Perception, 1992, volume 21, pages 765-777

Visual search: detection, identification, and localization

Marc Green
Computer Studies Programme, Trent University, Peterborough, Ontario K9J 7B8, Canada
Received 22 April 1991, in revised form 12 February 1992

Abstract. In two studies, observers searched for a single oblique target in a field of vertical distractors. In one experiment, target detection and identification (left versus right tilt) were compared. In another experiment, detection and localization were compared for the left versus the right half of the display. Performance on all three tasks was virtually identical: if a target could be detected, it could also be identified and localized. A review of previous studies generally supports the conclusion that performance on the three tasks is similar. This argues against current search theories, which rest heavily on data showing differences in identification and localization.

1 Introduction

Most scientific theories attempt to explain the world by decomposing phenomena into independent mechanisms, channels, or modules. This approach has naturally been applied to perception, where theories typically suggest that retinal image information is decomposed into separate feature modules to simplify processing. The divide-and-conquer strategy is assumed in several disciplines, including physiology (eg Barlow 1981, 1986; Cowey 1979, 1981; De Yoe and Van Essen 1988; Hubel and Livingstone 1985, 1987; Livingstone and Hubel 1987, 1988) and computational vision (eg Ballard 1984, 1986; Ballard et al 1983). In psychology, feature modules are used to explain visual search data (eg Treisman and Gelade 1980).

There are two main classes of search theories: feature-integration and 'interrupt' theories. Feature-integration theory (Treisman and Gelade 1980) suggests that retinal images are analyzed into separate features, each represented in a different 'map'. The maps code feature presence but not location, so individual occurrences of the feature sum spatially to create a 'pooled activity'. Location is held in a separate 'master map of locations' that has links into the feature maps. To combine individual features from different maps the observer must focus attention on a place in the master map of locations. This permits features linked to that point to be localized and 'glued' together. Features are spatially free-floating until tied to location and to each other through the master map.

Details about these feature maps are sketchy. It is unclear how the term 'map' is to be interpreted. For example, Treisman and Sato (1990) draw feature maps as two-dimensional spaces. However, the features themselves are seemingly represented without topographic information. The notion of pooled activity implies that map signals are not localized. Further, the theory presumes that features and their locations are represented independently until attention focuses on the master map (Treisman and Gelade 1980; Treisman and Schmidt 1982).

Three empirical findings support feature-integration theory. First, search is serial for targets defined by conjunctions of two or more features. This occurs because features must be glued together before the conjunction can be detected. Observers must therefore move focal attention from place to place in the master map. Second, illusory conjunctions—the miscombination of features in different spatial locations—occur independently of the spatial separation between the features. This shows that features float freely with no topographic constraints. Third, observers can sometimes identify but not localize targets. Treisman and Gelade (1980) state that "Locating a feature would ... be a separate operation from identifying it" (page 100) while Treisman and Gormican (1988) say that "When attention is divided over the entire field, only the presence of a unique feature can be detected, not its location" (page 18).

Feature-integration theory, however, produces an anomalous prediction—disembodied features. If different features of an object are represented in different maps, why do we not perceive disembodied features, eg color without shape? Treisman and Gormican (1988) suggest that features can be perceived and identified only after focal attention has combined them into an object. Attention apparently not only glues features, but also makes them perceptible.

In sum, the theory implies that search consists of three ordered processes: (i) detect: register pooled activity in a feature map; (ii) localize: use attention to link the activity to a location in the master map; and (iii) identify: perceive an object after all features are glued together. The order suggests that some observers might perform some tasks without the others. Detection can occur without localization or identification because it does not require attention. In fact, the theory of pooled activity rests on the assumption that spatial location is unavailable without focal attention. Localization, then, cannot occur until the observer applies attention to the master map. Last, after features are localized, they combine into identifiable objects.

'Interrupt' theories are the main competitors to feature integration. Several studies (Atkinson and Braddick 1989; Johnston and Pashler 1990; Sagi and Julesz 1985) claim that observers could localize but not identify targets. This is the reverse of the Treisman and Gelade (1980) result.

One interpretation of this finding (Nothdurft 1985; Sagi and Julesz 1987) is that in preattentive vision observers initially detect a 'difference signal' where there is a break in a feature gradient. But the discontinuity itself conveys no identity information. The source of the discontinuity cannot be identified until the observer moves focal attention to the location of the signal.

Johnson and Pashler (1990) offer a similar theory. Like Sagi and Julesz (1987), they found that localization could be better than identification and suggest that observers detect a featureless signal which cannot be identified without attention. Johnston and Pashler differ, however, by saying that the signal is an interrupt from a particular feature rather than from a texture discontinuity. The last two theories are less detailed than feature integration. Neither is specific about the mechanism for gluing features together, so it is unclear whether either is a feature-integration theory in the strict sense. However, both still insist that (i) feature identification requires attention and (ii) preattentive and attentive search represent distinct, sequential processing stages.

