

## Discriminability and dimensionality effects in visual search for featural conjunctions: A functional pop-out

STANISLAS DEHAENE

Laboratoire de Sciences Cognitives et Psycholinguistique, Paris, France

Treisman and Gelade's (1980) feature-integration theory of attention states that a scene must be serially scanned before the objects in it can be accurately perceived. Is serial scanning compatible with the speed observed in the perception of real-world scenes? Most real scenes consist of many more dimensions (color, size, shape, depth, etc.) than those generally found in search paradigms. Furthermore, real objects differ from each other along many of these dimensions. The present experiment assessed the influence of the total number of dimensions and target/distractor discriminability (the number of dimensions that suffice to separate a target from distractors) on search times for a conjunction of features. Search was always found to be serial. However, for the most discriminable targets, search rate was so fast that search times were in the same range as pop-out detection times. Apparently, greater discriminability enables subjects to direct attention at a faster rate and at only a fraction of the items in a scene.

Treisman and Gelade (1980) have proposed a feature-integration theory of visual attention. According to this theory, attention must be focused on a given spatial location before single features can be combined into a coherent object. The theory holds that prior to focused attention, isolated features are free-moving and may combine to produce *illusory conjunctions*—combinations of features that belong to different objects (e.g., a green O composed of the "green" feature of a green I and the "shape" feature of a red O).

This model has received extensive experimental support in recent years. It has been shown that single features, such as the color green, can be detected in parallel and can yield texture segregation (in contrast to some other color). Careful experiments have determined which featural dimensions lead to such "pop-out" effects. These dimensions include color, size, shape, orientation, and curvature. Consistent with Treisman and Gelade's (1980) model, conjunctions of these elementary features cannot be detected in parallel, but require serial, attentive search. Finally, with short exposure times, illusory conjunctions have been shown to occur more often than chance alone predicts (Treisman & Schmidt, 1982). Indeed, the precise conditions of their occurrence have been determined (Treisman & Paterson, 1984; Treisman & Schmidt, 1982).

Nevertheless, the feature-integration theory encounters some intriguing theoretical problems when it is applied to the perception of real-world scenes. Consider the task of looking for a chair in a room. Humans (with normal vision) are generally quite good at this task, and are able to find a seat within a few hundreds of milliseconds (if one excludes the difficult problem of finding a *free* seat in a crowded room—a problem that is indeed very similar to Treisman's conjunctive search). Yet, like most objects, chairs are not defined by a single feature that would otherwise be absent in the room. By this I do not mean that in a given environment, there is never any feature that suffices to separate chairs from nonchairs. Rather, there is no feature that *systematically* serves as a foundation for this distinction in all environments. A chair may well be the only blue object in a given room, but unless we know a priori that we are looking for a blue chair, a color-based search strategy may lead us to spend a very long time examining blue paintings in a room in which all the chairs are, as it happens, red.

Shape may be proposed as a feature that distinguishes chairs from other objects, since shape has been shown to yield a pop-out effect in several experiments (e.g., Treisman & Gelade, 1980). However, such pop-out effects occurred for conditions in which shape was confounded with more elementary parameters such as curvature or closure. To assume that there are preattentive detectors for the shape of a chair transfers the burden of shape recognition to a lower level without solving the problem. In any case, this example contradicts Treisman's views that object perception requires attention.

In real-world scenes, then, objects such as chairs cannot be found through a single-feature search. Rather, several features must be locally combined before the object can be recognized. Thus, according to Treisman,

I thank the personnel at the Laboratoire de Sciences Cognitives et Psycholinguistique, particularly Jacques Mehler and Peter Jusczyk, for their comments on the manuscript as well as for their participation in the experiment. A. Treisman, H. Pashler, and one anonymous reviewer made helpful remarks on an earlier version. Jean-Luc Aucouturier provided technical help. Reprint requests should be addressed to Stanislas Dehaene, Laboratoire de Sciences Cognitives et Psycholinguistique, 54 Bd Raspail, 75270 Paris Cédex 06, France.



search must be serial through all locations in a scene. At a rate of say 50–100 msec per location, such a search would take an enormous amount of time, especially if totally irrelevant but nevertheless salient locations, such as a crack in the wall, have to be scanned. This scenario seems to be in contradiction with the extreme speeds found in the perception of real-world scenes (e.g., Biederman, Rabinowitz, Glass, & Stacy, 1974).

In essence, the difficulty is that attention should be (and seems to be) directed only at potentially interesting objects. However, determining what and where these objects are requires focused attention. This is not a real paradox, but rather a problem of practical efficiency similar to the "frame problem" (Dennett, 1978); that is, if attention scans all salient locations of an image, a lot of time may be lost examining things that are irrelevant to the present context.

Treisman and Gelade (1980) have proposed an additional mechanism that may accelerate scene perception. They proposed that default values for features are assigned to objects in the absence of focused attention, so that "even when attention is directed elsewhere, we are unlikely to see a blue sun in a yellow sky" (p. 98). Thus, prior to complete scanning, the viewer could have a fairly correct idea of the attributes of several objects in a scene. Unfortunately, this proposal contains a real, logical paradox. Obviously, an object must be identified before default values can be assigned because the default value that will be chosen depends on the identity of the object. But in Treisman and Gelade's theory, identification requires attention, so paradoxically the assignment of default values also requires attention. The paradox may be partially resolved by making a distinction between *seeing* the sky as blue (which requires focused attention) and *knowing* that the sky is blue (which may involve default values that are independent of attention). But again, before applying such a priori knowledge, a brief but attentive analysis of the scene would be necessary, to make sure, for example, that part of the scene is indeed the sky.

