Filtering by Movement in Visual Search

Peter McLeod, Jon Driver, Zoltan Dienes, and Jennie Crisp
Oxford University, Oxford, England

Search for a target defined by a conjunction of movement and form (e.g., an X moving up in a display of intermingled Os moving up and stationary Xs) is parallel. This result is also found if (a) the moving Os and target X move in unpredictable directions so that the moving stimuli do not form a clear perceptual group or (b) the nontarget Xs also move but in a known, different direction from the Os and target X. In contrast, search is slow and serial if the target may be unpredictably among either moving or stationary stimuli. These results suggest that a component of the visual system operates as a movement filter that can direct attention to stimuli with a common movement characteristic. The filtering cue can be moving (vs. stationary), or movement in 1 particular direction. The results do not support the view that attention can only be directed to groups defined by common fate.

McLeod, Driver, and Crisp (1988) showed that search for a target defined by a conjunction of movement and form (e.g., an X moving up among an intermingled set of Os moving up and stationary Xs) was parallel (i.e., detection time was independent of the number of nontarget stimuli). This result, and others obtained with moving displays by Nakayama and Silverman (1986), run counter to the well-known predictions of Treisman and Gelade (1980). On the basis of experiments involving search for targets defined by conjunctions of color and shape in stationary displays, they claimed that search for targets defined by a conjunction of features should be serial (i.e., search time should increase linearly with set size). Logically, McLeod et al.'s task was exactly the same as those used by Treisman and Gelade, except that the color distinction (e.g., red vs. green) was replaced by a movement distinction (moving up vs. stationary). Viewing the displays used by McLeod et al. immediately suggests an account for the parallel search they found in terms of perceptual segregation. Even though the moving and stationary stimuli are randomly intermingled, they are seen as two groups, one moving up and one stationary. Within the moving group, the target is easy to detect because it is distinguished by a simple form difference from the other moving items (it is an X among Os). In other words, grouping by movement has turned the task from conjunction search to feature search, and so, as Treisman and Gelade would predict, it has become parallel.

McLeod et al. (1988) offered a different kind of explanation. They suggested that their results reflected the operation of a *movement filter* in the visual system. In its simplest form, this would be a subsystem in which all moving items are well represented, but stationary items would be represented weakly, if at all. Such a subsystem could allow visual attention to be directed to just the moving items in a visual array. Hence, a target among the moving items defined by a clear form difference, such as X versus O, could be easily detected.

The "perceptual grouping" account would dispense with the hypothetical subsystem and just cite the result as an example of the well-known potency of grouping by common fate (Wertheimer, 1923). Indeed, numerous authors have suggested that visual attention is directed to perceptual groups defined by Gestalt laws of organization (e.g., Driver & Baylis, 1989; Duncan, 1984; Kahneman & Henik, 1981; Neisser, 1967; Prinzmetal, 1981). Treisman (1988) has also noted the importance of grouping in visual search; displays that allow parallel conjunction search are often those in which nontargets with one feature in common form a global figure (i.e., a group) from which the target stands out. For example, search for targets defined by a conjunction of color and form can become parallel rather than serial if all the stimuli of one color or shape are adjacent, and hence spatially grouped together (Treisman, 1982). "Grouping" and "a movement filter" are not mutually exclusive explanations; rather, they operate at different levels. Grouping offers a description of phenomenology; the movement filter is a mechanism that might turn out to underlie the phenomenology of search in moving displays.

In this article, we will explore the conditions under which search for a target defined by a conjunction of movement and form is still parallel and when it becomes serial. What happens if the moving stimuli move in different directions? What happens if all stimuli in the display are moving? Is it possible to search more than one movement group simultaneously? We will show that an account of visual search in moving displays that relies on grouping by common fate is unsatisfactory because search can be easy in displays in which the moving items do not share common fate. In contrast, the results we will report are consistent with the idea of a movement filter in the visual system. Neurophysiological data and the visual search performance of a patient with extrastriate damage suggest that the medial temporal cortical area (MT) is involved in the processing performed by this filter.

Experiment 1

In the experiments reported by McLeod, Driver, and Crisp (1988) subjects searched for an X moving up among inter-

mingled Os moving up and stationary Xs; thus, the moving stimuli all shared a common direction of movement. Although the moving and stationary items were intermingled, the moving ones formed a clear subjective group moving up the screen. What happens if the moving stimuli no longer share common direction of motion but move in different directions? Because they will not form a common-fate group, the hypothesis that such grouping is necessary for parallel search suggests that search will be serial. The movement filter theory makes the opposite and somewhat counterintuitive prediction: Because all moving items will be represented in the filter, irrespective of direction of movement, it should be easy to direct attention to just the moving stimuli even when they move in random directions. Therefore, parallel search should be found because a simple form difference distinguishes the target (X) from the other moving stimuli (Os). The contrasting predictions of the movement filter and the grouping hypotheses were tested in Experiment 1. The task was to search for a single X moving in an unpredictable direction in a field of stationary Xs, and Os moving in unpredictable directions.

Method

The task. The subjects faced a display like that shown in Figure 1. The target was a single moving X, present on 50% of the trials. The subjects responded *Yes* (target present) with a button in one hand, or *No* (target absent) with a button held in the other hand. One half of the subjects responded *Yes* with their right hand and the other half, with their left. On target trials, there were an equal number of moving Os and stationary Xs, plus one moving X. On nontarget trials, the moving X was replaced by an extra O.