1.1 *Rationale for the present study*

Search theories are grounded in data showing the relationships between detection, localization, and identification. Feature-integration theory uses the presumed superiority of identification over localization as proof that observers detect pooled activity in nontopographic maps. Interrupt theories use the presumed superiority of localization over identification to demonstrate the existence of featureless signals in topographic representations.

The published data conflict. Moreover, none of the previous studies (see table 1, which will be discussed later) is definitive. Most authors equate detection with localization and do not measure them separately. Further, most previous search studies are difficult to interpret, both because the data are inconclusive and because they contain possible artifacts (see below and Johnston and Pashler 1990).

The purpose of the present study was to obtain conclusive data by using rigorous methodology, a two-alternative forced-choice procedure, and by employing stimuli which avoid the artifacts noted by Johnston and Pashler (1990). In the present study observers searched for tilted Gabor targets on a background of 2 - 32 vertical distractors. Results showed that accuracies on the three tasks were almost identical. This argues against any search theory which postulates that detection, identification, and localization are sequential operations, ie, any theory that postulates distinct preattentive and attentive visual processing stages.

Table 1. Comparison of identification - localization studies.

Study[a]	Features	Tasks	Results
[1]	color, form (8/10)	identification/localization	identification > localization
[2]	orientation (2 - 36)	detection/localization/ discrimination	detection = localization > discrimination
[3]	orientation (36)	identification/localization	coarse localization > identification ⩾ fine localization
[4]	orientation (2 - 6)	detection	detection = identification?
[5]	color (2 - 5)	identification	detection = identification?
[6]	color, form (8)	identification/localization	localization ⩾ identification
[7]	orientation (2 - 32)	detection/localization/ identification	detection = localization = identification

[a] [1] Treisman and Gelade (1980); [2] Sagi and Julesz (1985); [3] Atkinson and Braddick (1989); [4] Folk and Egeth (1989); [5] Duncan (1989); [6] Johnston and Pashler (1990); [7] present study.

2 Experiment 1: detection versus identification

2.1 *Observers*

Of the observers used, only one (DA) knew the purpose of the experiments.

2.2 *Display*

The displays contained elements differing only in orientation: (i) targets—oriented 45° left or right, (ii) distractors—vertical, and (iii) mask items—superposition of the two targets. (Mask items containing both targets and the distractor produce equivalent effects.) In most experiments, the display items were Gaussian modulated sine waves or 'Gabor functions'. These are, in simple terms, small patches of sine wave grating which are blurred at the edges and displayed on a background of the same mean luminance (see Green 1986 for details). The Gabors had a spatial frequency of 2.0 cycles deg^{-1}, peak contrast of 50%, diameter of about 0.8 deg, and mean luminance of 65 $\mathrm{cd\,m}^{-2}$. To ensure that the results were not peculiar to Gabor functions, however, I replicated one study with line segments. As described below, lines and Gabors produce similar results.

Each display consisted of 2, 4, 8, 16, or 32 items. As shown in figure 1a, there were 36 possible positions arranged in three concentric circles with radii of 1.0, 2.0, and 3.0 deg. On each trial, the computer selected a quasi-random subset of positions. Since there were never more than 32 items, the display always contained empty positions.

2.3 *Procedure*

Data were collected by means of a two-alternative forced-choice paradigm. Each trial consisted of two test intervals separated by 1 s. After viewing the trial, the observer pushed two switches: one indicated whether the target had been in the first or second interval (detection), and the other indicated whether the target had been oriented left or right (identification).(1) The order of response was counterbalanced across blocks of trials.

Figure 2 schematically shows the sequence of events on each trial. The first interval consisted of: (i) a one-raster (nominal duration 17 ms) test field with 2, 4, 8, 16, or 32 vertical distractors, (ii) a period during which a uniform grey field with the same mean luminance as the Gabors was shown, and (iii) a five-raster (nominally 84 ms) duration field containing a mask at each corresponding test field location. The second interval was similar except that the computer chose a new set of positions. That is, if the first interval contained 8 items, then the second interval also contained 8 items—but at a new set of locations. This prevented observers from anticipating the location of items in the second interval. In one of the intervals, a target filled one of the distractor locations. There was 0.5 probability of the target being in each interval as well as an independent 0.5 probability that it would be tilted left or right.

The experimenter usually collected data in four blocks, each comprising 750 trials: 5 display sizes crossed with 3 stimulus onset asynchronies (SOAs) times 50 observations. Each graphed data point therefore represents 200 judgments. In blocks one and four, one observer's first response indicated interval (detection) and second response indicated orientation (identification). In blocks two and three the order was reversed. The second observer made responses in the opposite order. This proved necessary since the order of response had a small effect: performance was 3%-4% better when a response was first rather than second.

