget cases (cf. line AC, Figure 3). With closer color spacing (high T-N similarity), effects of display size were rather larger and, in particular, a very large advantage for the middle-nontarget case emerged (line BD). With close color spacing and end nontargets, slopes were at least as high as those typically reported for conjunction search.

Of course, in most search studies, T-N and N-N similarities have been left to the chance selection of stimulus materials, producing a wide range of RT slopes (e.g., Duncan, 1983) with little apparent order. Comparing stimulus similarities across studies and materials (color patches, letters, words, etc.) is not generally meaningful. Where they have been directly tested, however, the proposals summarized in the search surface have been supported. Of particular importance is confirmation of these proposals in experiments using simple stimuli, for which the manipulation of similarity is relatively straightforward. There can be little doubt that the conclusions based tentatively on the present Experiments 1 to 4 in fact reflect principles determining search difficulty across a wide range of stimulus materials. In the next section, we turn to a theory of how T-N and N-N similarity exert their effects.

We should make it clear that we regard the search surface as a general, somewhat approximate summary of a wide range of search findings. For example, we have not attempted quantitative measures of stimulus similarity nor to establish with such measures that the search surface has a fixed shape across tasks or stimulus dimensions. Of course, such extensions of our work would be extremely interesting. In this article, though, the approximate empirical principles that we already have will prove sufficient for our theoretical development.

A Theory of Visual Selection

Overview

Our theory has several antecedents. In structure it resembles the theory of Duncan (1980b, 1985). There are three components: (a) a parallel stage of perceptual description, producing a structured representation of the input across the visual field and at several levels of spatial scale, (b) a process of selection by matching input descriptions against an internal template of the information needed in current behavior, and (c) entry of selected information into visual short-term memory (VSTM), allowing control of effectors (and corresponding to access to awareness). Our idea of an internal template comes partly from Naatanen's (1985) "attentional trace." Our account also owes much to Bundesen (1987; personal communication, October 27-30, 1987), with whom we have had many useful discussions.

Some key features of our theory may be mentioned in advance. It is general across search materials, dealing in the same way, for example, with feature and conjunction search though directly suggesting why conjunction search is often so difficult. An essential element is a distinction, not considered in previous work, between similarity of stimulus alternatives and similarity of stimuli within a display. This distinction leads to consider-

ation of hierarchical structure in perceptual representation and its influence on selection. Although our major motivation is to account for the search surface, as we proceed, we draw together a number of additional findings. We give an account of the relation between search and whole report and of the importance of perceptual grouping in search and selection. Finally, our account relates sensibly to selection in natural, complex visual environments.

The Theory

Description

We call the first stage of processing in our model *perceptual description* (Sutherland, 1968). This stage produces the data base for subsequent operations: a parallel, hierarchically structured representation of input across the visual field, from which a part can then be selected for control of immediate behavior. This representation is highly processed. For example, shapes have been categorized in such a way as to make contact with nonvisual properties (e.g., aspects of meaning) in memory. Similarly, there will be descriptions of color that are independent of the illuminant, of size independent of distance, and so on.

Like many others (e.g., Beck, Prazdny, & Rosenfeld, 1983; Marr & Nishihara, 1978; Palmer, 1977; Sutherland, 1968), we note that visual input has a natural hierarchical or part-whole organization. On a page of text, for example, relations between letters within a word are both more constrained and more important than relations between letters from different words; similarly, spatial relations between fingers on the same (seen) hand are more important and more constrained than relations between fingers on different hands. To capture these distinctions, a visual representation must be segmented. Parts that are to be described within the same whole must be linked together (Gestalt grouping) or, complementarily, boundaries must be drawn between parts that are to be described separately. Following Palmer (1977), we call the segments thus created *structural units*, although they resemble also the object files of Kahneman and Treisman (1984) and the 3-D models of Marr and Nishihara (1978).

The details of segmentation are not important here (see, e.g., Beck et al., 1983; Grossberg, 1987; Julesz, 1981). Following many others, we simply note that many different considerations combine to determine where segmentation occurs. Spatial factors, in particular the tendency to group by proximity, often dominate. Similarity grouping is also important, and will play a major role in our discussion of search. Thus, elements in a field with the same color, same motion, same shape, or other shared properties will tend to group together or, complementarily, boundaries between structural units will occur wherever there are local discontinuities in such properties. Perceived volumes may often be segmented into parts wherever there are sharp concavities in outline (Biederman, 1987). Probably, too, segmentation tends to produce familiar entities: We should expect rather different segmentation patterns in the strings THECAT and TACEHT (Prinzmetal & Millis-Wright, 1984).

A fully hierarchical representation is created by repeating segmentation at different levels of scale (Marr & Nishihara,

1978; Palmer, 1977). Each structural unit, contained by its own boundary, is further subdivided into parts by the major boundaries within it. Thus, a human body may be subdivided into head, torso, and limbs, and a hand into palm and fingers (Marr & Nishihara, 1978). Such subdivision serves two purposes. The description of a structural unit at one level of scale (animal, letter, etc.) must depend heavily on the relations between the parts defined within it (as well as on properties such as color or movement that may be common to the parts). Then, at the next level down, each part becomes a new structural unit to be further described with its own properties, defined among other things by the relations between its own sub-parts. At the top of the hierarchy may be a structural unit corresponding to the whole input scene, described with a rough set of properties (e.g., division into light sky above and dark ground below).

Within the hierarchy, we propose that each structural unit is described with a set of properties, including at least its relative location, motion, color, surface texture, size, and shape. As we have said, nonvisual properties (e.g., categorizations based on meaning) may also have been added from memory. We propose that the entire process of segmentation and description is parallel and resource free. In visual search, for example, the time taken to process one part of the field is not dependent on how many elements occur elsewhere. Although the whole input description is derived in parallel, however, it has no immediate control over behavior. Subjectively, this description is outside awareness. To serve as the focus for action, a part of the input description must be selected for access to VSTM. Here, we propose that structural units act as wholes, competing for and gaining access to VSTM with all their associated descriptions (Kahneman & Henik, 1977).

Selection

Access to VSTM is strictly limited. The classic experiment was reported by Sperling (1967). A row of letters was exposed for a variable duration, terminated by a mask. Subjects reported as many letters as possible. The number reported increased by roughly one per 10-ms increase in exposure, to a maximum of three or four. Once this maximum was reached, further increases in exposure had little effect.

Consider first the initial, steep increase in number of items reported with increasing exposure. Following Sperling (1967), we take this to reflect the entry rate of information into VSTM.³ Although a roughly linear increase might suggest a serial process, Sperling himself favored a parallel, limited-capacity model, and subsequent stronger evidence confirms this view (Pashler & Badgio, 1987).

Thus, we propose that access to VSTM is determined as follows. There exists some resource, limited in total amount, that can be divided in varying proportions among structural units in the input description (Rumelhart, 1970). Increased assignment of resources to any structural unit increases its speed and probability of access to VSTM (Bundesen, 1987; Shibuya &

³ In Sperling's (1967) model it is actually the "recognition buffer" that corresponds functionally to our visual short-term memory.

Bundesen, 1988). Resource assignment is strictly competitive: Increasing the assignment to one structural unit necessarily decreases the net assignment to others. Following Rumelhart (1970), we may think of each structural unit as having some weight, reflecting the strength with which it competes (cf. the "attentional strengths" of Shiffrin, Dumais, & Schneider, 1981; and the "impacts" of Bundesen et al., 1985). Note then that resource allocation depends on relative rather than absolute weights (Rumelhart, 1970). Total available resources are always used and are distributed across inputs in proportion to relative weights. For example, increasing or decreasing all weights equally has no great effect.

We intend *resource* to be understood as neutrally as possible. It might refer to any factor competitively distributed among structural units. For example, as proposed in a theory developed by Bundesen (1987; personal communication, October 27–30, 1987), the speed and probability of VSTM access for any structural unit could be determined by its degree of activation, with total activation across structural units fixed (e.g., through mutual inhibition). Then, increases and decreases in weight might correspond respectively to excitatory and inhibitory inputs.

For simplicity we assume that weights are initially set to some constant, average value for all structural units in the input description. For selection, then, the problem is to alter or bias these initial weights so that resources are assigned to the correct structural units. Such a selection system must be indefinitely flexible. Depending on present concerns, descriptions at any location or scale level may be behaviorally relevant. In this respect we propose that vision is typically *directed* by the search for some particular information (e.g., the next word in a page of text, an approaching vehicle at an intersection) whose advance specification, however, may vary substantially in precision. We call this advance specification a *template* of the information sought, although with no commitment to common ideas of template matching and with the explicit proposal that templates vary in both the nature and detail of the information they contain. For example, a template could specify only one attribute of the desired information, such as its location or color, or many combined attributes, including shape. Selection operates by matching input descriptions to current templates. The result is a change of weight for each structural unit that is proportional to the degree of match, good matches increasing the weight but poor matches decreasing it.