Os and Xs were initially distributed at random across the screen. The screen was divided into imaginary quadrants. Any moving item (i.e., an O or the target X) that started in the top left-hand quadrant could move either to the right or down. Different items in the quadrant were selected at random on different trials to move in either direction. Similarly, in the bottom right-hand quadrant, items could move either 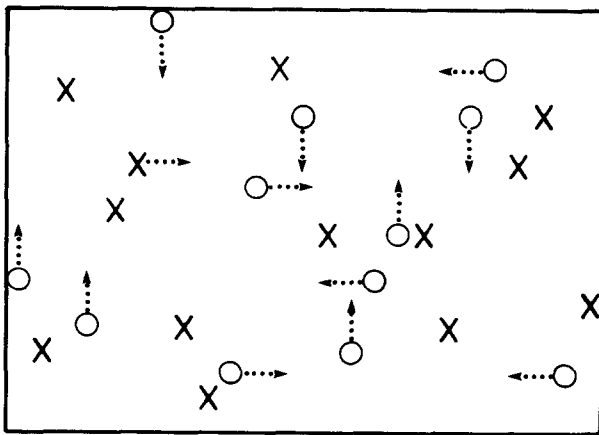to the left or up, and so on for the other quadrants. In this way, although the movement of any individual item was unpredictable before the start of the trial, all items remained on the screen during the display period. Moving items would sometimes intersect with and pass through other items.

The display. The display subtended 11° × 8° at a viewing distance of 30 cm. Individual letters subtended approximately 0.4°. Moving stimuli moved at 3.3° per second.

The stimuli were displayed on an Electronic Visuals oscilloscope with a P34 phosphor, driven by a Cambridge Electronic Developments minicomputer. The stimuli were letters drawn on a 9 × 5 dot matrix. The system takes 16 μs to display 1 point; approximately 6 ms to put up a display of 25 letters. Every 10 ms the display was refreshed with, according to direction of movement, either the vertical or horizontal coordinates of the moving items incremented by an amount that produced a change in position subtending roughly 0.03°. The stimuli appeared to move smoothly. Luminance levels of the screen and room were adjusted so that the moving stimuli left no apparent smear trail.

A trial started with a small fixation cross in the center of the screen for 500 ms, a blank period of 500 ms, and then the display for 1,200 ms. At the end of the display, or after a response from the subject, whichever came later, there was a pause of 1,500 ms before the start of the next trial.

Subjects. The subjects, aged between 18 and 45 years, were taken from the Oxford University Psychology Department subject panel. All experimental conditions reported in this article used a new set of 6 subjects drawn from this population.

Procedure. On successive blocks of 50 trials, there were either 5, 9, 17, or 25 items in the display. Each subject performed the search task at each of the four set sizes once in a random order. That order was then repeated a further seven times. The first two sets of four blocks were treated as practice. The data presented are the reaction times of correct responses on the last six blocks of 50 trials at each set size. The data from the first 10 trials of each block of 50 were discarded. Each data point represents the median performance on 120 trials per subject, less those on which errors occurred, out of 200 trials at that set size. The first four runs at each set size were performed on 1 day; the second four, on the next day.

The subjects were given feedback on their median reaction time and error rate at the end of each block. They were directed to aim at an error rate of between 5% and 10% and were encouraged to speed up if they were below 5% and slow down if above 10%.

Figure 1. A typical display in Experiment 1. (Subjects search for an X moving in an unpredictable direction among stationary Xs and Os moving in unpredictable directions. The dotted arrows do not appear on the display; they indicate moving stimuli.)

Results

The results are shown in Table 1. The slope of the regression between detection time and set size was 5 ms per item for Yes responses and 23 ms per item for No responses. These slopes are very close to those reported by Treisman and Gelade (1980, Experiment 1) for detecting single features of color or form. They reported average slopes of 3 ms per item for positives and 25 ms per item for negatives. Those results were interpreted by Treisman and Gelade as demonstrating parallel search. Presumably, the same interpretation should be made here. Such an interpretation of the Yes responses seems straightforward. Although slightly greater than the 3 ms per item slope found by McLeod et al. (1988) when the moving stimuli all moved in the same direction, the value is within the range that is usually interpreted as demonstrating parallel search (see Duncan & Humphreys, 1989).

The interpretation of the No responses is less clearcut because they show an effect of set size. Unfortunately, there

Table 1
Time (ms) to Detect the Presence or Absence of a Moving X in an Array of Randomly Moving Os and Stationary Xs as a Function of the Number of Items in the Display in Experiment 1

	Display size					Intercept	Slope
Response	5	9	17	25	r^2	(ms)	(ms/item)
Yes	688	691	731	785	0.96	654	5
No	682	772	923	1,143	0.99	563	23
Error (%)	3.6	3.1	3.5	4.7			

is no straightforward interpretation for the slope of No responses unless it is double the slope for Yes responses or equal to it. The first implies serial self-terminating search (and clearly does not apply in this case), the second implies that the same search process precedes both Yes and No responses. The difference between the slopes we observe for Yes and No responses suggests the intuitively plausible but regrettably untestable assumption that on a certain (unknown) proportion of trials in which no target is found the subject does a little rechecking before responding. This somewhat ad hoc proposal is the traditional one in this field (Duncan & Humphreys, 1989; Treisman & Gelade, 1980; Treisman, Sykes, & Gelade, 1977).

Conclusion

The time to detect the presence of an X in these displays is scarcely affected by set size. Thus, it appears that the visual system is able to separate all of the items that are moving, irrespective of direction, from all of those that are stationary and conduct a feature search among the former for an X. This result would be expected if the visual processing system contained a movement filter that represented all moving stimuli well but stationary stimuli weakly, if at all. Despite the fact that the moving stimuli moved in different directions, they would all be represented in the filter. In contrast, the result is surprising if one believes that grouping by common fate is required before parallel conjunction search in moving displays will appear. The randomly moving items of Experiment 1 do not form a clear perceptual group as they do when all the moving stimuli move in the same direction. Admittedly, search performance in Experiment 1 was not as efficient as it was when the moving stimuli moved in the same direction and a clear group was formed. McLeod et al. (1988) reported search rates of 3 ms per item for positive responses and 5 ms per item for negative responses when the moving stimuli had a common direction. However, the major difference between their results and the present findings is that it takes increasingly longer to be sure that a target is *not* present with increased set size in the case of multiple directions. Detecting the target when it *is* present produces parallel search whether the stimuli move together or not.