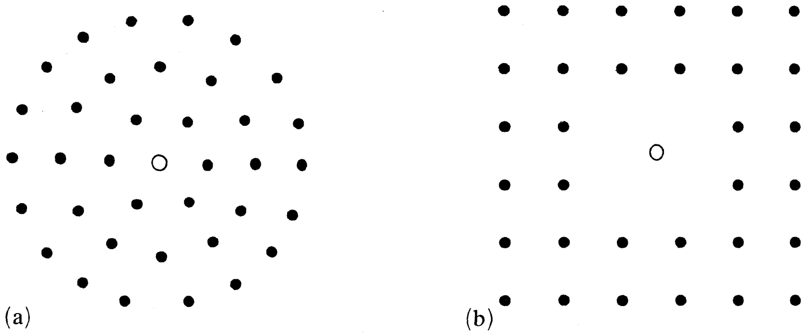

Figure 1. Schematic representation of the visual displays. Each dot represents the possible position of a target or distractor. Panel (a) shows the display used for detection-identification (experiment 1) and panel (b) shows the display used for detection-localization (experiment 2). The open circle in the middle of each display serves as a fixation mark.

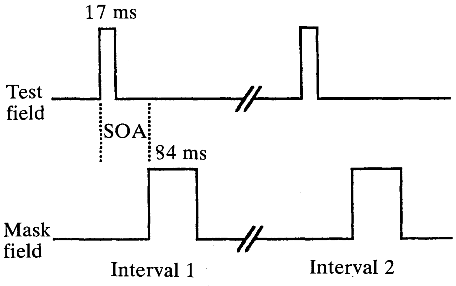

Figure 2. Schematic representation of the events on each trial. See text for details.

(1) One reviewer made the interesting observation that the temporal forced-choice task requires a *temporal* localization.

2.4 *Results*

Figure 3 shows the results for both observers. Each of the experimental variables, SOA and number of items, influenced performance. First, longer SOAs between target and mask improved performance. This is hardly surprising and has been reported previously (eg, Sagi and Julesz 1985).

The second variable was the number of display items. The purpose of this manipulation was to confirm that the paradigm had, in fact, produced parallel search. Instead of obtaining the usual flat function, however, accuracy for both detection and identification improved with increasing numbers of distractors. The effect is largest with the shortest SOA, low accuracy condition, but this is probably due to a ceiling effect at longer SOAs.

The major finding, however, is that detection and identification accuracies were almost identical. While there is a slight advantage for detection in most conditions, the difference is small and often disappears. The average detection superiority is 1.6% per data point. Roughly speaking, if the observer could detect the target, he could identify its orientation.

Visual-search experiments seldom use Gabors. Since several of these results are novel, I replicated the experiment with more conventional stimuli. The items were dark line segments with dimensions of 0.6 deg × 0.15 deg viewed against the same uniform grey background. As before, targets were 45° left/right, the distractors were vertical, and the mask items were Xs made from target lines. The results, shown in

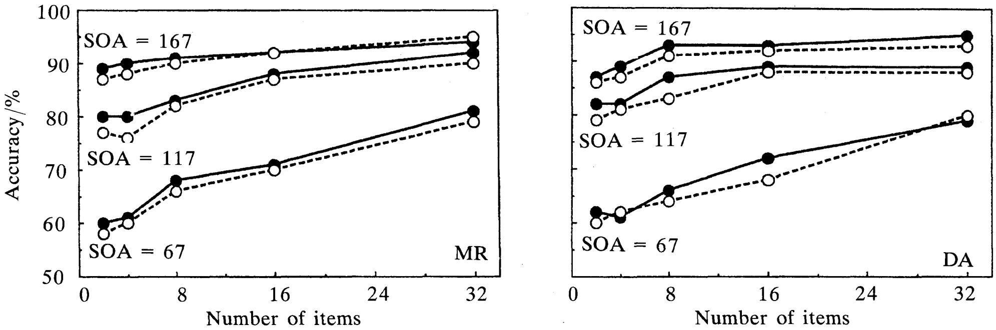

Figure 3. Detection and identification accuracies as a function of the numbr of Gabor items. SOA is the time in ms between presentation of the target and the mask. The two panels show data for different observers (MR and DA). Solid symbols show detection data, and open symbols represent identification.

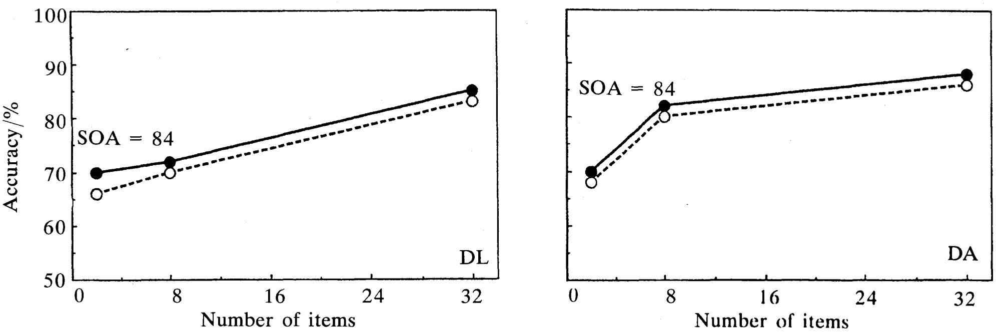

Figure 4. Detection and identification accuracies as a function of the number of line-segment items. SOA is the time in ms between presentations of the target and the mask. The two panels show data for different observers (DL and DA). Solid symbols show detection data, and open symbols represent identification.

figure 4, are similar to those obtained with Gabors. There is little difference between detection and identification, and accuracy improved with increased numbers of distractors. The average detection superiority was 2.5% per data point. Gabor results can thus be generalized to conventional stimuli.