One further consideration links this account of selection to the idea of hierarchical representation. In an artificial search display it may seem reasonable to limit consideration to the few stimulus elements that are presented by the experimenter, but in a realistic, natural image the problem is more complex. The field contains a multiplicity of structural units at different levels of scale. Suppose that one searches for red berries hanging in a holly tree. Is their template compared separately and independently with a description of every visible leaf, or even with every individual spike separated by concavities from other parts of each leaf? Working independently in this way would seem alarmingly inefficient. Instead, we propose that weights tend to change together to the extent that structural units are linked in the input description (i.e., according to the strength of grouping

between them). To put this another way, a change in weight for any one structural unit is also distributed to others proportionally to the strength of grouping. This process we call *weight linkage*.

Again, it would be easy to model weight linkage in terms of activation strengths in the input description. In connectionist models, for example, it is common for units consistent with the same perceptual hypothesis to support one another's activity (e.g., McClelland & Rumelhart, 1981). Accordingly such units tend to gain or lose activation together. Here, we might suppose a similar functional linkage, proportional to the strength of perceptual grouping, between different structural units in the input description.

The most important result of weight linkage is efficient rejection of strongly grouped nontargets through what we call *spreading suppression* between them. At one level of scale, each leaf of the holly tree is a distinct structural unit, but at the next level up, these are linked into homogeneous green areas broken by the branches and clusters of berries. Close association means that each reduction in weight for one leaf tends also to be distributed to others—or in other words, that all such reductions reinforce one another.

Note that we do not propose a reduction in any element's weight *simply because* it is similar to or grouped with others ("lateral inhibition"). Instead, we propose that weights *covary* according to grouping strength. There is no bias against selecting parts of the input that are grouped with others (Ullman, 1984), only a bias toward selecting or rejecting grouped parts together.

To sum up, two factors combine to determine selection weights. The first is match of each input against a template of currently needed information. Weights increase with increasing match. The second is weight linkage. Any change in weight for one input is distributed to others in proportion to the strength of perceptual grouping.

Visual Short-Term Memory

Access to VSTM allows a structural unit to become the focus of current behavior. It may be described, picked up, and so on. We need not decide *why* it is that access to VSTM is so severely limited. Along with Posner (1978) and Allport (1987), we may speculate that one major reason is the need to avoid overload and conflict in response systems, which ideally should *only* "know" as much as they currently need. Here, though, the simple existence of limitation is the important thing.

Returning to Sperling's (1967) experiment, we recall that, with increasing exposure duration, the number of letters reported from a tachistoscopic array asymptotes at three to four. Again, we may accept Sperling's view that this asymptote reflects the total capacity of VSTM. How to define capacity will be left open; for example, the limit might concern number of structural units or of descriptions and might perhaps be sensitive to the hierarchical structure of selected information. When VSTM is filled, it must be flushed before the entry of new information can begin. Perhaps flushing is normally accompanied by refixation, in accord with the suggestion of Loftus (1981) that fixations last for the (variable) amount of time needed to

acquire a fixed amount of information from the visual scene. Thus, limited VSTM access has both a parallel and a serial element, parallel within each period of filling the store, but serial from one such period to the next.

Account of Visual Search Data

From the preceding account it is clear that we shall have to consider two sorts of stimulus similarity in visual search. First, there is the problem of classifying each single element in a display as target or nontarget. In the aforementioned theory this involves matching each element against a template of possible targets. To understand this process we need to consider similarities within a task's list of stimulus alternatives (i.e., the possible targets and nontargets); these we call *interalternative similarities*. Second, there is the issue of perceptual grouping. Here the important consideration is the similarity of one stimulus to another within a display; this we call *within-display similarity*. We will show that both forms of similarity have major effects on search.

In outline, the approach we take may already be clear. Increasing T-N similarity is harmful because in the competition for VSTM access, the weight of each nontarget depends on its match to target templates. Decreasing N-N similarity is harmful because it reduces the opportunity for spreading suppression of grouped nontargets. In what follows we both supplement this view with other, less central arguments and show why in broad outline we prefer it to other possible accounts of similarity effects.

Selecting a Target Description

We begin by defining the most usual and general strategy for experimental search tasks. Structural units in the input description are compared against a template of the target(s). A positive response is made when a selected target description enters VSTM. Negative responses are made by default, as we consider in detail later.

Interalternative similarity and selection efficiency. Target detection times will be independent of display size if nontargets attract no resources. Otherwise, increasing the number of nontargets will be harmful, inasmuch as it must diminish resource availability for the target. Because the weight of each nontarget depends on how closely it matches a target template, the effect of T-N similarity follows readily. (By extension, too, we assume that background aspects of the input description, that is, features of the room other than the experimental display, have so little weight as to be ignored.)

Several previous workers have made similar proposals about T-N similarity. Most notable is Hoffman (1978), who directly proposed that nontargets resembling a possible target gain preferential access to limited-capacity perceptual systems, but in rather different terms, Estes (1972), too, suggested that one stimulus inhibits the perception of others in a display to the extent that it is like a possible target. This view is significantly different, however, from several others.

First, our account is in terms of interalternative similarity. We have considered not similarities between one stimulus and

another within an actual array, but rather the similarity of each nontarget to any possible target. In contrast, Bjork and Murray (1977) proposed an account of T-N similarity entirely in within-display terms. In their theory, characters in a display inhibit one another to an extent dependent on their activation of shared feature detectors. This predicts a form of camouflage: A target will be hard to see when a similar nontarget is nearby. An important experiment by Pashler (1987b) illustrates one case in which such within-display factors are unimportant, the resemblance of each nontarget to any possible target being the key consideration. Possible targets were the letters C and E; nontargets were a combination of letters chosen to resemble neither target (X and N), mixed (on some trials) with letters chosen to be confusable with one target or the other (Gs or Fs in different displays). The presence of confusable nontargets was harmful, and significantly, their effect was *independent of whether the target they resembled was the one actually present*. At least in this experiment, interalternative similarity was crucial and within-display similarity was immaterial.

Second, consider accounts based on classification *difficulty*. It is very plausible that the resource demand of a display element should depend on how hard this element is to classify. This is the straightforward prediction of any serial model in which the time spent on an element depends on how long it takes to accumulate sufficient evidence for an accurate decision. Similar proposals were made by Neisser (1963) and Broadbent (1970) in terms of the number of features to be extracted from search elements (see also, more recently, Fisher et al., 1988). Our account differs subtly. Nontargets similar to targets call for resources not because they are close to any classification boundary separating target and nontarget categories, but because they are close to the target category itself. The resulting prediction is obvious and has often been confirmed: Irrespective of classification difficulty, targets themselves make the greatest resource demands (Duncan, 1980b, 1985; see also Ostry, Moray, & Marks, 1976; Sorkin, Pohlmann, & Gilliom, 1973).

Before leaving our discussion of interalternative similarity, we must deal with one further matter. If we bear in mind that the important factor is similarity between each input description and a target *template* (as opposed, e.g., to the complete description of a target *input*), then a further consideration arises. What factors determine the nature of the template? Clearly it must be sufficiently elaborate to include all likely targets and to exclude all likely nontargets. It follows that selection will in all probability be impaired by increased heterogeneity either within the set of possible targets or within the set of possible nontargets.

In this article we are little concerned with heterogeneity of targets (T-T similarity), although of course it is known that increasing the number of possible targets (memory set size) often has a harmful effect (Schneider & Shiffrin, 1977). The more heterogeneous the targets, the more elaborate will be the template or set of templates needed to differentiate them from nontargets. (An exception occurs when targets can all be differentiated from nontargets on the basis of some common property. In this case, target heterogeneity might be unimportant.) Similarly, the more heterogeneous are nontargets, the less likely is it that a simple target description will exclude them all. For ex-

ample, suppose the target is a line varying randomly in length between 1.75 cm and 2.25 cm. If nontargets are all lines of length 3 cm, the target description could be *shorter than 2.5 cm*. Adding a second nontarget of length 1 cm, however, would require the more complex description *shorter than 2.5 cm but longer than 1.5 cm*.

There are two general approaches to the problem of why more elaborate templates should lead to poorer selection. The first is based on the complexity of the required matching operation. In the case of target lines from 1.75 cm to 2.25 cm, for example, a single criterion along the dimension of length will discriminate targets from 3-cm nontargets, whereas two criteria are needed when the 1-cm nontarget is added. The idea that more features might need to be checked with more heterogeneous sets of target and nontarget letters (cf. Rabbitt, 1967) is analogous. The second approach is more consonant with the present emphasis on similarity considerations. A distinction is often drawn between the shared and distinguishing attributes of two entities (e.g., Tversky, 1977). If the selection weight assigned to a nontarget depends largely on its shared attributes with the template, then the more attributes a template possesses, the more on average is it likely to share with each nontarget and the worse will selection be. For example, functional similarity between a line of 3 cm and the description *shorter than 2.5 cm* may be less than functional similarity with the description *shorter than 2.5 cm but longer than 1.5 cm*, because in the second case, the 3-cm nontarget in fact possesses one of the two attributes in the target description.