Experiment 2

Target detection was little influenced by set size in Experiment 1. The movement filter explanation is that the moving

and stationary items are represented in different components of the visual system. If this is correct, a task that requires search of both moving and stationary stimuli should be more difficult than one that requires search only among moving stimuli or only among stationary stimuli (on the assumption that it is more difficult to direct attention to two components of the system than to one). This prediction was tested in Experiment 2.

Four new groups of 6 subjects were used, 1 in each of four conditions. All subjects viewed a display that contained items moving up, items moving down, and stationary items. Thus, for every condition, the displays contained three groups of items defined by common fate: up, down, and stationary. In every condition, search was for an X among Os, with a background of nontarget Xs. What differed between the conditions was whether the Os were all moving, all stationary, or whether some were moving and some were stationary. Subjects in condition *Up–Stationary* viewed a display with Os moving up the screen, stationary Os, and Xs moving down the screen. They searched for a target X that could be unpredictably among either the Os moving up the screen or among the stationary Os. Subjects in condition *Up–Down* viewed a display with Os moving up, Os moving down, and stationary Xs. The target was an X that could be in either of the moving groups. Subjects in condition *Stationary* viewed a display with stationary Os, and Xs moving up and down the screen. Their target was a stationary X. Subjects in condition *Up* viewed a display with Os moving up the screen, and stationary and downward moving Xs. Their target was an X moving up. See Figure 2 for example displays from the four conditions.

The task for condition Up–Down was similar to that of Experiment 1. Subjects had to search among stimuli moving in different directions. They knew that the target would be moving, but they did not know in which direction. Extrapolating from Experiment 1, we could predict that this task would show little effect of set-size. The crucial contrast was between their performance and that of subjects in condition Up–Stationary. The movement filter hypothesis predicted that search in condition Up–Stationary would be more difficult. In condition Up–Down, the target could be found by attending just to the contents of the movement filter; however, in condition Up–Stationary attention had to be directed both to moving items in the filter and stationary items outside it. The grouping hypothesis predicted little difference. Both conditions required searching among two separate groups clearly defined by common fate; performance should therefore have been similar in each case. If Up–Stationary did prove to be harder than Up–Down it could be argued that this was simply because searching in a stationary group was harder than searching in a moving one. The performance of subjects in the Stationary condition allowed us to assess this. Finally, the Up condition gave an index of the difficulty of searching in a moving group alone, when both static and moving items must be filtered out.

Method

The display. The subjects faced displays similar to those shown in Figure 2. In each case, the upper half of the screen contained a
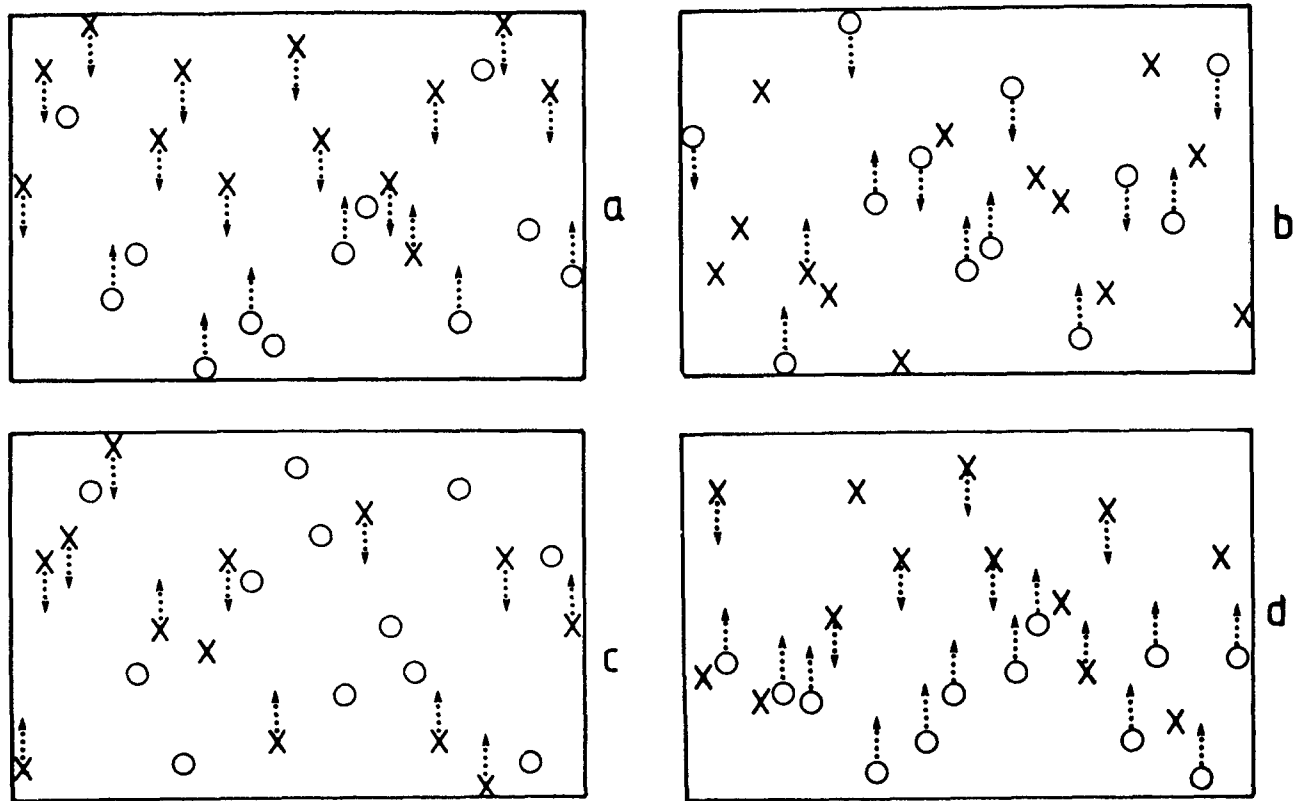
Figure 2. Typical displays for the four conditions of Experiment 2: (a) search for an X among the stationary and upward moving Os, (b) search for a moving X among the upward and downward moving Os, (c) search for a static X among the stationary Os, (d) search for an X among the upward moving Os. (The dotted arrows do not appear on the display; they indicate moving stimuli.)

random distribution of stimuli that moved down; the lower half contained stimuli that moved up. Stationary stimuli were distributed across the whole screen. The display was divided into 25 imaginary vertical columns with one stimulus in each, for a set size of 25. In smaller set sizes, an adjacent subset of these columns was used to produce an equivalent density of stimuli at different set sizes. On each trial, a new subset of columns was chosen at random. The horizontal separation was just sufficient to ensure that adjacent stimuli did not cross as the moving items moved up or down.