Several researchers (e.g., Biederman, Mezzanotte, & Rabinowitz, 1982) have shown that violations in the organization of real-world scenes impair the detection of objects. Their data seem to indicate that the discovery of relations between various components of a scene may precede and guide object identification (Biederman, Bickel, Teitelbaum, & Klatsky, 1988). Thus, knowledge of the habitual organization of world scenes, together with a preanalysis of the structure of the particular scene at hand, can guide attentional scanning. Nevertheless, the same problem as that described above applies to this preanalysis, although to a lesser degree; that is, according to feature-integration theory, there would still have to be a slow, serial scan.

Several other proposals can be made to enrich the initial theory. First, as certain objects acquire special relevance, the observer may become aware of new features that distinguish these objects from others more efficiently (Gibson & Gibson, 1955) and allow for an au-

tomatization of perceptual processing (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Second, it is possible that the visual system scans the world in a hierarchical fashion. The heuristic of segmenting a visual scene first into large blocks, and then into smaller details, may reduce identification times. For example, in looking for a chair, one could avoid attending to the titles of books that lie on a table.

Finally, one may take advantage of the many featural dimensions in a visual scene. Although objects in a scene are not singled out on the basis of only one feature, most of them—at least those that belong to "basic-level categories" in the sense of Rosch, Mervis, Gray, Johnson, & Boyes-Braem (1976)—are generally very distinct from each other in terms of the number of shared features. Thus, a chair presumably shares some, but very few, features with a bookshelf or a crack in the wall. Although a single feature rarely suffices to separate an object from the rest of a scene, if attention were first directed to locations that are globally highly discriminable from their background, an observer would be more likely to examine the relevant objects rapidly.

In the present experiment, the possibility that the scanning system takes advantage of global discriminability is explored. In most classical conjunctive-search experiments, only two dimensions (e.g., color and shape) are varied simultaneously, whereas all other dimensions (e.g., size, elongation, orientation, etc.) are fixed. As a consequence, the target and each one of the distractors often differ by only a single feature. This methodology stands in contrast with real-world scenes in which many dimensions vary simultaneously, and presumably any two objects differ along several of these dimensions at once. Smith (1981) has shown that in classification tasks, adults adopt very different sorting rules depending on the heterogeneity of the set of objects to be sorted; if the objects vary along few dimensions, the set is sorted according to one particular dimension, but if the set is heterogeneous (i.e., contains objects that vary in a variety of ways), "classifications that maximize within-category similarity on all varying dimensions" at once are constructed (p. 811). In a visual-search task, could the total number of dimensions used in the display similarly affect search strategies? For instance, would search be serial and self-terminating when dimensionality and target/distractor discriminability are low, but parallel if the number of features that distinguish target from distractor increases?

This question has been partially examined by Quinlan and Humphreys (1987). In Experiment 4 of their study, they examined the time taken to find a small green H under two conditions: when distractors shared, respectively, one feature and two features with the target. In both conditions, their subjects' search was serial; in agreement with Treisman and Gelade's (1980) views, target-distractor discriminability affected only search rates. However, in this experiment, discriminability was still not much greater than in a classical conjunction search; only three featural dimensions (letter shape, size, and



color) were used, and discriminability was low within the dimensions of size and shape (A vs. H). Thus, Quinlan and Humphreys may have failed to attain optimal conditions to observe a shift from serial to parallel search with increasing discriminability. The goal of the present paper is to extend Quinlan and Humphreys's results to a four-dimensional design (color, orientation, size, and shape being varied) with high discriminability within each dimension, although still under conditions of conjunctive search.

In the present experiment, reaction time (RT) was measured for the visual search of a target defined by a conjunction of features and placed among a varying number of distractors. Two experimental conditions were examined. In Condition 1, the number of features separating the target from the distractors (discriminability) was varied. In Condition 2, the primary variable was the total number of featural dimensions along which the objects varied (dimensionality). In all displays, the distractors were chosen so that no single feature was sufficient to isolate the target. Thus, according to the feature-integration theory of attention, search should always be serial (although search rates may vary as a function of target/distractor discriminability); one expects a linear dependency of RTs on display size for "present" as well as for "absent" responses, with a 2:1 ratio of the slopes reflecting a self-terminating search. If, on the other hand, target/distractor discriminability alters search strategies, patterns of RTs characteristic of parallel search (no dependency of RTs on display size for "present" responses) should be observed when many features distinguish the target from distractors.

## METHOD

### Subjects

Twenty subjects from the Laboratoire de Sciences Cognitives et Psycholinguistique participated in the experiment (the author was one of the subjects). Their ages ranged from 20 to 50 years. With the exception of the author, the subjects were not aware of the purpose of the experiment.

### Procedure

The subjects were exposed to Conditions 1 and 2 in random order on different days, with the same target for both conditions. One group of 10 subjects was told to search for a large, green, vertical segment. The other 10 subjects were told that the target was a small, red, horizontal rectangle. The subjects were asked to respond as fast as possible (but without guessing) by pressing the right-hand key if the target was present, and the left-hand key if it was not. Training was provided in the presence of the experimenter until the subjects decided they were ready to start. Each condition was divided into three parts. The subjects were allowed to take a short rest between the parts.

### Material

The experiment was run on an IBM-compatible Olivetti M-28 computer with a high-quality color screen. The computer presented the stimuli in random order and measured reaction times with a 1-msec precision.