In fact, it is a very general psychological truth that classification becomes easier with (a) decreased similarity between stimuli requiring different responses and (b) increased similarity between stimuli requiring the same response (Shepard & Podgorny, 1986). Furthermore, the latter point (b) is generally explained by appeal either to the complexity of required decision operations or to considerations of functional similarity between members of different response classes (Podgorny & Shepard, 1983). The arguments of the present section reflect an application of these general psychological principles to the particular act of selection between competing input descriptions.

In this section we have dealt with harmful effects of increased T-N similarity and decreased N-N similarity in interalternative terms. We turn next to evidence that these considerations alone are far from sufficient for a full account of similarity influences on search.

Within-display similarity and perceptual grouping. In the preceding section we argued that search would be impaired by increased heterogeneity either within the set of possible targets or within the set of possible nontargets. Recall, however, the experiment of Duncan (in press, Experiment 2), described in detail earlier. A set of four colors was divided either so that targets were relatively homogeneous and nontargets relatively heterogeneous, or so that targets were relatively heterogeneous and nontargets relatively homogeneous. Despite the fact that the distinction between targets and nontargets was exactly the same in the two cases, the task with homogeneous nontargets was substantially easier.

Although exactly comparable letter search studies have not

been done, there, too, a similar asymmetry may exist. Although increasing the number of targets from one to two is initially harmful, practice soon reduces the effect or eliminates it altogether (Schneider & Shiffrin, 1977). Increasing the number of nontargets from one to two, however, has large and lasting effects (Duncan, 1987).

Interalternative considerations suggest no reason for this asymmetry between T-T and N-N similarity. The obvious asymmetry in the tasks, however, is that displays contain multiple nontargets but not multiple targets. The implication, as our theory predicts, is that similarity of one stimulus to another within the display may be crucial. N-N similarity is much more important than T-T similarity because similar nontargets in a display tend to group perceptually. Because of weight linkage, they do not compete independently for VSTM access.

To reiterate, the more closely are two inputs linked in the hierarchical input description, the stronger is the tendency for their weights to covary. In search this means that reducing the weight of any one nontarget in a strong perceptual group tends correspondingly to reduce it for the rest. This is what we have called *spreading suppression*. Although separate groups of nontargets must be rejected independently, members of the same group tend to be rejected together.

Again, several previous workers have expressed similar views, in particular that nontargets in a strong perceptual group may be rejected as a unit (Bundesen & Pedersen, 1983; Farmer & Taylor, 1980). The strongest evidence for such a position comes from experiments dealing with nontarget grouping directly. In an experiment by Bundesen and Pedersen (1983), for example, a target of known color was presented in displays containing various mixtures of nontarget colors. When nontargets were grouped into local areas of common color, there was little effect (about 2-ms/item) of the number of items per group, but a much larger effect (about 12-ms/item) of the number of groups. In other words, adding a nontarget to a display had little effect when it simply increased the size of an already existing group, but a much larger effect when it created a new group. Similar results were reported by Farmer and Taylor (1980) for color search and by Treisman (1982) for targets defined by a conjunction of color and form. Indeed, adding nontargets to a display can even be helpful if it increases perceived grouping of the remainder (Banks & Prinzmetal, 1976). Such findings cannot be explained by considerations of interalternative similarity. They clearly indicate the importance of interactions between one element and another within a display.

Although these experiments deal with *nontarget* grouping, the tendency of targets to group with nontargets must also be important. A tendency for two weights to covary is helpful if both are to be set low (two nontargets) or high (two targets), but must be harmful if one is to be set low and the other high. Thus, our analysis of T-N similarity must also have its within-display aspect. Grouping between targets and nontargets will be harmful. We predict, for example, that targets could sometimes be camouflaged by placing them close to similar nontargets in an array.

In general, because access to VSTM is competitive, the most important consideration will be *relative* weights of targets and nontargets. Distributing a reduction in weight from one nontar-

get to other input elements will only be helpful if it reduces the weights of (some) other nontargets more than it reduces the weight of a target. The necessary condition for a benefit from spreading suppression must be that grouping between a set of nontargets be stronger than grouping between nontargets and targets. In the experiment of Bundesen and Pedersen (1983), for example, distributing a reduction in weight within a color group is helpful because weights within the group change more than all other weights. Although weight changes may also have some tendency to be distributed across color groups, this will be unimportant because its average effect on nontargets in these other groups will be equal to its effect on any target.

Effects of selection failure. Before leaving the issue of selection efficiency, we should consider the detailed implications of selection failure. Assignment of resources to nontargets has two different harmful effects. First, it reduces resources available for the target and, hence, the average speed of VSTM access. Second, in the extreme case, VSTM may be filled by nontarget descriptions before target access is achieved, so that the store must be flushed (by re-fixation?) and selection started again.

A study by Pashler (1987a) fits rather neatly with this analysis. In a detailed study of display size functions for typical conjunction search (e.g., green Ts among red Ts and green Os), Pashler (1987a) found that positive and negative RTs increased by roughly the same amount from display size 2 to 8, followed by diverging functions (2:1 ratio of negative:positive slopes) thereafter. He proposed that a limited-capacity parallel process deals with clumps of about eight items, but moves in a serial self-terminating search from clump to clump.

Finally, we should consider how negative responses are made. In all probability, the exact strategy is rather variable. At one extreme, if the assignment of resources is so selective that any target is always the first description entered into VSTM, then the negative response can safely be made either when any nontarget description enters or after a suitable time interval. At the opposite extreme, if resources are assigned equally to targets and nontargets, the negative response cannot safely be made until descriptions of all display elements have passed into (or through) VSTM. In fact, negative response functions are notoriously variable, and many authors have suggested that such responses are based on variable checking strategies (Corcoran & Jackson, 1977; Treisman & Souther, 1985). In general, however, when the search strategy is the one we have considered so far, negative responses will be slower than positive responses (because they are made by default), and any factor impeding selection will affect negatives at least as substantially as positives.

Special Case: Homogeneity Coding

We have said that for any input image, structural units will be described at various levels of scale. In most search arrays, one such unit will correspond to the whole patch of stimulus elements, described with its overall shape (in the present experiments, an irregular arc), brightness, texture, and so on. This immediately raises the possibility of a special strategy for carrying out the search task. Suppose that the presence of a target produces some characteristic change in the description of this whole-array unit. (For example, one can imagine tasks in which

a target would make the whole array larger or smaller or would alter the array's outline shape.) Then there is no need to match each element of the array against a template of the target, making a positive response when a target description enters VSTM. Instead, the whole-array unit can be selected (by match to a corresponding template, e.g., the template of a bright, irregular arc, which excludes background features of the room). A response can be based on the perceived properties of this whole-array unit.

With this strategy, we expect little affect of array size because elements do not compete individually for access to VSTM. Furthermore, there is no reason for negative responses to be slower than positive responses. Negative responses are not made by default. The whole-array unit simply has one perceived property if the target is present and another if it is absent. Such fast negatives are rarely observed with heterogeneous nontargets, even in the purest cases of "parallel" search (Duncan, in press, Experiment 1; Schneider & Shiffrin, 1977; Taylor, 1978; for an exception, see Duncan, 1983, Figure 2). With homogeneous nontargets, however, negatives can be as fast as positives (Corcoran & Jackson, 1977; Donderi & Case, 1970; Egeth et al., 1972; Humphreys et al., in press). Sometimes such negative RTs even decrease with increasing array size, especially from display size 1 (when whole-array homogeneity does not distinguish positive from negative trials) to 2 (Humphreys, Riddoch, & Quinlan, 1986; see also Donderi & Case, 1970; Donderi & Zelnicker, 1969; Polich, 1986). A final observation is that RTs can be independent of array size even when the task is simply to decide whether a whole array is homogeneous ("same") or not ("different"), with no advance specification of the identity of any mismatching element (Donderi & Case, 1970; Donderi & Zelnicker, 1969). All of these results suggest that, when nontargets are all identical, responses can sometimes be based on direct coding of homogeneity or heterogeneity at the level of the whole-array unit.

In fact, this is a reasonable suggestion. We have said that description of any structural unit includes a specification of how and where it is divided into parts (Marr & Nishihara, 1978; Palmer, 1977). In an array composed of multiple elements, major divisions into parts occur where adjacent elements are dissimilar in color, shape, or some other property (Beck et al., 1983; Julesz, 1981). Adding a target to a field of otherwise identical nontargets may indeed produce, at the level of the whole-array unit, a unique perceived division into subgroups.

This is a speculative possibility, and it makes little difference to our account of RT slopes. With homogeneous nontargets, slopes might be shallow either because responses are based on direct, perceived homogeneity of the whole array or because spreading suppression allows efficient target selection. Only absolute RTs may be much influenced by the choice of strategy. Furthermore, we would regard response to whole-array homogeneity as a special, rather degenerate case of search. In a natural visual environment it might be unlikely that we can detect relevant information simply because some higher level entity in the field is less homogeneous than otherwise expected. Still, the possibility of this strategy should perhaps be acknowledged in any complete account of N-N similarity effects.