3 Experiment 2: detection versus localization

Stimuli and procedures were similar to those used in the first experiment, except that items appeared in a square array. As shown in figure 1b, item positions formed a square with the center punched out. There were 32 possible positions, 16 to the left and 16 to the right of the fixation point. Although the number of items could vary from 2 to 32, there were always equal numbers of items on each side. The observers pushed switches to indicate detection (interval) and localization (left or right half of the display). Order of response was again counterbalanced across blocks.

3.1 *Results*

Figure 5 shows that the results of the second experiment replicated those of the first; accuracy again improved with increasing target-mask SOA and increased numbers of distractors. More importantly, detection and localization accuracies were almost identical. There was a slight superiority in detection for one observer (2% per point) but virtually no difference for the other (0.6% per point). If an observer could detect the target, he could localize it to one side of the visual field.

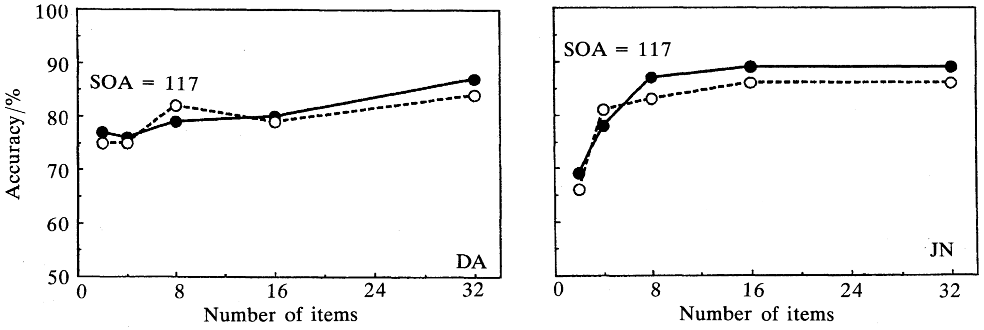

Figure 5. Detection and localization accuracies as a function of the number of Gabor items. SOA is the time in ms between presentations of the target and the mask. The two panels show data for different observers (DA and JN). Solid symbols show detection data, and open symbols represent localization.

4 Discussion

4.1 *Summary of results*

Observers detected, localized, and identified targets with similar accuracies.(2) This finding is consistent with the notion of 'labelled lines' (eg, Watson and Robson 1981) and suggests that there is no evidence of processing stages in which targets could be localized but not identified (Atkinson and Braddick 1989; Johnston and Pashler 1990; Sagi and Julesz 1985) or identified but not localized (Treisman and Gelade 1980). Although identification and localization were not directly compared in the same trials, it is reasonable to infer that they produce identical results. With detection serving as a baseline, both identification and localization exhibit the same relative accuracies.

---

(2) It could be argued that comparing identification and localization accuracies is, as one of the reviewers said, like comparing apples with oranges. Well, you can compare apples and oranges: both have weight, a property which permits direct comparison. In any event, there is much precedent (eg, Watson and Robson 1981) for using direct comparison of detection and discrimination accuracies to reveal the properties of visual mechanisms.

Observers obviously performed only a very crude localization, ie, to one side of the visual field. But observers were also confident that they could have localized the targets much more precisely if they had had the opportunity.

There is a slight superiority of detection over both localization and identification. This difference is probably unimportant for two reasons. First, the different is very small. The average superiority of detection over the other tasks was 1.7% per data point and, for the same condition, no data points differed by more than 4%. Second, preliminary studies showed that the order of responses affected the data. For example, if the localization response preceded the detection response, localization was slightly better than detection. Although the order of response was carefully counterbalanced across blocks, I cannot be sure that the effect was eliminated.

My localization results extend those of the study of Sagi and Julesz (1985) who found that observers could localize targets relative to one another. In their study observers were asked to say whether three targets formed a right triangle. Observers performed with high accuracy, showing the ability to localize targets. Note, however, that this required knowledge of relative but not absolute position in the visual field. This could be possible with a 'reversal of indices' mapping (Ballard 1986) where the feature module retains crude topographic information. The present results go further, since they show that observers have knowledge of visual field position as well.

Last, the results confirm the observation of Sagi and Julesz (1987) that detection improves with increased numbers of distractors. However, Sagi (1990) claimed that search deteriorates with display sizes of up to 8 elements but then improves with addition of more elements. My data do not show this effect: performance improved in many cases with increases from 2 to 8 display elements.

4.2 *Comparison with previous studies*

My results apparently conflict with data from previous studies. Since search theories rely heavily on data from detection, identification, and localization tasks, it is important to resolve the inconsistencies. My data alone are suggestive but insufficient, since they have the burden of trying to prove the null hypothesis—that there is no difference in performance on the three tasks. I therefore examined previous search studies (table 1) and found that many seemingly discrepant data actually agree reasonably well with the results reported here. Much of the apparent inconsistency is due not to the data themselves, but to the authors' interpretations.