The displays were designed to ensure that in each condition subjects had to search the same number of items to find the target. For each set size N (9, 17, or 25) there were $(N - 1)/2$ nontarget Xs, $(N - 1)/2$ nontarget Os, and 1 extra stimulus that was an X on target trials and an O on nontarget trials. In condition Up–Stationary, one-half of the Os moved up the screen, and the other half were stationary (Figure 2a). The target X (if present) was equally likely to be moving up or stationary. In condition Up–Down, one-half of the Os moved up the screen, and the other half moved down (Figure 2b). The target X (if present) was equally likely to be among either group. In condition Stationary all of the Os were stationary, as was the target X, if present. One half of the nontarget Xs moved up the screen, and the other half moved down (Figure 2c). In the Up condition all of the Os were moving up, as was the target X, if present. One half of the nontarget Xs moved down, and the other half were stationary (Figure 2d).

The display lasted 2.4 s. A longer display than that used in Experiment 1 was required to make it possible to perform the search task in condition Up–Stationary. The speed was reduced to 2.2° per

second to prevent the stimuli passing the edge of the screen during the longer display.

Subjects. Four new groups of 6 subjects were drawn from the Oxford University Psychology Department subject panel.

Procedure. With the following exceptions, the procedure was as in Experiment 1. In successive blocks of 50 trials there were either 9, 17, or 25 items in the display. Subjects performed the search task at each of the three set sizes once in a random order. That order was then repeated a further three times. This was performed in a single session lasting 1 hr. The first two sets of three blocks were treated as practice. The data presented are taken from the last two blocks of 50 trials at each set size. The first 10 trials of each block of 50 were discarded. The data recorded were the median Yes and No reaction times for correct responses on the last 40 trials of each block. Thus, each data point represents the mean of the individual subjects' median correct detection latencies over 40 out of 100 responses in that condition.

Results

The detection times and error rates are shown in Table 2. The crucial comparison is between conditions Up–Stationary and Up–Down. It is much harder to search the Up and Stationary groups than it is to search the Up and Down groups. It takes longer to search through a display of 9 items in the Up–Stationary condition than it does to search through

Table 2

Time (ms) to Detect the Presence or Absence of a Target as a Function of Display Size in Experiment 2

	Display size				Intercept	Slope
Condition	9	17	25	r^2	(ms)	(ms/item)
Up–Stationary						
Yes	846	997	1,245	0.98	605	25
No	1,348	1,926	2,290	0.98	854	59
Error (%)	0.8	3.5	7.3			
Up–Down						
Yes	610	660	737	0.99	530	8
No	706	914	1,129	1.00	467	26
Error (%)	1.9	2.5	4.5			
Stationary						
Yes	759	816	941	0.98	645	11
No	862	1,140	1,518	1.00	479	41
Error (%)	2.0	1.3	3.3			
Up						
Yes	633	696	762	1.00	560	8
No	786	913	1,074	1.00	618	18
Error (%)	3.9	5.0	5.4			

a display of 25 items in the Up–Down condition. Mann-Whitney comparisons between the slopes of the regressions for the individual subjects in each group show that the set-size effect is reliably greater for both Yes and No responses in the Up–Stationary condition. (Yes: 8 ms per item vs. 25 ms per item, $U[6,6] = 3$, $p < .01$; No: 26 ms per item vs. 59 ms per item, $U[6,6] = 3$, $p < .01$.)

It is also easier to search in the Stationary condition than in the Up–Stationary condition. (Yes: 11 ms per item vs. 25 ms per item, $U[6,6] = 5$, $p = .02$; No: 41 ms per item vs. 59 ms per item, $U[6,6] = 8$, $p = .07$.) Finally, performance in condition Up, in which subjects have to search within a single common-fate group, is very similar (e.g., producing the same mean Yes slope) to performance in condition Up–Down, in which they have to search within two common-fate groups.

Conclusion

The performance in the Up–Down condition was very similar to that found in Experiment 1 and confirmed the prediction made from that result. It is relatively easy to search for a moving X among two groups of Os moving in different directions, ignoring static nontarget Xs. However, the results for the Up–Stationary condition show that if one of the two groups in which the target might appear is stationary and the other moving, search is much more difficult. This result is not solely due to the difficulty of searching in a stationary group relative to a moving group. Performances in the Stationary condition and in the Up condition showed that although it is somewhat harder to search among a stationary group than among a moving one (a result also found by McLeod et al., 1988), the increase in difficulty is not adequate to explain the difference between the Up–Down and Up–Stationary conditions.

This pattern of results is predicted by the movement filter hypothesis. It is relatively easy to direct attention just to moving items (which are represented in the movement filter),

or just to stationary ones (which are represented elsewhere). But if search is required among both moving and stationary stimuli, the task becomes much harder. This result is difficult to reconcile with the grouping hypothesis. Because two common-fate groups had to be searched in both Up–Down and Up–Stationary conditions, there should have been little difference in performance.