The temporal constraints of the presentation were as follows. First, a yellow fixation cross appeared for 1 sec. The screen was then made blank for 300 msec, followed by the appearance of the display. The response was then recorded, with no upper boundary on response latencies. The display remained on for 900 msec longer, followed by 1 sec of blank screen, and again the fixation cross.

The display consisted of 1, 5, 16, or 36 items. Thirty-six locations for the items were chosen in a 144-cm<sup>2</sup> square, which subtended an angle of 6.9° when viewed from a distance of 1 m. The locations were irregularly spaced so that no excessive alignment effects were perceptible. When fewer than 36 items were used, a corresponding subset of locations was selected at random.

The items were displayed on a dark background. They varied along as many as four different dimensions: color, orientation, shape, and size. For each of these dimensions, a binary choice of features was selected: green versus red, vertical versus horizontal, segment versus rectangle, and large (14 mm long, 4.5 mm wide) versus small (7.5 mm long, 2.5 mm wide; drawing lines were 1 mm wide). Sixteen different objects resulted from the combination of these features. One object was the target throughout, and the other objects served as distractors. Each condition was divided into three parts, the order of which was fixed.

**Condition 1.** All four dimensions were used for the distractors. This means that any possible feature was represented in the distractors. Thus, there were green as well as red distractors, vertical as well as horizontal ones, and so forth, which ensured that no single feature distinguished the target from the distractors. What was varied was target/distractor discriminability. In Part 1A (see Figure 1), discriminability was three, that is, each distractor shared only one feature with the target. In Part 1B, discriminability was one, that is, each distractor shared three out of four features with the target. Part 1C controlled for a possible source of artifact: in Part 1B, 75% of the distractors possessed a given feature of the target (e.g., color); only 25% of the distractors possessed a given feature of the target in Part 1A. Part 1C was designed to separate the effects of this percentage of target/distractor discriminability. In Part 1C, one third of the distractors were chosen as in Part 1B (three out of four features in common with the target), and the other two thirds were chosen as exact opposites of the target, that is, not a single feature in common with it (see Figure 1). Thus, the minimal target/distractor discriminability was one, as in Part 1B, but in this case 25% of distractors shared a given feature with the target, as in Part 1A.

**Condition 2.** The dimensionality of the display was varied, while target/distractor discriminability was kept as large as possible in each case. Part 2A was identical to Part 1A: four dimensions were used, with a discriminability of three. In Part 2B, only three dimensions were used, and discriminability was two (reproducing Experiment 4 of Quinlan & Humphreys, 1987). The dimension abandoned was size. Thus, the distractors were the same size as the target (large for one group of subjects, small for the other). However, the objects still differed along the dimensions of color, orientation, and shape (see Figure 1). Finally in Part 2C, only two dimensions (color and orientation) were used, and discriminability was one. This case is identical to Treisman and Gelade's (1980) classical conjunctive search.

In each display, the number of distractors of each type was equalized as much as possible, except in Part 1C, in which the percentages of each type were approximated. When display size did not allow for a perfect equalization of distractor types (in particular at display size = 1), the type of supernumerary distractor(s) was chosen at random, and was different for each subject. In each part, each display size (1, 5, 16, and 36) was presented 14 times (seven times with the target absent and seven times with the target present). This resulted in 56 trials per part, the first 8 of which were consid-



PART	TARGET	DISTRACTORS
1A+2A	I	□ H □ □
	□	I □ H I
1B	I	H □ I I
	□	□ H □ □
1C	I	H □ I I □
	□	□ H □ □ I
		33% 67%
2B	I	□ H □
	□	I □ H
2C	I	H I
	□	□ □

Figure 1. Design of each experimental part in Conditions 1 and 2. The subjects' task was to search for a given target in a display that contained a variable number of corresponding distractors. Identical stimuli were used in Parts 1A and 2A. Unless otherwise noted, there was an equal number of distractors of each type. Black stands for green and grey for red.

ered training and were not included in later analyses. Each part lasted about 5 min.

## RESULTS

Reaction times associated with erroneous responses were not analyzed. The overall error rate was 4.2%. Errors tended to parallel RTs in all conditions, except that the error rates were higher when the target was present than when it was absent. This superiority of false negatives corroborated reports by some subjects who said they answered "absent" after searching for some time, even though they had not searched through the whole display.

Figure 2 shows RTs averaged across subjects in each condition. As shown in Table 1, a regression analysis of RTs with the number of items displayed (1, 5, 16, or 36) gave correlation coefficients that were significant in all cases. Data from Parts 1A and 2A were pooled since they corresponded to identical paradigms. The slopes of regression for absent and present responses in this condition were extremely small (5.1 and 3.0 msec, respectively, per

item when the target was the large, green, vertical segment; 8.5 and 5.2 msec, respectively, per item with the small, red, horizontal rectangle).

A 2 (target type)  $\times$  2 (presence/absence of target)  $\times$  5 (experimental part: 1A+2A, 1B, 1C, 2B, or 2C) ANOVA was performed on individual subjects' slopes obtained in a regression with display size. Target type had a small influence [ $F(1,18) = 4.69, p < .05$ ]; slopes were higher during the search for the small, red, horizontal rectangle than during the search for the large, green, vertical segment. However, target type did not interact with other factors and did not affect the global shape of the results. The experimental part drastically influenced search slopes [ $F(4,72) = 89.3, p < .001$ ]. In Condition 1, the ordering of parts in increasing order according to the overall slope (measuring the difficulty of the task) was 1A, 1C, 1B. In Condition 2 it was 2A, 2B, 2C. Finally, slopes were higher for absent than for present responses [ $F(1,18) = 84.3, p < .001$ ], and presence/absence of target interacted with experimental part [ $F(4,72) = 35.3, p < .001$ ].