In essence we have proposed that description of any struc-

tural unit includes information concerning the largest local discontinuities within it. Along related lines, it has been suggested by both Sagi and Julesz (1984) and Ullman (1984) that parallel visual processes are entirely confined to local discontinuity detection. Parallel processes indicate the element in the field that is most unlike its surroundings but are insufficient for any task requiring stimulus identification. Like us, Sagi (1988) has also recently shown that within-object conjunction search can be very easy when nontargets are all identical. According to our account, one sufficient condition for parallel search is indeed that the target be the display element most unlike its surroundings (line AB on the search surface). This is not a necessary condition, however (line AC; see the data reviewed earlier, and in particular, Duncan, *in press*, Experiment 1). If T-N similarity is sufficiently low for perfect selection, then N-N similarity is immaterial.

Size/Eccentricity Ratio

Our experiments suggest that decreasing the size/eccentricity ratio increases slopes in all letter search tasks. This follows straightforwardly from the theory. Decreasing size/eccentricity ratio makes all differences between letters harder to see. This will impair any discriminative operation based on such differences, including (a) discriminative setting of target and nontarget weights based on match to a target template and (b) any benefit from spreading suppression, arising when grouping between a set of nontargets is stronger than grouping of nontargets with a target.

T-T Grouping

The stronger is the perceived grouping of two inputs, the greater is the tendency for their selection weights to change together. Spreading suppression between grouped nontargets is one side to this. If arrays contained multiple targets, however, then increasing their perceived grouping should also allow spreading enhancement between them. An increase in weight for one target would be distributed to others proportionally to the strength of grouping. As we have said, N-N grouping is more important than T-T grouping only because of an asymmetry in the structure of typical search arrays, which contain multiple nontargets but not multiple targets. The analogue of a search task with multiple targets is partial report. Here several studies have confirmed that selecting a number of letters from a brief array is easiest when they form a strong perceptual group (Kahneman & Henik, 1977; Merikle, 1980; note, however, that in such studies T-T, T-N, and N-N grouping all necessarily covary).

Related to this is our hypothesis that structural units, with all their associated descriptions, are selected as wholes for VSTM access. In support, Duncan (1984) showed that two properties of the same object (e.g., brightness and tilt of a line) can be identified simultaneously in a brief display, without mutual interference, whereas properties of two different objects cannot. The different descriptions attached to a single structural unit are perhaps the most strongly grouped aspects of an input description; this may be why they are necessarily selected together.

Summary

The theory we have proposed both explains the general form of the search surface and reveals its limitations. Returning to Figure 3, line AC shows that, irrespective of N-N similarity, search time is always independent of array size as long as T-N similarity is sufficiently low. In the competition for VSTM access, reducing T-N similarity eventually renders nontarget weights negligible. Line AB shows that even increasing T-N similarity has relatively little effect when N-N similarity is maximal (identical nontargets). When N-N grouping is stronger than T-N grouping, selection is aided by spreading suppression. (Of course, as T-N similarity approaches N-N similarity, performance must eventually be impaired. Thus, although line AC is horizontal, AB must eventually turn up.) But, as we move along lines CD or BD, performance rapidly deteriorates. Nontargets increasingly match the target template, and spreading suppression is decreasingly helpful. The result is a continuous degradation in search efficiency.

One major limitation of the search surface is its bundling together of two quite different similarity considerations. We have proposed that, in the competition for VSTM access, two different factors combine to determine selection weights. The first is a tendency of each element to attract resources in proportion to its match to a target template. Here, the relevant consideration is interalternative similarity. With reference to standard classification principles (Shepard & Podgorny, 1986), we argued that appropriately assigning weights will be enhanced by (a) increasing similarity within the set of possible targets, (b) increasing similarity within the set of possible nontargets, and (c) decreasing similarity between possible targets and nontargets. The second factor is perceptual grouping. To the extent that two elements are linked in the input description, their weights tend to covary. Here, the relevant consideration is within-display similarity. We argued that selection will be enhanced by (a) increasing grouping between simultaneous nontargets and (b) decreasing grouping between targets and nontargets. If displays contained multiple targets, then increasing their perceived grouping would also be helpful. Although these analyses of interalternative and within-display similarity have obvious and interesting correspondences,⁴ it is important to keep them apart. Suppose, for example, that we wished to extend the search surface to give a quantitative account of similarity effects

⁴ In brief, much the same principles may be at work in different representational domains. For analysis of interalternative similarity, the relevant domain is the representation in memory of alternative possible stimuli. Distance in this representation concerns the dimensions of variation used in mentally classifying stimuli. We may think of classification as drawing a boundary in this space of stimulus alternatives to distinguish target and nontarget areas. Classifying any given element is easier the further it is from this boundary (T-N similarity) and the simpler the boundary is (T-T and N-N similarity). For analysis of within-display similarity, the relevant domain is representation of the input across visual space. Distance corresponds to grouping strength. We may think of selection as drawing a boundary in this input representation to distinguish selected and unselected areas. This is easier the further apart are selected and unselected areas (T-N grouping) and the more compact is each one (T-T and N-N grouping).

in some stimulus domain. We have no guarantee that interalternative confusability and interelement grouping will respond in exactly the same way to any given manipulation, whether of color, shape, or any other attribute. Indeed, the considerations arising in analyzing these two problems are necessarily somewhat different. This is our major reason for offering the search surface as a first, approximate summary of experimental findings. To go further, we need independent measurement and understanding of stimulus principles governing interalternative confusability in input-template matching and grouping strength in perceptual description.

A second limitation of the search surface is neglect of T-T similarity. As we have shown, N-N similarity is in fact more important in search as it is usually investigated (Duncan, in press, Experiment 2). In the more general case of multiple targets and nontargets, however, T-T and N-N effects may well be symmetrical.

Across-Object Conjunctions

Experiments 1-4 dealt with within-object conjunctions. In those experiments, individual nontargets differed from the target only in the arrangement of strokes within the letter. Among other things, this allowed us to examine the extreme case of nontarget homogeneity (identical nontargets). As yet, however, we have said nothing about across-object conjunctions. One of the key findings in support of feature integration theory is the difficulty of search when a target can be formed by recombining attributes of the different, heterogeneous nontargets (e.g., search for a red O among blue Os and red Xs). Can our theory suggest an account of this result?

According to the theory there are two sufficient conditions for efficient visual search. The first is that N-N similarity should be appreciably higher than T-N similarity (Figure 3, line AB). Because grouping between all nontargets is stronger than grouping of any nontarget with the target, there is maximal benefit from spreading suppression. In typical across-object conjunction search, this condition is clearly violated. The two alternative nontargets each share one attribute with the target, but no attributes with one another.

The second sufficient condition is that T-N similarity should be so low that nontarget weights are negligible (Figure 3, line AC). In analyzing this we must recall two key points. First, the important consideration is similarity between nontargets and the target template, not between nontarget and target stimuli. Second, selection may be expected to suffer as templates become more elaborate. In particular, the more attributes are specified in a template, the more on average will be shared with each nontarget.

In their classic experiment, Treisman and Gelade (1980) compared two tasks. Nontargets in each case were green Xs and brown Ts. In feature search, the target was either blue (in which case it was either an X or a T) or it was an S (in which case it was either green or brown). Either target could occur on any given trial. In conjunction search, the target was a green T. The experiment was carefully designed to equate both N-N and T-N similarity across tasks, construing T-N similarity as the number of attributes shared between each nontarget and the

most similar possible targets (e.g., a green X shares one attribute with either a blue X or a green S; similarly, it shares one attribute with a green T). The conjunction task, nevertheless, was substantially more difficult.

According to our theory, the variable not matched across tasks was similarity between each nontarget and the target template. If we assume that templates included *only* those attributes distinguishing targets from nontargets, then in feature search, each nontarget shared no attributes with either of the templates *blue* and *S*, whereas in conjunction search, each nontarget shared one attribute with the template *green T*. If we suppose that allocation of resources to nontargets is determined partly by the number of attributes shared with target templates (as opposed just to the number of distinguishing attributes), the relative difficulty of conjunction search readily follows.

"Search asymmetries" reported by Treisman and Souther (1985) may be explained by similar means. For example, it is very much easier to find a circle with an intersecting line among circles without a line, than to find a circle without a line among circles with lines. If the target is a circle with a line, the template may simply specify *line* or *intersection*, but if the target is the circle with no line, only a complete template *circle and absence of line* can be used. (Obviously, the subject cannot search just for "no line"—most parts of the field will contain no line.) Other asymmetries reported by Treisman and Gormican (1988) may be explicable in the same way. For example, it is easier to find a slightly tilted line among vertical lines than the reverse. Treisman and Gormican suggested that a tilted line may be described perceptually as the standard (vertical) plus the deviation (tilt). Then, in search for a tilted line, only the deviation need be specified in the template, whereas in search for the vertical, the whole (*standard and absence of deviation*) must be included.