In Experiment 1, condition Up–Down in Experiment 2, and in all of the experiments of McLeod et al. (1988), the nontarget items sharing the form of the target (i.e., the nontarget Xs) could be rejected because they were stationary rather than moving. In each case, search for a target X among the moving Os was little affected by set size. We have attributed this result to the operation of a movement filter that segregates the moving and stationary stimuli, turning the task into a feature search among the moving stimuli. If parallel search were only found when one group is moving and the other stationary, it might not be movement per se that is crucial but the fact that some stimuli are transient and others are sustained. It appears that there may be separate channels in the visual pathway for transmitting information about transient and sustained stimuli (e.g., Enroth-Cugell & Robson, 1966; Tolhurst, 1973; Yantis & Jonides, 1984). Perhaps our results are a consequence of the existence of these channels, rather than evidence for a specific movement-filter mechanism.

However, results in condition Up from Experiment 2, in which subjects had to search for an X moving up among Os moving up, stationary Xs, and downward moving Xs, fell within the parallel range. Hence, successful filtering by movement does not require that all the discarded items be stationary. If subjects were simply attending to all transient items while ignoring sustained items, they would have been incapable of filtering out the downward moving Xs. We can therefore dismiss the suggestion that all our results simply reflect the existence of separate channels in the visual system for conveying information about sustained and transient stimuli.

If we wish to maintain the theory that a movement filter underlies these results, it is necessary to specify one of its properties: Within the movement filter, attention can optionally be restricted to stimuli moving in a prespecified direction. The visual system has the flexibility to attend either to all moving items regardless of direction (as in Experiment 1 and condition Up–Down from Experiment 2), or just to those items moving in a particular direction (as in condition Up from Experiment 2).

Experiment 3 further examined our ability to restrict search to items moving in one direction, to the exclusion of items moving in the opposite direction.

Experiment 3

The purpose of our final experiment was to examine whether the revisions of feature integration theory suggested by Treisman (1988) and by Wolfe, Cave, and Franzel (1989) can accommodate the cases of parallel searching for conjunctions of movement and form that the preceding experiments establish. The original formulation of feature integration the-

ory (Treisman & Gelade, 1980) proposed that information about individual features such as colors, orientations, directions of movement, and so on is extracted preattentively in parallel across the visual array and is represented in independent topological "maps." These give the spatial distribution of particular features such as blue, vertical, and so on. An attentive process then combines the information about the features present at a particular spatial location across different maps to allow recognition of multifeature objects but can only do so for one locus at a time. Thus, information about the individual features in the visual array is available preattentively, but information about conjunctions of features is only available after the operation of an attentive process. Hence, single features can be detected in parallel across the visual field but detecting their conjunction requires a serial process that explores each object location in turn.

Treisman (1988, p. 226) subsequently modified this theory in order to accommodate evidence that conjunction search can sometimes be parallel (McLeod et al., 1988; Nakayama & Silverman, 1986; Nakayama, cited in Treisman, 1988). Can this modification cope with the cases of parallel conjunction search established in the present experiments? The essence of Treisman's proposal is to allow the maps representing features that define nontargets to inhibit locations on the "master map," which represents all stimuli. Thus, she would explain a hypothetical case of parallel search for a red X among red Os and green Xs by suggesting that the feature map for green might be able to suppress all green locations in the master map, leaving the red ones more active. The target would then be defined by a simple form difference (X vs. O) and hence should "pop out." Wolfe et al. (1989) made a similar proposal, except they suggested that, rather than nontarget features being inhibited from the feature-maps, target features might be excited. At first sight, the inhibitory and excitatory proposals seem to make identical predictions. On the inhibitory account, conjunction search should become parallel if inhibition of nontarget features is possible from feature maps; on the excitatory account, parallel conjunction search should be found if excitation of target-features is possible. We hoped to tease the excitatory and inhibitory proposals apart in Experiment 3.

The logic was as follows. Search in Experiment 2 was most laborious in condition Up–Stationary. This result showed that subjects could not readily restrict search to the upward moving and stationary items, to the exclusion of downward moving items. On the inhibitory revision of feature integration theory, this suggests that subjects cannot readily inhibit the downward moving items. On the excitatory revision, it suggests that subjects cannot readily excite upward moving and static items simultaneously.

Consider a task in which the subject has to search for an upward moving X in upward moving Os and downward moving Xs. The inhibitory account suggests that this task will prove difficult, because to produce parallel search the subject would have to inhibit downward moving Xs, and the results from condition Up–Stationary in Experiment 2 demonstrate the difficulty of such inhibition. By contrast, the excitatory account predicts parallel search for the new task, because the subject simply has to excite the upward-moving items, and

condition Up from Experiment 2 demonstrated the ease of such excitation. To test these contrasting predictions, we examined search in the new task.

Experiment 3a

Method

The display. Subjects faced a display like that in Figure 3. The upper two-thirds of the screen contained Xs that moved down the screen; the lower two-thirds contained Os (and the target X, if present) that moved up the screen. Stimuli moved in imaginary vertical columns as in Experiment 2. Other display details were as in Experiment 1. The lowest of the downward moving stimuli and the highest of the upward moving stimuli just reached the edge of the display at the end of the display period.

Subjects. A new set of 6 subjects was drawn from the Oxford University Psychology Department subject panel.

Procedure. The procedure was as for Experiment 2, except that the set sizes were 7, 15, and 25.

Results

The results are shown in Table 3. The linear regression between detection time and set size averaged across subjects gave a slope for Yes response of 0 ms per item. The slope for No responses was 9 ms per item.