Planned comparisons were performed in order to assess the presence of a 2:1 ratio of absent-to-present slopes. As predicted by Treisman and Gelade's (1980) theory, a global comparison with vector (1, -2) was not significant [ $F(1,18) = 0.004$ ], whereas a 1:1 ratio, tested with

Table 1  
Results of Regression Analyses of Reaction Time on Display Size

		Percent Linearity	Slope	Intercept	Slope Ratio
Target: large, green, vertical segment					
1A	Absent	99.0	5.9‡	520	1.97
	Present	92.0	3.0*	458	
1B	Absent	99.3	64.7‡	600	2.23
	Present	97.6	29.0‡	539	
1C	Absent	98.7	22.7‡	484	1.99
	Present	98.3	11.4‡	472	
2A	Absent	95.0	4.3*	517	1.48
	Present	99.9	2.9‡	456	
2B	Absent	99.2	9.6‡	524	1.60
	Present	97.4	6.0‡	456	
2C	Absent	99.5	20.4‡	553	1.91
	Present	95.5	10.7*	506	
1A+2A	Absent	97.7	5.1‡	518	1.70
	Present	97.2	3.0‡	457	
Target: small, red, horizontal rectangle					
1A	Absent	99.5	9.6‡	571	1.78
	Present	97.3	5.4‡	497	
1B	Absent	99.7	72.8‡	636	2.17
	Present	98.0	33.6‡	595	
1C	Absent	99.2	28.8‡	524	1.72
	Present	100	16.7‡	486	
2A	Absent	99.6	7.5‡	560	1.50
	Present	98.1	5.0‡	485	
2B	Absent	99.6	12.0‡	563	1.85
	Present	96.2	6.5*	480	
2C	Absent	99.7	22.6‡	609	1.92
	Present	94.7	11.8*	550	
1A+2A	Absent	99.6	8.5‡	565	1.63
	Present	97.8	5.2‡	491	

\* $p < .05$ . † $p < .02$ . ‡ $p < .01$ . § $p < .001$ .



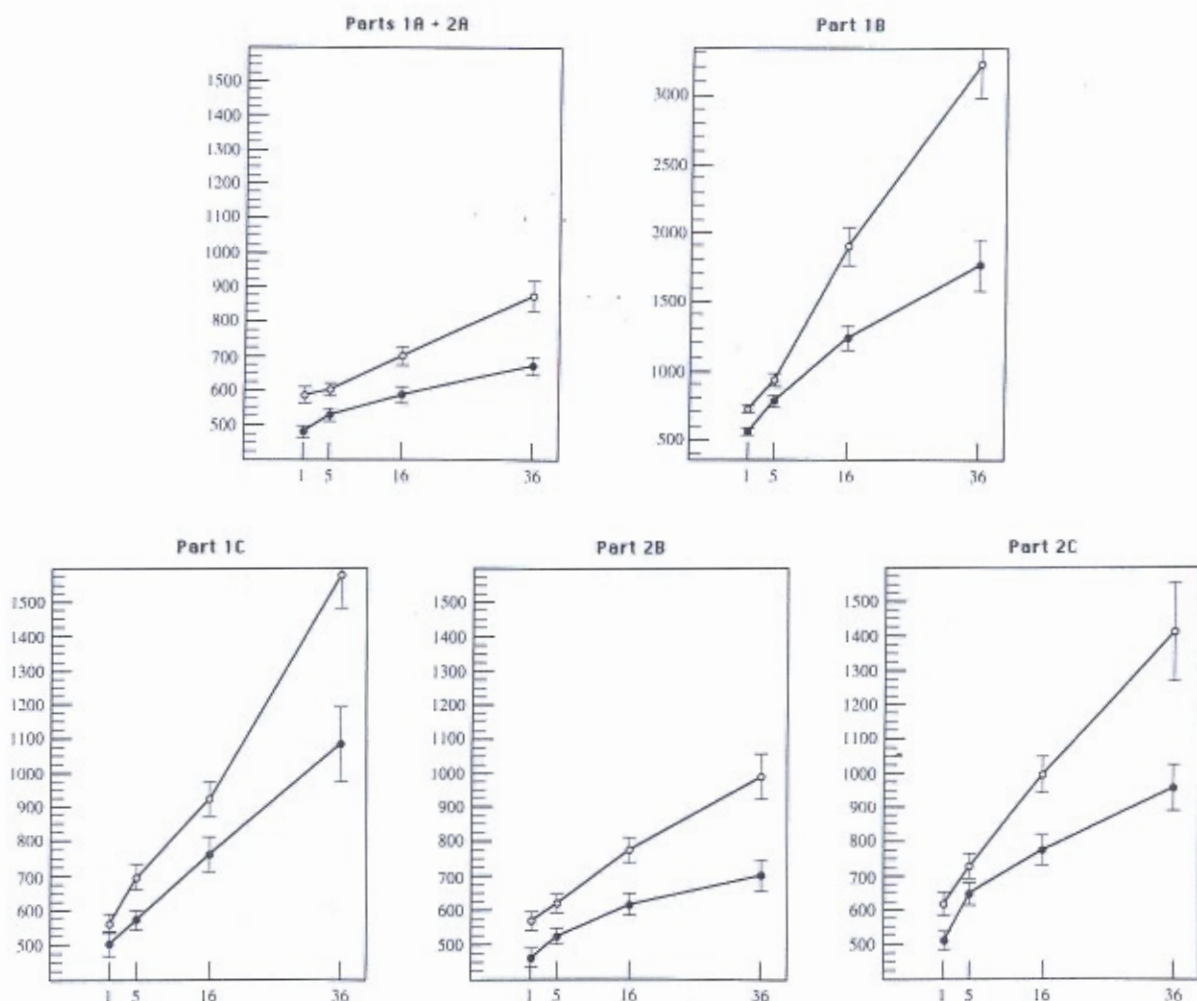


Figure 2. Reaction times (RTs) for absent (open circles) and present (closed circles) responses as a function of the number of objects displayed. Data from Parts 1A and 2A are pooled. Part 1B is drawn to a separate scale. The bars around each circle represent two standard deviations of the mean. Each point is averaged from about 120 individual RTs (about 240 RTs for Part 1A+2A).