Of all previous studies, that of Sagi and Julesz (1985) is most similar to the present experiments. They measured detection as well as localization and discrimination for oriented lines. However, there are difficulties in interpreting their results. First, observers performed discrimination, not identification. While discrimination might imply that targets must first be identified, there is the additional task of comparing the targets. Second, as noted, the localization task tested for relative rather than absolute location. The study established that observers can localize targets with respect to one another but not necessarily with respect to their location in space.

Sagi and Julesz (1985) say that observers perform detection and localization at the same level but that discrimination takes more time. However, the data do not unequivocally support their conclusions (see also Atkinson and Braddick 1989). First, only one of the two observers showed a clear superiority of localization over detection. Second, the data show that detection and discrimination can be equivalent. The detection task required observers to indicate the number of targets: 1/2, 2/3, or 3/4. The discrimination task was a same-different judgment of orientation. With 2 targets, there was virtually no difference between detection and discrimination. The difference between detection and discrimination occurred only with 3 and 4 targets.

Apparently, it is possible to discriminate up to 2 targets preattentively. However, this requires the observer to compare the identity of two targets at different locations. This should be impossible if attention is conceptualized as a single beam.

It is true that discrimination performance was poorer with 3 and 4 targets. However, discrimination likely requires two steps: identification followed by comparison. Perhaps the comparison is a serial process, and the longer thresholds reflect increased difficulty of comparison, not identification. Of course, it is possible to argue that discrimination occurs without prior identification. Folk and Egeth (1989) proposed that the difference signals themselves can be compared. However, they did not explain what dimension would be useful to make the comparison.

The study of Johnston and Pashler (1990) is the most recent to compare detection and identification. They noted that many previous studies may have suffered from subtle artifacts, and after presumably correcting these, reported little difference between localization and identification. However, like Sagi and Julesz (1985) and Atkinson and Braddick (1989), they suggested that it may sometimes be possible to localize without identifying the target. Johnston and Pashler estimated a maximum 10% superiority of localization over identification, but acknowledged that even this could be an overestimate.

Table 1 also includes two experiments where observers performed only one task. Both studies are relevant, however, because they demonstrate target identification in parallel search. Folk and Egeth (1989) attempted to show that identification occurred in parallel search and that the positive slope for discrimination found by Sagi and Julesz was due to postperceptual processes. The relevant experiment (3) tested detection of a single line segment on a background of diagonal distractors. The display also contained 'pseudotargets' scattered in the distractors. If the target was vertical, the pseudotargets were horizontal and vice versa. The observer could not rely on a different signal for detection or identification because the target differed from both pseudotargets and distractors. To confirm that observers were performing a parallel search, the number of pseudotargets was varied from 1 to 5. Results showed that target detection was independent of the number of pseudotargets. Folk and Egeth (1989) suggested that same-different discrimination tasks require postperceptual processing which causes the appearance of sequential search. Unfortunately, the specific model they proposed is very complex and difficult to assess.

Duncan (1989) tested target identification for a specific colored patch among 2 to 5 distractors. All items had different colors, so observers could only perform the task by absolute color identification. Response time was independent of display size, showing that identification could be accomplished with a parallel search. Duncan (1989; see also Duncan and Humphrey 1989) also posited a search theory based on signal detection theory: targets (signal) pop out if the distractors (noise) are significantly different. Search slows if the target-distractor differences decrease (signal and noise are moved closer) or if the distractors are varied from each other (increasing variability of the noise distribution).

The current results fit nicely with all of these data. I found virtually no difference (i) between detection and identification, (ii) between detection and localization, and (iii) by association with detection, no difference between identification and localization (although this inference requires some assumptions). The results of the current experiments therefore agree very well with those of Duncan (1989) and Folk and Egeth (1989), and reasonably well with those of Johnston and Pashler (1990). My results are also compatible with the data of Sagi and Julesz (1985) but not with Sagi and Julesz's interpretation. Taken together, these studies suggest no compelling difference among detection, identification, and localization tasks.

However, there are also two studies which report seemingly incompatible results. Treisman and Gelade (1980) concluded that observers may identify a target without knowing its location. As shown in table 1, no study has replicated this finding. Possible reasons will be discussed below.

In addition, Atkinson and Braddick (1989) contradicted both Treisman and Gelade (1980) and the other studies by finding that localization could be superior to identification. Observers identified targets and performed a coarse or fine localization in a feature search. Coarse localization was superior to identification, which in turn was perhaps slightly better than fine localization. Atkinson and Braddick postulated that the coarse localization is preattentive but that identification and fine localization require that attention be directed toward the target.