Experiment 3b

Experiment 3a yielded unequivocally parallel search. However, there is a possible artifact. The upward-moving items, including the target, if present, began the trial in the lower two-thirds of the screen, whereas the downward-moving nontarget Xs began in the upper two-thirds of the screen. As in our previous experiments, this prevented items scrolling off of the edge of the screen during the display period. However, it may have allowed subjects to detect some targets simply by
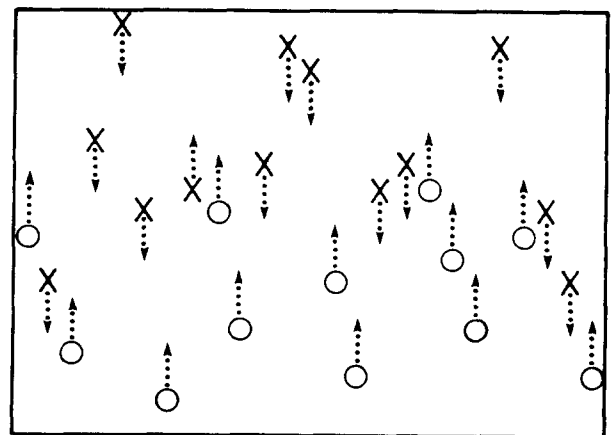

Figure 3. A typical display in Experiment 3. (Subjects search for an X moving up the screen among Os moving up and Xs moving down. The dotted arrows do not appear on the screen; they indicate moving stimuli.)

Table 3

Time (ms) to Detect the Presence or Absence of an Upward Moving X Among Upward Moving Os and Downward Moving Xs as a Function of Display Size in Experiment 3a and 3b

	Display size				Intercept	Slope
Experiment	7	15	25	r^2	(ms)	(ms/item)
3a						
Yes	609	586	615	0.07	596	0
No	615	639	772	0.95	536	9
Error (%)	4.3	3.1	4.2			
3b						
Yes	591	580	643	0.79	562	3
No	792	930	1,073	1.00	688	16
Error (%)	2.3	3.3	5.4			

registering an X towards the bottom of the screen at display onset. The present experiment eliminates this possible artifact.

Method

The only difference in method from Experiment 3a was that the vertical starting position was random for each item. Items reaching the edge of the screen scrolled around; for example, upward-moving items disappeared at the top and reappeared moving upward from the bottom of the screen with the same horizontal position. Six new subjects from the same source as before performed the search task.

Results

The results are shown in Table 3. The linear regression between detection time and set size averaged across subjects gave a slope for Yes responses of 3 ms per item. The slope for No responses was 16 ms per item. Thus, search was parallel, although the slope for No responses was larger than in Experiment 3a. Presumably the sudden changes in the display as items scrolled around the screen led to a little uncertainty on some trials in which no target was found.

Conclusion

Detection of an X moving in a specified direction among Os moving in the same direction and Xs moving in the opposite direction is independent of set size. Successful filtering by movement can be achieved with these displays. This confirms the finding of the Up condition in Experiment 2, namely, that filtering by movement does not require that the discarded items should be stationary. It is only necessary that there is a movement characteristic that distinguishes the non-targets that share the form of the target. This can be stationary versus moving, as in most of the earlier experiments, but it can equally well be a different direction of movement.

The Os form a clearly defined common-fate group in the displays of Experiment 3, just as they did when the background Xs were stationary in the experiments of McLeod et al. (1988). So this result alone would be consistent with an account in terms of grouping by common fate. However, recall that such an explanation was eliminated by Experiment

1 (in which parallel search was found even when moving items went in different directions), and Experiment 2 (in which search difficulty was not simply a function of the number of common-fate groups to be searched).

The ease of search in the present experiment supports the excitatory revision of feature integration (Wolfe et al., 1989) rather than the inhibitory revision originally suggested by Treisman (1988). The problem for the inhibitory account is that Experiments 2 and 3 force it to make contradictory predictions. The parallel search observed in the present conjunction task suggests that inhibition of downward-moving items can readily be performed, thus turning the conjunction task into a feature search among the upward-moving group. However, in the Up–Stationary condition of Experiment 2, a similar group of downward-moving items had to be excluded, and yet search was slow and serial, suggesting that downward-moving items cannot be inhibited. The excitatory account can accommodate the results without making such contradictory assertions. In the present task, parallel search is observed because the subject can excite the master-map representations of the upward-moving items. Search is more difficult in the Up–Stationary condition presumably because the subject cannot readily excite upward-moving items and static items at the same time. Of course, the latter assertion corresponds to the central claim of our own movement-filter account.

General Discussion

Summary

Experiment 1 showed that it is easy to direct attention to just the moving items in an array of moving and stationary stimuli, even if the moving stimuli move in different directions and thus do not form a single group defined by common fate. Experiment 2 showed that, in contrast, attention cannot easily be directed to both moving and stationary stimuli simultaneously. Experiment 3 showed that attention can be directed to stimuli moving in one direction despite the presence of stimuli moving in a different direction. Taken together, these results suggest that some part of the visual system acts as a movement filter that allows attention to be directed to parts of the visual scene with a particular movement characteristic. It can separate moving items from a stationary background or it can separate stimuli moving in a particular direction from items moving in a different direction.

Filtering by movement. It is clear that the ability to filter by movement would be a useful property of a visual pathway. A key task for the visual system is to ensure that focal attention is directed at the most important parts of the visual field surrounding the organism. The problem is to know which parts are important so that attention (and often gaze) can be directed at them. Objects that are moving are particularly likely to be important because they may represent either food or danger. Activity in the movement filter would represent the presence of a moving object in the environment even if attention were currently directed elsewhere. A simple movement filter with the property of registering all moving parts of the environment would be better than nothing, but would

cease to help when the organism itself was moving. Then the whole optic array would move (Gibson, 1966). A more sophisticated movement filter would be one that could detect anomalous movement in the environment. This could pick up movement in one direction against a general optic flow in a different direction and thus identify an object moving relative to the observer even when the observer was moving.

The idea of filtering (i.e., selecting a subset of items from a display because they possess a particular physical characteristic) is a familiar one in cognitive psychology. It was the key method of selection in Broadbent's (1958) seminal information processing model. Why should we suggest that a specific mechanism exists for filtering by movement rather than treat this as one application of a general-purpose mechanism for filtering in the nervous system? One reason is that evidence from other disciplines also suggests that a specific mechanism exists. Neurophysiological data show that the MT has some of the properties required of a movement filter. Converging evidence linking the neurophysiology and the behavioral studies is provided by a patient with bilateral damage to the presumed human homologue of MT who has a selective deficit of filtering by movement.