vector (1, -1), was rejected at  $p < .001$  [ $F(1,18) = 84.0$ ]. In a condition  $\times$  condition analysis, the null hypothesis of a 2:1 ratio was accepted in all cases except condition 1A+2A [ $F(1,18) = 4.85$ ,  $p < .05$ ]. However, in the latter condition, the observed ratio was 1.64:1, and the hypothesis of a 1:1 ratio was rejected as well [ $F(1,18) = 12.8$ ,  $p < .005$ ].

To test the linearity of increases in RTs with increased display size further, individual subjects' mean RTs were submitted to a 2 (target type)  $\times$  2 (presence/absence of target)  $\times$  5 (experimental part)  $\times$  4 (display size: 1, 5, 16, or 36) ANOVA. The last three within-subject factors and all of their interactions were significant at  $p < .001$ ; there was also a significant effect of target type ( $p < .05$ ) and an interaction of target type and display size ( $p < .01$ ), all of which confirms the findings of the previous ANOVA. Specific linear, quadratic, and cubic contrasts, taking into account the irregular sampling of display size (1, 5, 16, 36), were applied separately for absent

and present responses in each condition. The linear component was always significant beyond  $p < .001$ . The quadratic component was significant at  $p < .01$  for all "present" responses except in Part 1C, in which it held for "absent" responses. Finally, the cubic component was significant at  $p < .05$  for "absent" responses in Parts 1B and 1C, and for "present" responses in Parts 1A+2A and 2C. These deviations from linearity are perceptible in Figure 2.

Some of these deviations from linearity may stem from a variation in the ratio of absent-to-present slopes, which would be 1:1 for small display sizes but 2:1 for large display sizes (Pashler, 1987a). To test this hypothesis, slopes of increase of RTs with display size were computed for each subject separately for the intervals 1-5, 5-16, and 16-36. These data were submitted to a 2 (target type)  $\times$  2 (presence/absence of target)  $\times$  5 (experimental part)  $\times$  3 (interval: 1-5, 5-16, or 16-36) ANOVA. The last three within-subject factors and their interactions were sig-



nificant at  $p < .01$ . In focused analyses, the influence of interval on slopes was significant for "present" responses [ $F(2,36) = 16.3, p < .001$ ], but not for "absent" responses [ $F(2,36) = 2.60, n.s.$ ]. In condition  $\times$  condition analyses, the effect appeared significant at  $p < .05$  for all "present" responses except Part 1C, and for "absent" responses in Parts 1B and 1C (these parts were precisely those in which a significant quadratic and/or cubic component had previously been found). Planned comparisons were performed to assess the null hypothesis of a 2:1 ratio of absent-to-present slopes in each of the 1-5, 5-16, and 16-36 intervals. The 2:1 ratio was rejected in the 1-5 interval for Parts 1A+2A, 1B, 2B, and 2C; in the 5-16 interval for Part 1C; and in the 16-36 interval for Part 1B. In all but two of these cases, a 1:1 ratio was an acceptable hypothesis. The two exceptions were Part 1A+2A (interval 1-5), in which the observed ratio (0.310:1) was significantly lower than 1:1, and Part 1B (interval 16-36), in which the observed ratio (2.56:1) was significantly larger than 2:1.

### DISCUSSION

To ascertain whether search was serial and self-terminating in all conditions, as Treisman and Gelade's (1980) model predicted, I shall first apply several criteria of seriality to the data, and will then discuss the objections raised by Pashler (1987a).

A first hint that processing was serial is that response time increased with increased display size for both "absent" and "present" responses. This effect was significant for all parts in both conditions, with the rate of linearity around 95%. Had there been a pure pop-out effect, it would have yielded a flat curve for "present" responses as a function of display size. A second, more stringent criterion for serial, self-terminating search is the 2:1 ratio of absent-to-present slopes. This was observed in all conditions except Part 1A+2A. Even in this case, however, the observed ratio of 1.64:1 was incompatible with a 1:1 ratio, which reflects either exhaustive serial search or parallel search. Furthermore, since that ratio was less than 2:1, a weaker form of pop-out, in which RT for "present" responses would increase as a function of display size (albeit much more slowly than for "absent" responses), can also be rejected.

A closer analysis of the data reveals that the 2:1 ratio holds only for large display sizes, not for smaller ones. This finding confirms the results of Pashler (1987a; see also Houck & Hoffman, 1986), who noted that slope ratios approach 1 when fewer than eight items are displayed. Pashler interpreted this finding as evidence against serial, self-terminating search. He suggested that search was parallel within clumps of items, and serial self-terminating from one clump to the next. Pashler's data are definitely not questionable, since the effect can be observed *a posteriori* in earlier data and is reproduced in the present study. However, I would like to argue that his observations do not necessarily undermine the hypothesis of serial processing.