Because (by definition) templates will never be the same in feature and conjunction search, it may be impossible to achieve a perfectly controlled comparison between the two. According to our theory, a better approach is to examine wide variations of T-N and N-N similarity in both sorts of task. We have already reviewed experiments on search for color patches showing that, as T-N similarity increases and N-N similarity decreases, results come closely to match those typically reported for conjunction search (e.g., Duncan, in press, Experiment 2). Experiments 1-4 dealt with maximal N-N similarity in within-object conjunction search. Experiment 5 deals with T-N similarity in across-object conjunction search.

Experiment 5

Treisman and Gelade (1980) showed that it is easier to find an R among mixed Ps and Bs than among mixed Ps and Qs. In the first case, the target has a unique stroke (the diagonal), whereas in the second it is unique only in its conjunction of strokes. Thus, the results are consistent with feature integration theory. According to our theory, the R-PB task is easier because a simpler template (a description just of the diagonal) can be used.

In Experiment 5, we compared the R-PQ task with a modification designed to reduce T-N similarity while maintaining

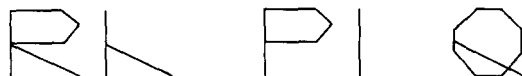


Figure 4. Stimulus set in Experiment 5. (Targets are on the left, nontargets on the right.)

the condition that all target strokes be contained in the nontargets. Much of the shared contour between an R and a P lies in their common loop. We asked what happens when the bulk of this loop is simply omitted. The target becomes a caret (Figure 4, second from left), whereas nontargets become I and Q (Figure 4, right). How does this reduction in T-N similarity—as assessed again by amount of shared contour (Townsend, 1971)—affect search?

As we have said, although the number of attributes distinguishing each nontarget from the target template may be similar in feature and conjunction search, the number of shared attributes is not. The more complex templates of conjunction search may be harmful because selection weights depend partly on shared attributes. In this light it may be noted that our manipulation of T-N similarity left the contour distinguishing targets and nontargets the same. Only shared contour was altered.

A preliminary experiment asked how our manipulation affects the time to classify single shapes. In one task the two alternative stimuli were R and P, whereas in the other they were the caret and I. Shapes were identical to those used in Experiment 5, presented in the same display locations. A single shape appeared on each trial, and the subject indicated which it was by pressing one of two alternative keys. Reaction times were 53 ms shorter in the caret-I task, $F(1, 3) = 45.7, p < .01$.

Method

Tasks. Stimuli are shown, drawn to scale, in Figure 4. The two alternative targets—an R and a caret—are on the left, and the three nontargets—P, I, and Q—are on the right. There were two tasks: search for an R among Ps and Qs, and search for a caret among Is and Qs. The only difference was that, in the second task, the loop shared by an R and a P was omitted.

Except for the I, which occupied the same position as the upright in

Table 11
Experiment 5: Reaction Times (in Milliseconds)
as a Function of Display Size

Condition	Heterogeneous			Slope (ms/item)
	2	10	18	
Caret in IQ				
Present	450	453	502	3
Absent	494	576	608	7
R in PQ				
Present	512	691	916	25
Absent	595	996	1516	58
Caret in PQ				
Present	439	473	484	3
Absent	482	593	652	11

Table 12
Experiment 5: Error Proportions

Condition	Heterogeneous		
	2	10	18
Caret in IQ			
Present	.021	.016	.011
Absent	.011	.011	.011
R in PQ			
Present	.005	.042	.032
Absent	.026	.016	.026
Caret in PQ			
Present	.000	.016	.016
Absent	.016	.011	.005

a target, each letter measured 29' arc square. Displays contained either 2, 10, or 18 letters, positioned at random on two imaginary circles of radius 1° 23' (letter height/eccentricity = 1/2.9) and 1° 57' (letter height/eccentricity = 1/4.1) centered on fixation. There were 12 possible letter positions on each circle, evenly spaced, starting at 12 o'clock for the inner circle and at 12:30 for the outer circle. In the previous experiments, nontargets had always been shifted slightly from their positions on the display perimeter, so that a target's strokes fell in positions that were possible for nontarget strokes. This procedure was unnecessary here because the experiment concerned across-object rather than within-object conjunctions. Instead, display regularity was avoided by assigning letters randomly to display locations.

Other procedural details were copied from Experiment 3.

Design. Each subject attended for three sessions. Each session was divided into two blocks, one for each task, with the order counterbalanced across subjects. Each block had one run of 24 practice trials followed by three runs of 96 experimental trials. In the third session only, after completion of the main experiment, there was a further block of 24 practice plus 3 × 96 experimental trials of search for a caret among Ps and Qs.

Subjects. The 4 subjects, all women, were between 21 and 33 years of age.

Results

Mean RTs from the third session appear in Table 11. There was a very large difference between the two tasks. Search for a caret among Is and Qs was scarcely affected by display size. Search for an R among Ps and Qs replicated the results of Treisman and Gelade (1980): Large slopes in a ratio of about 2:1 for target-absent and target-present responses. An ANOVA comparing these two tasks showed significant effects of task, $F(1, 3) = 100.2, p < .005$; display size, $F(2, 6) = 44.3, p < .001$; target presence, $F(1, 3) = 22.9, p < .02$; and every interaction: Task × Display Size, $F(2, 6) = 60.0, p < .001$; Task × Target Presence, $F(1, 3) = 116.6, p < .005$; Display Size × Target Presence, $F(2, 6) = 36.0, p < .001$; and Task × Display Size × Target Presence, $F(2, 6) = 41.0, p < .001$. Last, the data at the bottom of Table 11 show that results did not depend on any special characteristic of the nontarget I. Search for a caret was equally easy whether nontargets were Is and Qs or Ps and Qs.

Error data appear in Table 12. They suggest no change in conclusions.

Discussion

To account for these results in terms of feature integration theory, we should have to propose that adding a loop to the caret caused it to *lose* some elementary feature shared by no other letter in the stimulus set. Again, we can find no plausible suggestion that cannot be eliminated by other findings.

One journal reviewer suggested three candidate features distinguishing the caret from I and Q: triangularity, intermediate amount of contour, and number of free line terminators (3). The last seems doubtful because our R also had more free terminators (2) than either P or Q (1 each, Figure 4), but to eliminate all three suggestions, we tested four unpracticed subjects in search for a caret among mixed Is, Qs, Ys, and triangles (whose corners were in the positions of the caret's three free line ends). On target present trials, slopes were almost identical to those obtained in the basic task of search for a caret among Is and Qs. A second reviewer suggested that the caret's unique feature might be the roughly T-shaped junction of its two component lines. Again, we were doubtful because locally this intersection was identical in the caret and the Q (Figure 4). Still, we modified the task of search for a caret among Is, Qs, Ys, and triangles by subtracting from the Q all those lines contributing to its closed loop, except the leftmost three (Figure 4). This left a shape rather similar to the caret, but which in the neighborhood of the junction was identical to the previous Q. Now performance became very poor, the slope for target-present trials increasing by a factor of four.

From the perspective of our theory, the results confirm that decreasing T-N similarity can make even across-object conjunction search very easy. With one caveat they also support our hypothesis that selection weights are determined partly by attributes *shared* between nontargets and the target template. The R-PQ and caret-IQ tasks differ only in the amount of contour shared between target and nontargets, not in distinguishing contour. This is important because it helps to explain why conjunction search is often more difficult than feature search.

The caveat concerns the following thought. To account for the difference between feature and conjunction search, we have proposed that templates need not specify whole target shapes. For example, in search for R among Ps and Bs, only the diagonal stroke might be specified. If this is so, then in Experiment 5, why did the target template not specify only a caret even when the actual target was an R?

The explanation may be linked to the following observation. We asked 18 subjects to divide an R (cf. Figure 4) into parts in the most natural way possible. The diagonal stroke appeared as a natural part in 15/18 answers, whereas the caret appeared in 0/18. (Indeed, the vertical and the loop may be more closely linked than the vertical and the diagonal. Some subjects broke the vertical and made a single part out of the whole closed loop, whereas some broke the loop and made a single part out of the vertical. One way to put this is that the figure contains both a vertical and a closed loop, which are partly coextensive.) Obviously, we need further work on this issue. Still, it seems likely that, in search for shapes, templates can only be those natural parts or wholes that appear in the hierarchical input description.

As we have said, however, the main conclusion is simply that decreasing T-N similarity can make even across-object conjunction search very easy. Related results were reported by Quinlan and Humphreys (1987). Their target was a small green H; in two conditions of conjunction search, nontargets either shared one attribute each with the target (large green As, large orange Hs, small orange As) or two attributes (small green As, small orange Hs, large green Hs). Slopes for positive and negative trials were 12- and 29-ms/item, respectively, with one shared attribute, as compared with 37- and 83-ms/item for two shared attributes. Reducing T-N similarity reduced slopes by a factor of three. Very recently, Wolfe, Cave, and Franzel (in press) have used a similar method to reduce conjunction search slopes almost to zero.

There remains little reason to suppose that feature and conjunction search differ in principle. By manipulating T-N and N-N similarity we can make feature search arbitrarily difficult or conjunction search arbitrarily easy. In either case, stimulus similarities are the basic consideration determining search difficulty.