It is seen from table 1 that there are two major differences across the experiments. The first is the particular types of items. I primarily used Gabors whereas the discrepant experiments employed colored letters or line segments. This probably does not explain the different results because I replicated some Gabor results with line segments and because my results agree with Johnston and Pashler's (1990), who used the more conventional colored letters. Second, Treisman and Gormican (1988) and Johnston and Pashler (1990) suggested that uniformity of the distractor backgrounds might be important. Uniform background might promote aggregation of the distractors, so the observer is actually detecting not a single item, but a break in a texture (cf Sagi and Julesz 1985). But the pattern of results across studies does not correlate with distractor uniformity. Specifically, my data agree well with those of Johnston and Pashler (1990) and Folk and Egeth (1989) although my distractors were homogeneous and theirs were not. Moreover, this leaves the equality of detection and identification still unaccounted for. It seems more likely that the divergent results are due to differences in methodology. Treisman and Gelade (1980) suffer the 'negative information problem' noted by Johnston and Pashler (1990). Suppose that in an experiment shape and color are used as the target features. If these are not equally detectable, observers can identify targets above chance with a simple decision strategy: if the target is not seen, then assume it was the less detectable of the two features. This strategy is less likely to be successful for determining location because there are many possible locations and they may differ little in detectability. In any event, no search study other than that of Treisman and Gelade (1980) has shown a superiority of identification over localization.

The findings of Atkinson and Braddick (1989) are more difficult to explain away. The superiority which they found of coarse localization over identification contradicts my findings. There are many differences between the two studies. One major procedural difference is that Atkinson and Braddick measured localization and identification in different trials. Perhaps observers adopted different strategies for the two tasks. Although I also collected identification and localization data in different trials, detection served as a common baseline. Another possible source of difference may relate to their stimuli. Because of aliasing in their display, the target differed from the distractors, not only in orientation, but also slightly in length, width, and probably contrast. If the observers detected targets with these other attributes, then orientation would not be available at detection threshold. On the other hand, there may simply be sets of circumstances which promote interrupt detection over feature detection. For example, discontinuities might be detected from the frequency spectrum of a texture rather than from its individual elements (eg, Beck et al 1987; Nothdurft 1990). This would more likely occur with line segments than with Gabor functions.

To summarize the studies in table 1: there is no compelling evidence to suggest that observers perform differently on detection, localization, and identification tasks. Most discrepancies can be explained by categorical interpretation of small effects or by procedural artifacts.

4.3 *Implications for theories of visual search*

The data presented here, along with those from previous studies, provide little evidence that detection, identification, and localization tasks are sequentially performed operations. There is no evidence that they represent distinct preattentive and attentive processing stages. This argues against feature-integration theory and its major competitors. For the purpose of discussing the implications of the data in more detail, search theories are divided into four classes: (i) feature-integration theory, (ii) guided-search theories, (iii) interrupt theory, and (iv) alternative theories.

4.3.1 *Feature-integration theory.* Feature integration is a complicated and detailed theory which is comprised of separate subtheories on representation, algorithm, and architecture. The current results, along with data from other studies, are inconsistent with all three. I outlined elsewhere (Green 1991) the architectural issues, so I restrict the ensuing comments to representation and algorithmic components of the theory.

The representational component posits nontopographic feature maps linked to a topographic master map of location. Features float free of location until focal attention activates the links to the master map. The theory rests in part on data showing that observers can identify targets that cannot be localized. This result has not been replicated and may be traced to artifacts (Johnston and Pashler 1990). Moreover, other data (reviewed in Green 1991) argue convincingly against nontopographic feature maps.

The second is an algorithmic theory which describes the procedure for performing search. Observers detect pooled activity, localize by focusing attention, and then glue raw features into perceptible objects. But there is little evidence to prove that detection, localization, and identification tasks represent three distinct processing stages.

The weight of evidence supports the view (Nakayama 1990; Tsal 1989) that there is no sharp distinction between 'preattentive' and 'focal attention' vision. Early vision is not governed by a two-stage parallel–serial process in which features are first detected and then integrated.

Some still hold that serial and parallel search are distinct processes. Braun and Sagi (1990) found that a central attentive task (identification) interferes with a peripheral attentive (identification) task but not a peripheral preattentive task (simple detection). However, their experiment is hard to interpret. First, the peripheral identification task used no distractors, while the peripheral detection task required a field full of distractors. This makes the experiments incomparable. Further, observers performed the central task better when they had to detect rather than identify the peripheral target. That is, detection may have been unaffected because the central task was easier and required less concentration of attention in the center of the field.

4.3.2 *Guided-search theories*. There are several recent revisions to feature-integration theory. Many studies (eg, Krose and Julesz 1989; Pashler 1987; Treisman and Sato 1990; Wolfe et al 1989) suggest that observers need not search item by item to detect conjunctions. To accommodate this finding, some feature-integration theories have mutated into 'guided-search theories', which suggest that parallel and serial search are connected. Information collected by a parallel search restricts the subsequent serial search to items likely to be targets. Most guided-search theories retain the central role for attention as perceptual glue, and the distinction between preattentive and focal search. The new idea is that the preattentive and focal searches are separate but not independent.

In the most recent change to feature-integration theory, Treisman and Sato (1990) suggested that observers, in addition to using the moving spotlight, might employ an attentional strategy which inhibits entire feature spaces. Acting through the links to the master map, this reduces the activity in some conjunctions and prevents inspection during serial search. It is unclear what, if any, relationship this spatially independent attention holds to the standard spotlight.