First, Pashler's (1987a) basic observation should be qualified, since at small display sizes, the ratio of "absent" to "present" slopes is not always 1:1. In the present study, the ratio in Part 1A+2A was significantly smaller than 1 at display sizes between one and five. Ratios smaller than 1 are also evident in Pashler's Experiments 2B, 3A, 3B, and 4, with display size between two and four. In fact, in this interval, RTs for "absent" responses seem to be almost independent of the number of items presented. Pashler's data support the hypothesis of a continuous variation of slope ratios with increased display size, from a value near 0 to a 2:1 ratio. This is also clear in the results of Houck & Hoffman's (1986) Experiment 3.

Hence, it is possible that two different effects are superimposed in the data: (1) the effect of the classical serial, self-terminating search, reflected in the 2:1 ratio; and (2) a spurious effect that should appear only when few items are presented. Where would this additional effect stem from? In predicting a linear relationship between search times and the number of items, several approximations are made. Scanning rate is usually identified with the mean time for comparison of an item with the target. Yet subjects must also locate the items, shift their attention, keep track of previously scanned items, and make sure that all items have been scanned. With small displays, it is possible that subjects first scan all the items presented, and if the target has not been found, then spend time checking that they have not missed any item. This should selectively slow down target-absent responses at small display sizes, and thus reduce the target-absent slopes. At large display sizes, this double-check strategy should be abandoned or even reversed (large error rates for target-present trials suggest that at large display sizes, subjects sometimes respond "absent" before having scanned all the items).

If this analysis is correct, then ratios smaller than 2:1 stem from spurious factors, and serial self-terminating search is the rule. The results of Pashler's (1987a) Experiment 4 may be reinterpreted within this framework. In his experiment, subjects had to search a display of two, four, or eight items for a given target. Displays could contain zero, one, or two occurrences of the target. In addition to a 0.81:1 ratio of the "absent" to the "single target present" slopes, Pashler found that there was an advantage of double-target displays over single-target ones. This finding is incompatible with an exhaustive serial-search paradigm. However, the ratio of the double-target slopes to the single-target slopes was very near 2:3. This is exactly the ratio predicted by serial self-terminating search: single-target RTs should be proportional to half the number of items, whereas double-target RTs should be proportional to one third of this number. Thus, these data are perfectly compatible with serial search, if one accepts that the ratio of "absent" to "single target present" slopes may deviate from a 2:1 ratio for small display sizes.

Whether the time to search through few items is adequately predicted by the double-check hypothesis outlined



above, or whether search through few items is parallel, remains an open question. Until new experiments separate these two accounts, the conservative hypothesis of a serial, self-terminating search remains viable even if a 2:1 ratio of "absent" to "present" slopes is not observed for small display sizes. The conclusion of the present study is that processing was serial, self-terminating in all conditions, as Treisman and Gelade's (1980) theory predicts. This result reproduces and extends the findings of Quinlan and Humphreys (1987). Even with very high target-distractor discriminability, search remains serial, provided that no single feature suffices to indicate the target.

At first glance, the slopes of Parts 1A and 2A of the present experiment appear extremely small for serial search. With the large target, search rate averaged around 6 msec per item, a figure that has been classified as typical of parallel search by Treisman and Souther (1985; e.g., Table 2, p. 301, medium gap size). This indicates the difficulty in distinguishing real serial search with a small rate from cases in which 2:1 slope ratios are obtained by chance. In Treisman and Souther's work, there was a convergence of results from several experiments that suggested parallel search. In the present study, in contrast, continuity in the results of Parts 2A, 2B, and 2C, as well as in the slope ratios and linearity coefficients, support the hypothesis of a serial search in Parts 1A and 2A. Nevertheless, the speed of search makes it functionally equivalent to a pop-out effect.

Treisman and Souther (1985) and Treisman and Gormican (1988) have explained rates of as little as 13 msec by assuming that subjects scan small groups of items instead of focusing on each item individually. However, this explanation may be dismissed here; post hoc analyses revealed that in nearly all displays of the type in Parts 1A and 2A, each of four distractor types was present in the immediate neighborhood of the target. Thus, attending to small, local groups of items did not (in most cases) reduce our subjects' conjunctive-search task to a local, single-feature search, and could not help the subjects reduce their search rates.

The very fast scanning observed in Parts 1A and 2A can be explained by hypothesizing that the subjects scanned only some of the items displayed. Indeed, most subjects reported that they paid attention only to items of the correct color in the display. Correcting for the percentages of items that shared the color of the target, we may then get estimates of real search rates. Raw and corrected rates are shown in Table 2. The corrected figures, all above 20 msec per item, are more compatible with the literature. It is worth noting that with the large, green, vertical segment as the target, corrected rates for Parts 1B and 1C become comparable. Thus, if we divide the number of distractors very similar to the target by 3, we also divide the slopes of RTs by the same factor, which strengthens the hypothesis that only the distractors similar to the target were scanned. However, with the other target, the small, red, horizontal rectangle, scanning rates for Parts 1B and 1C still differ after the correction. Most

Table 2  
Mean Scanning Rate (msec/item) in Each Experimental Condition

Condition	1A+2A	1B	1C	2B	2C
Target: large, green, vertical segment					
Raw rate	6	61	23	11	21
Corrected rate	22	81	92	33	42
Target: small, red, horizontal rectangle					
Raw rate	9	70	31	12	23
Corrected rate	38	93	124	37	46

Note—Raw rate is the mean of the slope for "absent" responses and two times the slope for "present" responses. Corrected rate is computed assuming that the subjects scanned only items that shared one particular feature with the target (e.g., color).

likely, when the target was small and red, the subjects could not help rapidly scanning large or green distractors. Similar search asymmetries have already been reported in the literature (Treisman & Gormican, 1988; Treisman & Souther, 1985).