Neurophysiological evidence for a movement filter. The major projection from retina to cortex goes to the striate cortex (area V1). From there information passes along a variety of routes to different cortical areas. One of these routes, through the thick stripes of area 18 to the MT area, appears to be specialized for the handling of information about stimulus movement (Livingstone & Hubel, 1987). A comparison of the response of cells within MT to that of cells in cortical area V1 to moving and stationary slits shows that cells in MT are relatively more sensitive to moving stimuli than are those in V1 and relatively less sensitive to stationary ones (Albright, 1984; Felleman & Kasse, 1984). Cells in MT also show a selective response to particular directions and speeds of stimuli. Thus, it seems that cells in MT have some of the properties required of a movement filter. Allman, Miezin, and Mc-Guinness (1985) have shown a particularly important property of some cells in this area. They are defined as *movement cells* because they are tuned to the direction and speed of stimuli falling within a classical receptive field. However, this response can be attenuated if a large area of background outside the receptive field moves in the same direction as the target, or it can be enhanced if the background moves in the opposite direction. In other words, the cells do not just respond to moving objects as opposed to stationary ones. They respond to an object within their receptive field moving in a way that is different to the general background motion. This is the property that a movement filter would require if it were to signal movement relative to the observer even when the observer was moving.

The idea that MT may be involved in the processing required by a movement filter is supported by visual search data from a patient, L.M., with bilateral lesions to the suspected human homologue of area MT in monkey visual cortex. Like normals, she is capable of detecting feature targets defined by either unique form or unique movement, across the visual field in parallel. But when she has to detect a conjunction of these two features, her search, unlike that of normals, becomes slow and serial. Her detection time increases by about 160 ms per item, with the slope for No responses roughly twice as steep as that for Yes responses (McLeod, Heywood, Driver, & Zihl, 1989). McLeod et al. demonstrated that her failure to perform this task in parallel, as normals do, is because of her inability to separate moving and stationary stimuli and then exclude the latter from search. Thus, the stationary Xs interfered with her search for the moving X. In other words, one consequence of her damage to the suspected human homologue of MT was that she lacked the ability to filter by movement. Her search for conjunction targets that did not include motion as a feature (e.g., a conjunction of color and form) was within the normal range, showing that the damage to her visual pathway had produced an impairment specific to detecting conjunctions involving motion, not to detecting conjunction targets in general.

Implications for Theories of Visual Attention

Our results show that a number of theories of visual attention, based on experiments with stationary displays, do not generalize successfully to moving displays.

The spatial spotlight. A common metaphor for visual attention has been that of a spatial spotlight (e.g., Eriksen & Eriksen, 1974; LaBerge, 1983; Posner, 1980). The crux of this metaphor is the idea that attention can only be directed to contiguous areas of the visual field. It is clear that attention *can* be directed towards specific areas of space (e.g., Posner, 1980). However, our experiments demonstrate that movement can allow attention to be directed to stimuli in noncontiguous regions of space, to the relative exclusion of interleaved items with a different motion. This conclusion has also been reached by Driver and Baylis (1989) using a variation of the Eriksen and Eriksen response competition paradigm, and by Baylis, Driver, and McLeod (1990) using the illusory conjunction paradigm of Treisman and Schmidt (1982).

Feature integration theory. Our three experiments each demonstrate parallel search for conjunctions of movement and form and are therefore inconsistent with a strong interpretation of Treisman and Gelade (1980), on which conjunction search should always be serial. However, our results can be accommodated by the revised version of feature integration theory, on which information from motion feature maps could "guide" search in the master map, which represents multidimensional information (Treisman, 1988; Wolfe et al., 1989). The results of Experiment 3 and condition Up–Stationary from Experiment 2 suggest that such guidance should be considered as the excitation of target-feature representations on the master map (Wolfe et al.), rather than the inhibition of nontarget features. Were the guidance inhibitory, it should have been able to inhibit downward moving Xs in both Experiment 3 and condition Up–Stationary from Experiment 2, producing parallel search for the target X in both cases. However, parallel search was only observed in the former case. The excitatory account is therefore preferred, although to explain the slow serial search observed in condition Up–Stationary this account has to assume that moving and static feature maps cannot both guide search at the same

time. This corresponds to the central claim of our movement-filter account, namely, that different components of the visual system are specialized for the representation of moving and static stimuli, and they cannot readily be interrogated for their contents at the same time.

Similarity theory (Duncan & Humphreys, 1989). These authors offer an alternative framework for understanding experiments in which subjects search for targets defined by conjunctions of features. They propose that two factors govern the ease of visual search: (a) increasing the similarity of targets and nontargets makes search more difficult, and (b) increasing the similarity of the nontargets to each other makes search easier. These two factors interact to give a continuous gradient of task difficulty. At a general level it seems unnecessary to leave the armchair to be convinced that the two factors identified by Duncan and Humphreys will contribute to the ease of search in the way they suggest. At an empirical level, it is also true that we find continuous variation in task difficulty across tasks (as gauged by the slope of the detection time vs. set-size regressions), which they predict.

However, at a more detailed level, their general proposals about visual search do not provide clear predictions about the relative ease and difficulty of our experimental conditions. The essence of the Duncan and Humphreys (1989) approach is the importance of grouping; if it is easy to group nontargets, and easy to do so without including the target, then search will be easy. But, as we showed in Experiments 1 and 2, grouping in terms of the established principles of common fate does not offer a good explanation of our results for moving displays. In Experiment 1, target detection is easy in a display in which the moving items all move in different directions. In Experiment 2, there seem to be no a priori grounds for predicting that grouping will be easier in any particular condition, because there are three common-fate groups in each, two of which must be searched in both the Up–Down and Up–Stationary conditions. Yet there are large differences between the ease of search in these conditions.