Wolfe et al (1989) offered a similar model in which features sum to produce activity in the master map. Attention is directed toward locations with this activity. Here, attention amplifies rather than inhibits. Both theories still retain the dichotomy between attentive and preattentive search, the need for attention to glue features together, and the independence of feature identity and location.

As noted, the results of the present experiments, supported by data from previous studies, show that targets are detected and identified at the same level of performance. This contradicts any theory which postulates distinct serial mechanisms for the three tasks, and any theory which gives focal attention a central role in feature localization and identification. The guided-search theories retain all of these concepts.

4.3.3 *Interrupt theories*. Several authors postulate the following search theory: a feature (Johnston and Pashler 1990) or gradient discontinuity (Atkinson and Braddick 1989; Sagi and Julesz 1985, 1987) produces an 'interrupt' signal, pops out, and pulls attention to the target location. The attentional beam falls on the target, which can then be identified. Both sets of authors developed the theory to account for the superiority of detection over identification. But, as the preceding discussion shows, observers perform similarly on the two tasks.

Sagi and Julesz have buttressed the theory with two other pieces of evidence. First, observers can localize targets in parallel search (Sagi and Julesz 1985). Second, detection improves with increased numbers of distractors (Sagi 1987; Sagi and Julesz 1990). They attributed this latter result to a local process which compares neighboring elements. Greater numbers of items decrease processing time by reducing the distance across which the comparison must be made.

My data are only partially consistent with their interpretation. I confirmed and extended the finding that observers can localize targets in parallel search. Further, both detection and identification improved with larger display sizes. This does not conflict with the notion that local constraints are important in target detection. But if the only output from preattentive vision is a difference signal, it should be impossible to distinguish the left from the right oblique targets. However, both discrimination and localization improved with increased numbers of distractors. Whatever the reason for the facilitating effect of increased display size, it cannot be explained by a featureless difference signal from a texture gradient.

Of all current theories, that of Johnston and Pashler (1990) is the best fit to the data. They found a 10% localization and identification difference and noted that even this small difference might be an overestimate due to artifacts. Their theory differs significantly from feature-integration theory in saying that attention is not needed to glue features together. Instead, conjoining of features occurs in parallel across the entire visual field. Search data support both conclusions. But Johnston and Pashler also suggest that attention plays an important role by transferring identity to 'central processes'. This idea stems from the questionable assumption that localization is better than identification; it is also extremely vague.

4.3.4 *Alternative theories*. None of the current theories is entirely consistent with existing data. One way to generate alternatives is to start relaxing the base axioms of current theories. The major axiom of most theories is that features are represented in separate maps. A new theory might start by denying the existence of separate feature representations. Navon (1990) has suggested, for example, that many of the current search data could be explained by assuming that features are always conjoined. However, Green (1991) reviews many arguments favoring feature representations.

A new theory might retain feature maps but relax some secondary axioms. Feature-integration theory has two: (i) feature maps have no useable topographic information so features float free until localized by attention and (ii) feature maps are independent with no interconnections. Together, the axioms create a theory which requires modules to be tied together through a central location map. This forces the need for a central control (attention) mechanism to integrate the modules. In the computational literature (eg, Erman et al 1980) this would be termed a 'blackboard architecture' because the information sources (feature modules) can communicate with each other only by 'posting messages on a central blackboard' (the master map).

What happens if the secondary axioms are relaxed? The theories of Sagi and Julesz (1985) and Johnston and Pashler (1990) drop the first by suggesting that feature modules are topographic. This may repair some weaknesses in feature-integration theory, but does not force much alteration in the basic tone. Attention is still required for target identification and attention may still serve to glue features together.

Elsewhere (Green 1991), I have suggested a new class of theories which relaxes both axioms. It retains the concepts of separate feature modules and topographic representation but suggests that the modules are directly connected to one another. This creates a very different theory with feature detection and integration occurring in a single operation. There is no central location map, and no central control process to integrate maps. Instead, control is distributed, so focal attention does not play any role in conjoining features.

4.4 *Conclusion*

The data reported here show that detection, identification, and localization tasks are performed at the same levels of accuracy, a conclusion generally supported by a review of previous studies. This argues against both feature-integration theory and interrupt theory, since these assume that the tasks can only be performed in a particular order. However, both classes of theories find support in other phenomena. For example, feature-integration theory invokes illusory conjunctions and the slopes of search curves as supporting evidence. Close examination (Green 1991; Navon 1990; Tsal 1989) of these phenomena, however, reveals that they do not support feature-integration theory.

Acknowledgements. This work was supported by NSERC Grant OPG0046297.