The conclusion that subjects scan only some of the items in the display is not novel. Egeth, Virzi, and Garbart (1984) showed that partial scanning occurs even in a classical conjunction search. In their study, the target was a red O placed among red Ns and black Os. When the subjects were told to attend only to the red items, the number of black Os did not influence response times, but the number of red Ns did. Schematically, the converse held when the subjects were told to attend only to the Os. Finally, in the absence of attentional instructions, the subjects' response times were very similar to the condition in which only red items were scanned. The slopes of response times as a function of the number of distractors should have been twice as steep if subjects in the "no instructions" group really scanned all the items in the display. The authors concluded that subjects can restrict their search to a subset of the stimuli.

The idea that a preattentive, parallel selection of potential candidates precedes serial examination has been proposed by Hoffman (1978, 1979) as a refinement of Neisser's (1967) hypothesis. According to Hoffman, in an initial parallel stage, the subject estimates the similarity of each item with the target(s), selecting or rejecting the item as a candidate for the second stage of focused attention. This view is consistent with our data and with the experiments of Egeth, Virzi, and Garbart (1984), provided that one hypothesis is relaxed. Hoffman (1978) states that the initial parallel stage corresponds to the "identification of each display element" (p. 9). This finding is incompatible with Treisman and Gelade's (1980) experiments, which showed that conjoining of features, a prerequisite of identification, requires serial processing. In the light of these results, it may be better to say that the first stage filters out, in parallel, the candidates for serial search.

The present results shed some light on the rules of this parallel filter. Two such rules may be opposed a priori: either the selection of items depends on the presence or absence of a single feature, or the global similarity of each



item with the target is measured along several dimensions at once. The second hypothesis is rejected by the results of Part 1A+2A. In that condition, target-distractor discriminability was very high, yet the search was not parallel. This indicates that the parallel stage does not use global target-distractor similarity as a basis for filtering. Rather, the filtering must be based on a single dimension.

Do the present data also constrain the second processing stage of the model, namely, the serial comparison of each candidate item with the target? Even after correction for the number of scanned candidates, large differences between the scanning rates in the various conditions can be noted. Part 2C (classical conjunction) appears to have been more difficult than Part 2B (conjunction with discriminability = 2), itself more difficult than Part 2A (conjunction with discriminability = 3). Thus, target/distractor discriminability still affects the corrected rates. The findings of Treisman and Gelade (1980) and Treisman and Gormican (1988), that target/distractor discriminability within a given dimension (e.g., size) affects search rates but not search strategies, is thus extended to multiple dimensions.

It is not possible to account quantitatively for these differences in scanning rates with a model of attentional scanning. Rather, a model of the comparison of an individual item with the target is needed. Such a model should explain, for instance, the subtle asymmetry patterns between the search for a large target among small items and vice versa (see Treisman & Souther, 1985). Since the dimensions used in the present study were chosen quite arbitrarily, I do not intend to develop such a model (but see Treisman & Gormican, 1988; Tversky, 1977). Let me just mention how complex such modeling is likely to be. In the present experiment, there is a large difference in the corrected scanning rates for Parts 1C and 2C, yet the distractors in Part 2C were also used in Part 1C. Two explanations can be proposed. First, it may be that the additional distractors used in Part 1C were less discriminable from the target than those used in Part 2C. This would imply that some dimensions are easier to discriminate than others. Each featural dimension would then have its own peculiarities, which a model of comparison should specify (this approach is exemplified by Treisman & Gormican, 1988). The second possible basis for the differing rates in Parts 1C and 2C is that only two featural dimensions were varied in Part 2C, whereas all four dimensions were varied in Part 1C. The subjects might have been able to use this restriction in the number of varied dimensions to accelerate their responses by a factor of two or three. The similarity of an item with the target would then vary with context; a difference of one feature would be more discriminable in a display in which two dimensions are varied than in a display with four varied featural dimensions. This possibility deserves exploration.

In summary, the model most compatible with the current data is a two-stage model along the lines of that of Hoffman (1978, 1979). The first stage consists in a

parallel preselection of candidates with a single-feature filter. All the items selected enter the second stage of serial, self-terminating examination. During the latter stage, each item is compared to the search target. The comparison latency increases as a function of the number of features shared with the target but might also be modulated by the total number of dimensions used in the display.

Such a model is fully compatible with the theory of Treisman and Gelade (1980) because it postulates that only single-feature searches can be carried out in parallel. This model also makes sense in the light of Julesz's experiments. Sagi and Julesz (1985) showed that the position and number of features that pop out of a display can be determined in parallel, but that the identity of these features cannot be reported with presentations of very short duration. The number and position of the candidates must be computed in parallel in the preattentive stage of Hoffman's (1978, 1979) model in order to drive the next stage of focused attention.<sup>1</sup>

Hoffman's (1978, 1979) two-stage model has recently received more support from experiments by Pashler (1987b). Pashler showed that search times increased when target-confusable distractors were added, even though the total number of distractors remained constant. This finding confirms the results of Egeth, Virzi, and Garbart (1984), and is similar to the difference between Parts 1B and 1C of the present study. In Pashler's (1987b) Experiment 2, which was a two-target search task, RTs to one target were slowed to a similar degree when distractors similar to that (present) target, or to the other target (not itself present) were added to the display. Pashler concluded that *decision noise*, namely the difficulty of reducing false alarms when target-confusable distractors are present, was responsible for the effect of target-distractor similarity. This finding is consistent with the two-stage model, which predicts that more candidates will have to be serially scanned when confusability increases.