We would not suggest that our results falsify Duncan and Humphreys's (1989) position. From their perspective, our experiments could be considered an investigation into how grouping operates on one visual dimension (motion), showing that in this domain the established Gestalt principle (common fate) does not provide a complete account. We simply note that it was difficult to derive clear predictions from the Duncan and Humphreys (1989) theory and that if the operation of grouping in visual search is revealed only by doing visual-search experiments, then their thesis that grouping determines visual search may be unfalsifiable.

Why do theories that account for the ease of search in stationary displays (e.g., Duncan & Humphreys, 1989; Treisman & Gelade, 1980) fail to generalize successfully to moving displays? One possible explanation is suggested by the organization of the visual processing system. It consists of a series of independent, parallel routes to different areas of the cortex (Cowey, 1985). These routes handle different sorts of information. For example, cells in area V4 are sensitive to wavelength rather than to movement, whereas the opposite is true for cells in area MT. Some independence of these pathways starts even before information reaches V1. The information

that will eventually reach V4, for example, has been processed by different sets of retinal ganglion and lateral geniculate nucleus cells. The cortical processing routes are also different, involving different layers of V1 and V2 (Livingstone & Hubel, 1987). Given the striking independence of these routes physiologically, it would be rather surprising if conclusions based on the processing of static colored displays did extend without qualification to moving monochrome displays. Both Treisman and Gelade (1980) and Duncan and Humphreys (1989) offer general theories of visual attention that are supposed to apply to stimuli in any pathway. Some general principles may exist, but our data suggest that at a detailed level different accounts of visual attention will be needed for the different pathways in the modular primate visual system.

The role of the movement filter. We have identified one component within the visual system, the movement filter. Two of its functions are apparent. First, it allows attention to be directed toward any stimulus in the environment associated with anomalous movement. One biological role for this system is obvious: Predators and prey will frequently be associated with anomalous movement relative to the background. The filter's second function is to segregate different parts of the visual array on the basis of movement cues. A key role for the visual system is to parse the visual array into separate objects (i.e., to work out which parts of the scene go together). Motion cues provide a useful heuristic for solving this problem. Parts of the visual display that move together may well be parts of the same object.

One apparent problem with a movement filter is that it will represent all moving items in a display (as in Experiment 1). Different parts of the visual world will frequently move in different directions because they correspond to distinct objects moving independently. A system that necessarily linked all moving parts of the visual display would be highly misleading. So it is not surprising that attention can be directed to just one direction of motion within the filter, as we showed in Experiment 3. However, parts of the visual display that move independently sometimes are linked by a higher order connection (e.g., the different limbs of an approaching animal, the independent birds of a flock, or the expanding edges of an approaching object). To be prevented from linking such visual events because of their different direction of movement would be inefficient. We demonstrated such linking in our first experiment. But we find the visual system has difficulty in attending simultaneously to stationary and moving parts of the display. This suggests that there may be relatively few naturally occurring visual events of biological significance that require the visual system to attend to static and moving stimuli simultaneously.

References

Albright, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology, 52,* 1106–1130.

Allman, J., Miezin, F., & McGuinness, E. (1985). Direction and velocity specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception, 14,* 105–126.

Baylis, G., Driver, J., & McLeod, P. (1990). *Movement and proximity*

constrain conjunction errors of color and form. Manuscript submitted for publication.

Broadbent, D. E. (1958). *Perception and communication.* London, England: Pergamon Press.

Cowey, A. (1985). Aspects of cortical organization related to selective attention and selective impairments of visual attention. In M. Posner & O. Marin (Eds.), *Attention and performance XI* (pp. 41-62). Hillsdale, NJ: Erlbaum.

Driver, J., & Baylis, G. (1989). Movement and visual attention: The spotlight metaphor breaks down. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 448-456.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113,* 501-517.

Duncan, J., & Humphreys, G. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433-458.

Enroth-Cugell, C., & Robson, J. (1966). The contrast sensitivity of the retinal ganglion cells of the cat. *Journal of Physiology, 187,* 517-552.

Eriksen, B., & Eriksen, C. (1974). Effects of noise-letters on identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16,* 143-149.

Felleman, D., & Kasse, J. (1984). Receptive-field properties of neurons in Middle Temporal Area (MT) of owl monkeys. *Journal of Neurophysiology, 52,* 488-513.

Gibson, J. (1966). *The senses considered as perceptual systems.* Boston: Houghton Mifflin.

Kahneman, D., & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy & J. Pomerantz (Eds.), *Perceptual organization* (pp. 181-211). Hillsdale, NJ: Erlbaum.

LaBerge, D. (1983). Spatial extent of attention to letters and words. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 371-380.

Livingstone, M., & Hubel, D. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement and depth. *The Journal of Neuroscience, 7,* 3416-3468.

McLeod, P., Driver, J., & Crisp, J. (1988). Visual search for a conjunction of movement and form is parallel. *Nature, 332,* 154-155.

McLeod, P., Heywood, C., Driver, J., & Zihl, S. (1989). Selective deficit of visual search in moving displays after extrastriate damage. *Nature, 339,* 466-467.

Nakayama, K., & Silverman, G. (1986). Serial and parallel processing of visual feature conjunctions. *Nature, 320,* 264-265.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.

Posner, M. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3-26.

Prinzmetal, W. (1981). Principles of feature integration in visual perception. *Perception & Psychophysics, 30,* 330-340.

Tolhurst, D. (1973). Separate channels for the analysis of the shape and the movement of a moving visual stimulus. *Journal of Physiology, 231,* 385-402.

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and objects. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 194-214.

Treisman, A. (1988). Features and objects. The 14th 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., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology, 14,* 107-141.

Treisman, A., Sykes, M., & Gelade, G. (1977). Selective attention and stimulus integration. In S. Dornic (Ed.), *Attention and performance VI* (pp. 333-361). Hillsdale, NJ: Erlbaum.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt [Laws of organization in perceptual forms]. *Psychologische Forschung, 4,* 301-350.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 419-433.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 601-621.