References

Atkinson J, Braddick O J, 1989 "'Where' and 'what' in visual search" *Perception* **18** 181 - 189

Attneave F, 1974 "Apparent motion and the what-where connection" *Psychologia* **17** 108 - 120

Ballard D, 1984 "Parameter nets" *Artificial Intelligence* **22** 235 - 267

Ballard D, 1986 "Cortical connections and parallel processing: structure and function" *The Behavioral and Brain Sciences* **9** 67 - 120

Ballard D, Hinton G, Sejnowski T, 1983 "Parallel visual computation" *Nature (London)* **306** 21 - 26

Barlow H, 1981 "The Ferrier Lecture: Critical limiting factors in the design of the eye and visual cortex" *Proceedings of the Royal Society of London B* **212** 1 - 34

Barlow H, 1986 "Why have multiple cortical areas?" *Vision Research* **26** 81 - 90

Beck J, Sutter A, Ivry R, 1987 "Spatial frequency channels and perceptual grouping in texture segregation" *Computer Vision, Graphics and Image Processing* **37** 299 - 325

Braun J, Sagi D, 1990 "Vision outside the locus of attention" *Perception & Psychophysics* **48** 45 - 58

Cowey A, 1979 "Cortical maps and visual perception: The Grindley memorial lecture" *Quarterly Journal of Experimental Psychology* **31** 1 - 17

Cowey A, 1981 "Why are there so many visual areas?" in *The Organization of the Cerebral Cortex* Ed. F Schmitt (Cambridge, MA: MIT Press)

DeYoe E A, Van Essen D C, 1988 "Concurrent processing streams in monkey visual cortex" *Trends in Neurosciences* **11** 219 - 226

Duncan J, 1989 "Boundary conditions on parallel processing in human vision" *Perception* **18** 457 - 469

Duncan J, Humphrey G, 1989 "Visual search and stimulus similarity" *Psychological Review* **98** 432 - 458

Erman L, Hayes-Roth F, Lesser V, Reddy R, 1980 "The Hearsay-II speech understanding system: integrating knowledge to resolve uncertainty" *Computing Surveys* **14** 213 - 253

Folk C L, Egeth H, 1989 "Does the identification of simple features require serial processing?" *Journal of Experimental Psychology: Human Perception and Performance* **15** 97 - 110

Green M, 1986 "What determines correspondence strength in apparent motion?" *Vision Research* **26** 599 - 607

Green M, 1991 "Visual search, visual streams and visual architectures" *Perception & Psychophysics* **50** 388 - 403

Hubel D, Livingstone M, 1985 "Complex-unoriented cells in a subregion of primate area 18" *Nature (London)* **315** 325 - 327

Hubel D, Livingstone M, 1987 "Segregation of form, color and stereopsis in primate area 18" *Journal of Neuroscience* **7** 3378 - 3417

Johnston J, Pashler H, 1990 "Identity and location in visual feature perception" *Journal of Experimental Psychology: Human Perception and Performance* **16** 843 - 856

Krose B, Julesz B, 1989 "The control and speed of shifts of attention" *Vision Research* **29** 1607 - 1619

Livingstone M, Hubel D, 1987 "Psychophysical evidence for separate channels for the perception of form, color, motion and depth" *Journal of Neuroscience* **7** 3416 - 3468

Livingstone M, Hubel D, 1988 "Segregation of form, color, movement and depth: anatomy, physiology and perception" *Science* **240** 740 - 749

Nakayama K, 1990 "The iconic bottleneck and the tenuous link between visual processing and perception", in *Vision: Coding and Efficiency* Ed. C Blakemore (Cambridge: Cambridge University Press) pp 43 - 61

Navon D, 1990 "Does attention serve to integrate features?" *Psychological Review* **97** 453 - 459

Nothdurft H, 1985 "Sensitivity for structure gradient in texture discrimination tasks" *Vision Research* **25** 1957 - 1968

Nothdurft H, 1990 "Texture segregation by associated differences in global and local luminance distribution" *Proceedings of the Royal Society of London B* **239** 295 - 320

Pashler H, 1987 "Detecting conjunctions of color and form: reassessing the serial search hypothesis" *Perception & Psychophysics* **41** 191 - 201

Sagi D, 1990 "Detection of an orientation singularity in Gabor textures: Effects of signal density and spatial frequency" *Vision Research* **30** 1377 - 1388

Sagi D, Julesz B, 1985 "'Where' and 'what' in vision" *Science* **228** 1217 - 1219

Sagi D, Julesz B, 1987 "Short-range limitations on detection of feature differences" *Spatial Vision* **2** 39 - 49

Treisman A, Gelade G, 1980 "A feature-integration theory of perception" *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, Sato S, 1990 "Conjunction search revisited" *Journal of Experimental Psychology: Human Perception and Performance* **16** 459 - 478

Treisman A, Schmidt H, 1982 "Illusory conjunctions in the perception of objects" *Cognitive Psychology* **14** 12 - 31

Tsal Y, 1989 "Do illusory conjunctions support feature-integration theory? A critical review of theory and findings" *Journal of Experimental Psychology: Human Perception and Performance* **15** 394 - 400

Watson A, Robson J, 1981 "Discrimination at threshold: labelled detectors in human vision" *Vision Research* **21** 341 - 347

Wolfe J, Cave K, Franzel S, 1989 "A modified feature-integration model for visual search" *Journal of Experimental Psychology: Human Perception and Performance* **15** 419 - 433