The two-stage model is also compatible with recent reports of fast, quasiparallel, search for conjunction targets (McLeod, Driver, & Crisp, 1988; Nakayama & Silverman, 1986; Steinman, 1987; see also Sagi, 1988). This seems to happen only with particular featural dimensions, which allow for a perceptual segregation of all items that share one of the target's features (Treisman, 1988). For example, a conjunction of binocular disparity and color can be detected in parallel because disparity preattentively defines two planes of items that can be examined separately (Nakayama & Silverman, 1986). Rapid search for conjunctions is compatible with Treisman's model if one accepts that search can sometimes be totally restricted to a subset of items (Treisman, 1988).

Finally, a two-stage model partially resolves the difficulty mentioned in the introduction, that is, can we reconcile the slow processing predicted by serial scanning with the extreme speed observed in the processing of real-world scenes? The present data show that Treisman's views are basically correct; when a target for search



is defined by a conjunction of features, some serial scanning is necessary, even if dimensionality and target/distractor discriminability are large. However, the rate of this search, inherently slow, can be considerably improved by a parallel preselection of candidates. Although still serial, the search then becomes functionally equivalent to a parallel pop-out, which enables subjects to search through a display of 36 items in less than 700 msec.

## REFERENCES

- BIEDERMAN, I., BLICKLE, T. W., TEITELBAUM, R. C., & KLATSKY, G. J. (1988). Object search in nonscene displays. *Journal of Experimental Psychology: Learning, Memory & Cognition*, **14**, 456-467.
- BIEDERMAN, I., MEZZANOTTE, R. J., & RABINOWITZ, J. C. (1982). Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, **14**, 143-177.
- BIEDERMAN, I., RABINOWITZ, J. C., GLASS, A. L., & STACY, E. W., JR. (1974). On the information extracted from a glance at a scene. *Journal of Experimental Psychology*, **103**, 597-600.
- DENNETT, D. (1978). *Brainstorms*. Cambridge, MA: MIT Press.
- EGETH, H. E., VIRZI, R. A., & GARBART, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 32-39.
- GIBSON, J. J., & GIBSON, E. J. (1955). Perceptual learning: Differentiation or enrichment? *Psychological Review*, **62**, 32-41.
- HOFFMAN, J. E. (1978). Search through a sequentially presented visual display. *Perception & Psychophysics*, **23**, 1-11.
- HOFFMAN, J. E. (1979). A two-stage model of visual search. *Perception & Psychophysics*, **25**, 319-327.
- HOUCK, M. R., & HOFFMAN, J. E. (1986). Conjunction of color and form without attention: Evidence from an orientation-contingent color aftereffect. *Journal of Experimental Psychology: Human Perception & Performance*, **12**, 186-199.
- MCLEOD, P., DRIVER, J., & CRISP, J. (1988). Visual search for a conjunction of movement and form is parallel. *Nature*, **332**, 154-155.
- NAKAYAMA, K., & SILVERMAN, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, **320**, 264-265.
- NEISSER, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.
- PASHLER, H. (1987a). Detecting conjunctions of color and form: Reassessing the serial search hypothesis. *Perception & Psychophysics*, **41**, 191-201.
- PASHLER, H. (1987b). Target-distractor discriminability in visual search. *Perception & Psychophysics*, **41**, 285-292.
- 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 & Psychophysics*, **41**, 455-472.
- ROSCH, E., MERVIS, C. B., GRAY, W., JOHNSON, D., & BOYES-BRAEM, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, **7**, 573-605.
- SAGI, D. (1988). The combination of spatial frequency and orientation is effortlessly perceived. *Perception & Psychophysics*, **43**, 601-603.
- SAGI, D., & JULESZ, B. (1985). "Where" and "what" in vision. *Science*, **228**, 1217-1219.
- SCHNEIDER, W., & SHIFFRIN, R. M. (1977). Controlled and automatic human information processing: 1. Detection, search, and attention. *Psychological Review*, **84**, 1-66.
- SHIFFRIN, R. M., & SCHNEIDER, W. (1977). Controlled and automatic human information processing: 2. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, **84**, 127-190.
- SMITH, L. B. (1981). Importance of the overall similarity of objects for adults' and children's classifications. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 811-824.
- STEINMAN, S. B. (1987). Serial and parallel search in pattern vision? *Perception*, **16**, 389-398.
- 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 & 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.
- TVERSKY, A. (1977). Features of similarity. *Psychological Review*, **84**, 327-352.

## NOTE

1. In Sagi and Julesz's (1985) experiments, targets "automatically" popped out of the displays. In contrast, I propose that in visual-search experiments, subjects can voluntarily restrict their search to a prespecified subset of the display. Sagi and Julesz's findings should be replicated in the context of a voluntary restriction of visual search, in order to reconcile them fully with Treisman's model. In such a replication, a display would be briefly flashed and the subjects would then be asked to report either the number, the location, or the identity of the targets. The very same display might be used several times with different criteria for targets (e.g., color or orientation). I predict that whatever feature separates targets from nontargets, the number and location of targets will be determined in parallel, whereas determining the targets' full identity will require longer viewing times. This would show that the hypothesized stage of parallel preselection works indifferently in two modes: an automatic mode in which it signals all items that significantly differ from background, and a voluntary mode in which it selects all items that possess a predefined target feature.

(Manuscript received April 11, 1988;  
revision accepted for publication December 19, 1988.)