Relation to Other Theories

Feature Integration Theory

The theory we have proposed differs from feature integration theory in two major respects. First and most obvious, we suggest that feature and conjunction search are in principle very similar. By manipulating T-N and N-N similarity, we can make either one very easy or very difficult. Second, feature integration theory is concerned almost exclusively with the relation between targets and nontargets. (For a partial exception see Treisman, 1982.) Scant attention is given to interactions between elements within a display: The overwhelming consideration is the procedure of classifying each separate input as target or nontarget. For this reason, the theory has little to say about why the nontarget-nontarget relation is so important in both feature (Duncan, in press, Experiment 2) and conjunction (this article, Experiments 1-4) search.

It is also useful, however, to mention points of contact between the two theories. One of these concerns Treisman and Gormican's (1988) suggestion that arrays can be searched in clumps of items varying in size. This is not the same as our idea that strongly grouped nontargets tend to be rejected together. In Treisman and Gormican's (1988) model, clumps of items do not have to be alike. Clump size is determined only by the average amount of noise activation that each nontarget produces in the target map (i.e., by the average similarity of targets to nontargets). Their proposal that search can be parallel within clumps but serial from clump to clump does relate, however, to our view that search has both a parallel and a serial element, parallel within each period of filling VSTM but serial from one such period to the next.

We may also note a modification to feature integration theory suggested recently by Wolfe et al. (in press; see also Treisman, 1988). Recall that, according to the theory, elementary features are initially registered in multiple, independent maps of the visual field. According to Wolfe et al. (in press), outputs from

these maps might then be combined to determine the focus of attention. Locations in each feature map are linked to corresponding locations in a single master map (Treisman & Souther, 1985). The most activated location in the master map becomes the focus of attention. In feature search, only inputs from the target map to the master map are enabled. If search is for a red item, for example, only activity in the redness map is passed to the master map, and attention is drawn at once to red targets. In conjunction search, by extension, inputs from two feature maps to the master map might be enabled. In search for a red O among blue Os and red Xs, for example, the master map would receive activation from both redness and curvature maps. The region occupied by the target would again be the most activated in the master map, although against a background of partial activation from other regions. According to Wolfe et al. (in press), how harmful this background activity is depends on target–nontarget discriminability.

This of course is a rather profound modification of feature integration theory. As Wolfe et al. (in press) noted, it makes feature and conjunction search very similar in principle. Information about the *combination* of an object's attributes comes together to determine the focus of attention. And of course, even this modification has nothing to say about interactions between one nontarget and another and, hence, about the effects of N–N similarity and nontarget grouping. We do see this, however, as a plausible account of how input–template matching is achieved when the target is defined by at least some conjunctions of different attributes. The target template is distributed in the form of enabled connections between two or more different feature maps and the master map. The model gives a good account of why *shared* attributes between inputs and this template are harmful; inputs activating any of the maps whose master map connections are enabled add noise to the process of directing attention (Treisman, 1988; Wolfe et al., in press). Finally, retaining the idea of initially distinct representation for (some) different visual attributes is consistent with neurophysiological evidence (Maunsell & Newsome, 1987) and may allow us still to account for the range of different psychological phenomena (e.g., illusory conjunctions) to which feature integration has been applied (Treisman & Gelade, 1980; Treisman & Schmidt, 1982).

In this article we have dealt only with selection efficiency and, hence, whether adding nontargets to a display has a generally large or small effect. Two other important questions concern (a) the shape of the search function and (b) the relation between functions for negative and positive trials. In Treisman's work, functions for conjunction search are often steep and linear, with a negative/positive slope ratio of 2/1. This suggests serial, self-terminating search. Functions for feature search are often flatter and negatively accelerated, with a negative/positive ratio of more than 2/1. This is taken to indicate parallel search. Commonly, however, the different criteria for serial and parallel search do not cohere. For example, we noted before that functions for positive and negative responses in conjunction search may be parallel up to a display size of about eight, only diverging thereafter (Houck & Hoffman, 1986; Pashler, 1987a). Flat display size functions for positive responses may sometimes be associated with much steeper functions for negatives, but just as

often, negatives also are quite flat (e.g., Treisman & Souther, 1985; Duncan, in press, Experiment 1). Linearity is hard to assess with relatively flat functions because restriction of range assures that the proportion of variance accounted for by linearity is likely to be low. Even in conjunction search an average 2/1 slope ratio for negative and positive trials can be characteristic of few individual subjects (Wolfe et al., in press).

We have suggested that the variable relation between positive and negative RTs reflects variable response strategies. Similarly, we would suppose that a variety of factors influences linearity, including the probability that VSTM is filled before a target is found, introducing a serial element into search (Pashler, 1987a), and the detailed relation between number of elements and number of perceived homogeneous groups in a heterogeneous display (Bundesen & Pedersen, 1983). We regard both the relation between positive and negative RTs and the exact shape of search functions as complex issues in their own right. More work would be needed before our theory could give a good account of them.

Late Selection

In architecture, the present account resembles Duncan's (1980b) late selection theory of attention. A defining feature of late selection theories concerns how highly processed are input descriptions even at the parallel, preselective stage. According to these theories, in particular, categorical processing of input patterns is sufficient for selection to be guided by such considerations as alphanumeric class (Duncan, 1980b) or (for spoken words) semantic category (Ostry et al., 1976).

In contrast, early selection theories propose that, before selective stages of processing, many perceptual distinctions are simply not "analyzed" and hence are not available for any discriminative purpose. A somewhat corresponding point in the present work is that access to VSTM becomes increasingly indiscriminate with increasing similarity between targets and nontargets, a principle that seems to hold whatever the selection criterion, from a simple attribute such as color to a more complex attribute such as the conjunction of strokes in a letter. It would be wrong to conclude, however, that fine perceptual distinctions are unavailable for any purpose before resources are allocated for VSTM access. The same fine distinctions that are used only poorly in selection (e.g., the distinction between Ls rotated by 90°) may be used perfectly in grouping (see Experiment 3). Different uses of the same information have differing accuracy (Duncan, 1985).

In most visual search studies, the visual details of targets are known in advance of the display. Under these circumstances, visual similarity considerations may always dominate. We need more work on selection when target information is specified more loosely, perhaps just by rather nonspecific relevance to a task like driving a car or shopping for a present.

Automatic and Controlled Processing

We have not dealt here with the substantial effects of practice on search (Schneider & Shiffrin, 1977). Practice increases the selection weight of stimuli that are usually targets and decreases

the weight of those that are usually nontargets (Schneider & Fisk, 1982). This may be seen as learning an indwelling selection tendency, supplementary to the short-term tendency to select momentarily relevant information that we have discussed here. Correspondingly, it is well-known and desirable that important information can draw attention even when not anticipated (or relevant) in the immediate context (Moray, 1959).

Another part of Schneider and Shiffrin's (1977) view is that continuous variations in search efficiency reflect varying mixtures of serial and parallel processing. This is an idea that we have worked out in some detail here.

Segmentation and Grouping

Our theory is inconsistent with the common view that perceptual grouping is sensitive only to a few, simple stimulus characteristics such as the length, orientation, and color of line elements (e.g., Julesz, 1981). According to our account, grouping must be influenced even by complex characteristics of shapes, at least when display elements are sufficiently large (see, e.g., Experiment 3). Of course, studies dealing with perceptual segmentation commonly use large fields consisting of very many elements. As an unintended result, the size/eccentricity ratio of a single element is often much less than the ratios used here.

Much more consistent with our theory is the idea that, whatever the characteristic distinguishing elements in the parts of an array, the strength of segmentation depends partly on simple discriminability (Enns, 1986).

Extensions

Interleaved Nontarget Groups

According to our theory, weight linkage depends on the strength of perceptual grouping. This, in turn, is strongly influenced by spatial proximity. Recall, for example, the experiment of Bundesen and Pedersen (1983). The main factor influencing search time was neither the number of nontargets (N_n) in a display, nor the number of different nontarget colors (N_c), but rather the number of homogeneous color groups (N_g). For given N_n and N_c , RT was minimized by spatial arrangements of nontargets than minimized N_g . (Indeed, RT could be well predicted from subjects' direct judgments of how many color groups a display contained.) We have taken this as strong evidence for the importance of perceived nontarget grouping and weight linkage. Beyond this, it emphasizes the role of spatial factors in grouping. It is *adjacent* elements of the same color that form a strong perceptual group, or in other words, elements of the same color may not group strongly when differently colored elements are placed between them. If this had not been so, then the number of judged groups (and RT) would have depended only on N_c .

Sometimes, however, perception creates structural units whose parts are spatially interleaved in the input image. If an animal runs behind a bush, common movement ensures coherence of its parts despite interleaving in the image with bush parts. An experiment by Nakayama and Silverman (1986) makes the same point for visual search. Stimuli were patches of

moving dots, presented stereoscopically at two apparent depths. The target was unpredictably either a far patch whose dots moved upward or a near patch whose dots moved downward; nontargets were a mixture of near/upward and far/downward patches. Although this is conjunction search, there was little effect of display size. Subjects reported that the field could be segmented perceptually into two interleaved groups, separated in depth. Each group, in turn, was simply checked for homogeneity. Related results using other grouping dimensions have been reported by Steinman (1987) and McLeod, Driver, and Crisp (1988).

These results indicate one way in which perceptual grouping and weight linkage might allow performance to be independent of array size even in conjunction search. More generally, if the weights of all identical nontargets in an array can be closely linked, irrespective of spatial arrangement, then RT will depend only on the number of different *types* of nontarget that are present.⁵

It is perhaps interesting that subjects in Nakayama and Silverman's (1986) study reported focusing on each depth plane in turn. Absolute RTs were indeed very long, suggesting a time-consuming switch from one perceptual organization to another. When grouping is complex or weak, perhaps it must be supported by some top-down influence. For example, increasing the selection weight (activation?) for all elements in one depth plane might allow them to be linked into a new structural unit, a unit then perceived as foreground against the background of all other elements. Although we have treated segmentation and selection independently, there may well be important interactions between them.

Nontarget Heterogeneity on Relevant and Irrelevant Dimensions

The dimension(s) along which targets are distinguished from nontargets may be called the relevant dimension(s) of a search task. Nontargets might then differ from one another on either relevant or irrelevant dimensions. Some preliminary results reported by Treisman (1988) suggest that nontarget heterogeneity on irrelevant dimensions has little effect. If the target is defined by color, it is harmful to mix different nontarget colors in a display but not to mix different nontarget orientations. If the target is defined by orientation, it is mixed nontarget orientations that are harmful.

One suggestion might be that the processing of irrelevant dimensions is simply prevented or inhibited. Results described recently by Pashler (1988) show that this cannot be. The task was to detect a single element differing in shape from the others in a large array (e.g., a / in a field of Os). Although random variation in color had little effect (cf. Treisman's, 1988, find-

⁵ This is not a likely explanation for our findings concerning search for a caret among either mixed Is and Qs or mixed Ps and Qs. Even if the target is an R, array size has little effect when nontargets are homogeneous Ps (unpublished observations). If perceptual segmentation could organize a mixed P-Q array into two interleaved groups, one of Ps and one of Qs, then we should expect even search for an R to be very easy.

ings), a single element that differed in color from the rest was extremely harmful.

In analyzing these results, we must remember a key point. Because access to VSTM is competitive, the most important consideration is the *relative* selection weight of targets and nontargets. Correspondingly, spreading suppression has little effect if it influences targets and nontargets equally. When we analyze the effects of N-N similarity and grouping, the most important consideration is the strength of N-N linkage relative to T-N linkage.

Now consider manipulation of nontarget heterogeneity on a relevant dimension. As an example, let the target be a yellow horizontal line, whereas nontargets are green and red horizontals. In the homogeneous condition all nontargets in one display are the same, whereas in the heterogeneous condition the two nontarget colors are mixed. Here we have altered the average strength of N-N linkage but left T-N linkage the same. A large difference between conditions may be expected.

Nontarget heterogeneity on an irrelevant dimension is a different matter. Let the target be a vertical line, whereas nontargets are horizontal lines. In the homogeneous condition, all elements in a display are the same color (all red or all green). In the heterogeneous condition, colors of both targets and nontargets vary randomly (mixed reds and greens). This time the effect is simply a net change in the mean linkage of any element, target or nontarget, to the remainder. Relative N-N to T-N linkage is not affected. As before, spreading suppression between nontargets will be weaker in the heterogeneous condition because each nontarget differs in color from one half of the others. To exactly the same extent, however, spreading suppression from nontargets to the target will also be weaker. The difference between our manipulations of nontarget heterogeneity on relevant and irrelevant dimensions lies in their effect on T-N, not on N-N linkage!

By the same reasoning, a single nontarget that differs from others on an irrelevant dimension could have a profound effect. Now, the task might be to find a green / in a large field of green Os, with an occasional red O inserted. Because they suffer less from spreading suppression, both the green / and the red O will be favored in competition for VSTM access.

Parametric studies of nontarget heterogeneity on relevant and irrelevant dimensions remain to be done. Here we have a framework within which such studies might be analyzed.

Multiple Targets

In this article we have largely been concerned with the processing of multiple nontargets. How do we deal with multiple targets? There are two possibilities. As we have noted, if the task requires that separate structural units be entered into VSTM for each target, then increasing the number of simultaneous targets should be very harmful (as it usually is; see Bundesen et al., 1985; Duncan, 1980b, 1985). Occasionally, though, this may be avoided. As an extreme case, suppose that one to four target lines are presented in an otherwise blank field and the subject is to count them. Very plausibly, they will be perceived as forming a single structural unit whose shape (dot, line, triangle, quadrilateral) usually indicates the number of targets. Because

only this one unit need enter VSTM, there is no competition for resources.

Findings consistent with this analysis were reported by Sagi and Julesz (1984), although their targets were mismatching elements in a field of otherwise identical lines. (For a failure to replicate under these circumstances, see Folk, Egeth, & Kwak, 1988.) For the future, the important hypothesis is that perceiving properties of simultaneous targets without mutual interference depends on functional grouping. This contrasts with rejecting simultaneous nontargets in parallel, for which strong grouping is a sufficient but not a necessary condition.

Natural Visual Environments

The view of visual selection that we are proposing seems to us to fit sensibly with both the requirements and the experience of normal vision. If selection is to be effective in the usual rich visual environment, it must be possible to discard the vast majority of information efficiently. We tend only to experience selection difficulties when objects in the environment are indeed very similar: when we stare at a drawer of cutlery searching for a particular fork or at a supermarket shelf searching for a particular tin. Only in such constrained environments may demands on the selection system approach those of laboratory search tasks using similarly constrained stimulus materials. Presumably, such experiences strike us because we are used to a selection system that works.

We think that, in the present theory, we have the beginnings of a general account of how stimulus factors influence the selection of behaviorally relevant information from visual scenes. Explicit generalization beyond laboratory search tasks is one of the most important challenges for the future.

References

- Allport, D. A. (1987). Selection for action: Some behavioral and neurophysiological considerations of attention and action. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 395-419). Hillsdale, NJ: Erlbaum.
- Banks, W. P., & Prinzmetal, W. (1976). Configurational effects in visual information processing. *Perception and Psychophysics*, 19, 361-367.
- Beck, J. (1966). Effect of orientation and of shape similarity on perceptual grouping. *Perception and Psychophysics*, 1, 300-302.
- Beck, J., & Ambler, B. (1973). The effects of concentrated and distributed attention on peripheral acuity. *Perception and Psychophysics*, 14, 225-230.
- Beck, J., Prazdny, K., & Rosenfeld, A. (1983). A theory of textural segmentation. In J. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and machine vision* (pp. 1-38). London: Academic Press.
- Bergen, J. R., & Julesz, B. (1983). Parallel versus serial processing in rapid pattern discrimination. *Nature, London*, 303, 696-698.
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94, 115-147.
- Bjork, E. J., & Murray, J. T. (1977). On the nature of input channels in visual processing. *Psychological Review*, 84, 472-484.
- Broadbent, D. E. (1970). Stimulus set and response set: Two kinds of selective attention. In D. Mostofsky (Ed.), *Attention: Contemporary theories and analysis* (pp. 51-60). New York: Appleton-Century-Crofts.
- Bundesen, C. (1987). Visual attention: Race models for selection from multielement displays. *Psychological Research*, 49, 113-121.

- Bundesen, C., & Pedersen, L. F. (1983). Color segregation and visual search. *Perception and Psychophysics*, 33, 487-493.
- Bundesen, C., Shibuya, H., & Larsen, A. (1985). Visual selection from multielement displays: A model for partial report. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI* (pp. 631-649). Hillsdale, NJ: Erlbaum.
- Carter, R. C. (1982). Visual search with color. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 127-136.
- Corcoran, D. W. J., & Jackson, A. (1977). Basic processes and strategies in visual search. In S. Dornic (Ed.), *Attention and performance VI* (pp. 387-411). Hillsdale, NJ: Erlbaum.
- Donderi, D. C., & Case, B. (1970). Parallel visual processing: Constant same-different latency with two to fourteen shapes. *Perception and Psychophysics*, 8, 373-375.
- Donderi, D. C., & Zelnicker, D. (1969). Parallel processing in visual same-different decisions. *Perception and Psychophysics*, 5, 197-200.
- Duncan, J. (1979). Divided attention: The whole is more than the sum of its parts. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 216-228.
- Duncan, J. (1980a). The demonstration of capacity limitation. *Cognitive Psychology*, 12, 75-96.
- Duncan, J. (1980b). 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 and Psychophysics*, 34, 221-232.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, 113, 501-517.
- Duncan, J. (1985). Visual search and visual attention. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI* (pp. 85-105). Hillsdale, NJ: Erlbaum.
- Duncan, J. (1987). Attention and reading: Wholes and parts in shape recognition. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 39-61). Hillsdale, NJ: Erlbaum.
- Duncan, J. (in press). Boundary conditions on parallel processing in human vision. *Perception*.
- Egeth, H., Jonides, J., & Wall, S. (1972). Parallel processing of multielement displays. *Cognitive Psychology*, 3, 674-698.
- Enns, J. (1986). Seeing textons in context. *Perception and Psychophysics*, 39, 143-147.
- Eriksen, C. W. (1953). Object location in a complex perceptual field. *Journal of Experimental Psychology*, 45, 126-132.
- Eriksen, C. W., & Spencer, T. (1969). Rate of information processing in visual perception: Some results and methodological considerations. *Journal of Experimental Psychology Monograph*, 79 (2, Pt. 2).
- Estes, W. K. (1972). Interactions of signal and background variables in visual processing. *Perception and Psychophysics*, 13, 278-286.
- Farmer, E. W., & Taylor, R. M. (1980). Visual search through color displays: Effects of target-background similarity and background uniformity. *Perception and Psychophysics*, 27, 267-272.
- Feldman, J. A. (1985). Four frames suffice: A provisional model of vision and space. *The Behavioral and Brain Sciences*, 8, 265-289.
- Fisher, D. L., Duffy, S. A., Young, C., & Pollatsek, A. (1988). Understanding the central processing limit in consistent-mapping visual search tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 253-266.
- Folk, C. L., Egeth, H., & Kwak, H. (1988). Subitizing: Direct apprehension or serial processing? *Perception and Psychophysics*, 44, 313-320.
- Gordon, I. E. (1968). Interactions between items in visual search. *Journal of Experimental Psychology*, 76, 348-355.
- Grossberg, S. (1987). Cortical dynamics of three-dimensional form, color, and brightness perception: I. Monocular theory. *Perception and Psychophysics*, 41, 87-116.
- Hoffman, J. E. (1978). Search through a sequentially presented visual display. *Perception and Psychophysics*, 23, 1-11.
- Houck, M. R., & Hoffman, J. E. (1986). Conjunction of color and form without attention: Evidence from an orientation-contingent color after-effect. *Journal of Experimental Psychology: Human Perception and Performance*, 12, 186-210.
- Humphreys, G. W., Quinlan, P. T., & Riddoch, M. J. (in press). Grouping processes in visual search: Effects with single- and combined-feature targets. *Journal of Experimental Psychology: General*.
- Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1985). Interactive processes in perceptual organization: Evidence from visual agnosia. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI* (pp. 301-318). Hillsdale, NJ: Erlbaum.
- Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1986). *Evidence for textures derived from combined-feature information*. Unpublished manuscript.
- Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature, London*, 290, 91-97.
- 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.
- Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 29-61). Orlando, FL: Academic Press.
- Kleiss, J. A., & Lane, D. M. (1986). Locus and persistence of capacity limitations in visual information processing. *Journal of Experimental Psychology: Human Perception and Performance*, 12, 200-210.
- Loftus, G. R. (1981). Tachistoscopic simulations of eye fixations on pictures. *Journal of Experimental Psychology: Human Learning and Memory*, 7, 369-376.
- Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society London B*, 200, 269-294.
- Maunsell, J. H. R., & Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience*, 10, 363-401.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, 88, 375-407.
- McIntyre, C., Fox, R., & Neale, J. (1970). Effects of noise similarity and redundancy on the information processed from brief visual displays. *Perception and Psychophysics*, 7, 328-332.
- McLeod, P., Driver, J., & Crisp, J. (1988). Visual search for a conjunction of movement and form is parallel. *Nature, London*, 332, 154-155.
- Merikle, P. M. (1980). Selection from visual persistence by perceptual groups and category membership. *Journal of Experimental Psychology: General*, 109, 279-295.
- Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, 11, 56-60.
- Naatanen, R. (1985). Stimulus processing: Reflections in event-related potentials, magnetoencephalogram and regional cerebral blood flow. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI* (pp. 355-373). Hillsdale, NJ: Erlbaum.
- Nakayama, K., & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature, London*, 320, 264-265.
- Neisser, U. (1963). Decision-time without reaction-time: Experiments in visual scanning. *American Journal of Psychology*, 76, 376-385.
- Ostry, D., Moray, N., & Marks, G. (1976). Attention, practice and se-

- mantic targets. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 326-336.
- Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, 9, 441-474.
- Pashler, H. (1987a). Detecting conjunctions of color and form: Reassessing the serial search hypothesis. *Perception and Psychophysics*, 41, 191-201.
- Pashler, H. (1987b). Target-distractor discriminability in visual search. *Perception and Psychophysics*, 41, 285-292.
- Pashler, H. (1988). Cross-dimensional interaction and texture segregation. *Perception and Psychophysics*, 43, 307-318.
- Pashler, H., & Badgio, P. C. (1987). Attentional issues in the identification of alphanumeric characters. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 63-81). Hillsdale, NJ: Erlbaum.
- Podgorny, P., & Shepard, R. N. (1983). Distribution of visual attention over space. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 380-393.
- Polich, J. (1986). Hemispheric processing of multi-element displays. *Acta Psychologica*, 61, 137-151.
- Posner, M. I. (1978). *Chronometric explorations of mind*. Hillsdale, NJ: Erlbaum.
- Prinzmetal, W., & Millis-Wright, M. (1984). Cognitive and linguistic factors affect visual feature integration. *Cognitive Psychology*, 16, 305-340.
- Quinlan, P. T., & Humphreys, G. W. (1987). Visual search for targets defined by combinations of color, shape, and size: An examination of the task constraints on feature and conjunction searches. *Perception and Psychophysics*, 41, 455-472.
- Rabbitt, P. M. A., (1967). Learning to ignore irrelevant information. *American Journal of Psychology*, 80, 1-13.
- Rumelhart, D. E. (1970). A multicomponent theory of the perception of briefly exposed visual displays. *Journal of Mathematical Psychology*, 7, 191-218.
- Sagi, D. (1988). The combination of spatial frequency and orientation is effortlessly perceived. *Perception and Psychophysics*, 43, 601-603.
- Sagi, D., & Julesz, B. (1984). Detection versus discrimination of visual orientation. *Perception*, 13, 619-628.
- Sagi, D., & Julesz, B. (1987). Short-range limitation on detection of feature differences. *Spatial Vision*, 2, 39-49.
- Schneider, W., & Fisk, A. D. (1982). Degree of consistent training: Improvements in search performance and automatic process development. *Perception and Psychophysics*, 31, 160-168.
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84, 1-66.
- Shepard, R. N., & Podgorny, P. (1986). Spatial factors in visual attention: A reply to Crassini. *Journal of Experimental Psychology: Human Perception and Performance*, 12, 383-387.
- Shibuya, H., & Bundesen, C. (1988). Visual selection from multielement displays: Measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 591-600.
- Shiffrin, R. M., Dumais, S. T., & Schneider, W. (1981). Characteristics of automatism. In J. B. Long & A. D. Baddeley (Eds.), *Attention and performance IX* (pp. 223-238). Hillsdale, NJ: Erlbaum.
- Shiffrin, R. M., & Gardner, G. T. (1972). Visual processing capacity and attentional control. *Journal of Experimental Psychology*, 93, 72-83.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84, 127-190.
- Sorkin, R. D., Pohlmann, L. D., & Gilliom, J. (1973). Simultaneous two-channel signal detection: III. 630 and 1400 Hz signals. *Journal of the Acoustical Society of America*, 53, 1045-1051.
- Sperling, G. (1967). Successive approximations to a model for short-term memory. In A. F. Sanders (Ed.), *Attention and performance I* (pp. 285-292). Amsterdam: North-Holland.
- Steinman, S. B. (1987). Serial and parallel search in pattern vision? *Perception*, 16, 389-398.
- Sutherland, N. S. (1968). Outlines of a theory of visual pattern recognition in animals and man. *Proceedings of The Royal Society, B171*, 297-317.
- Taylor, D. A. (1978). Identification and categorization of letters and digits. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 423-439.
- Townsend, J. T. (1971). Theoretical analysis of an alphabetic confusion matrix. *Perception and Psychophysics*, 9, 40-50.
- Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 194-214.
- 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.
- Treisman, A., & Paterson, R. (1984). Emergent features, attention, and object perception. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 12-31.
- Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, 14, 107-141.
- Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, 114, 285-310.
- Treisman, A., Sykes, M., & Gelade, G. (1977). Selective attention and stimulus integration. In S. Dornic (Ed.), *Attention and performance VI* (pp. 333-361). Hillsdale, NJ: Erlbaum.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 327-352.
- Ullman, S. (1984). Visual routines. *Cognition*, 18, 97-159.
- von Wright, J. M. (1970). On selection in visual immediate memory. In A. F. Sanders (Ed.), *Attention and performance III* (pp. 280-292). Amsterdam: North-Holland.
- Warren, R. M., & Warren, R. P. (1968). *Helmholtz on perception: Its physiology and development*. New York: Wiley.
- Wolfe, J. M., Cave, K. R., & Franzel, S. L. (in press). A modified feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*.

Received April 21, 1988

Revision received November 14, 1988

Accepted November 29, 1988 ■