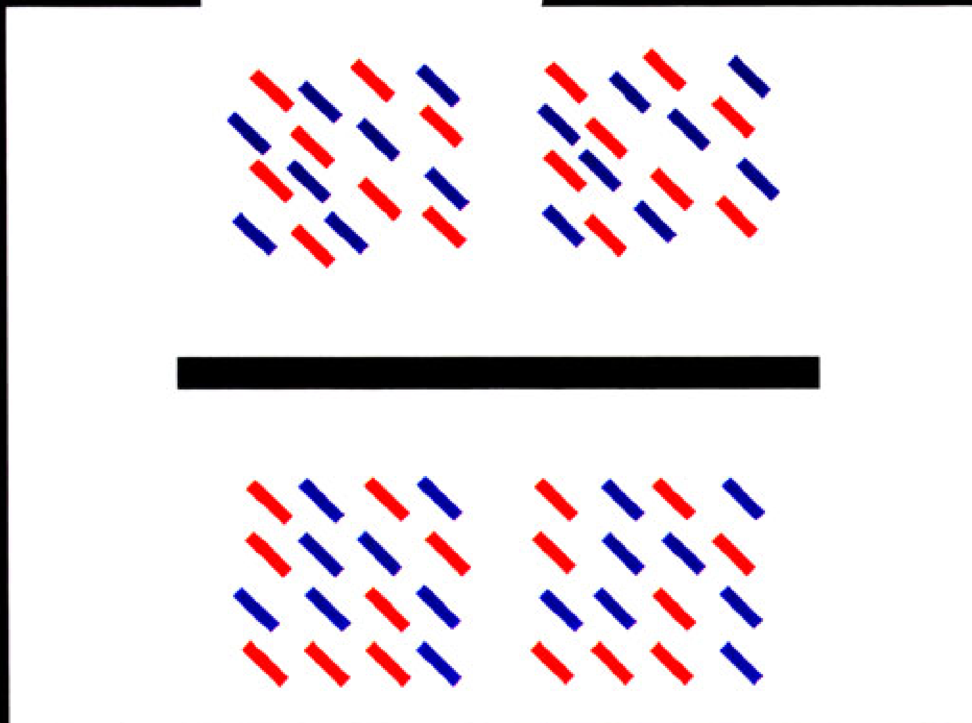
neuroreport

For Rapid Communication of Neuroscience Research

Volume 10 Number 1 18 January 1999 ISSN 0959-4965

Wash. U. Medical Library
22-FEB-99



WASHINGTON UNIVERSITY MEDICAL LIBRARY
3 2201 20235 6578

LIPPINCOTT WILLIAMS & WILKINS

On the cover of this issue:

Stereoscopic demonstration of the role of perceptual grouping in visual search. Two stereograms are shown, one in which perceptual grouping of search elements is allowed (top) and the other in which such grouping is made harder (bottom). When viewed dichoptically, i.e., so that each eye sees a different half of a given stereogram, the two halves of the stereogram fuse to yield a 3-D percept of an array of bars. The bars of a given color are all located in a given depth plane (near or far, depending on the color) except for the target (second row from top, third column from left), which is located in the opposite depth plane given its color. The target is easily visible when distractors in a given depth plane from perceptual groups of the same color. When such grouping is degraded, the target becomes harder to find.

Reproduced by kind permission from J. Hegdé (from the article entitled 'The popout in some conjunction searches is due to perceptual grouping' which appears on pages 143–148 of this issue).

NeuroReport 10, 143–148 (1999)

The popout in some conjunction searches is due to perceptual grouping

Jay Hegdé[1,2,CA] and Daniel J. Felleman[1]

[1]Department of Neurobiology and Anatomy, University of Texas Medical School, Houston, TX 77030; [2]Department of Anatomy and Neurobiology, Washington University School of Medicine, Box 8108, St. Louis, MO 63110, USA

[CA,2]Corresponding Author and Address

THE target in a visual search task usually pops out if it can be distinguished from its background on the basis of only one visual feature but not if the target represents a conjunction of two or more features. However, several recent reports suggest that in certain cases, search targets defined by a conjunction of two features also pop out. We have reinvestigated three pairs of such features to determine whether the popout in these cases can be attributed to perceptual grouping. We find that that in all three cases, popout no longer occurs when perceptual grouping is degraded, suggesting that the popout is the result of perceptual grouping and not of novel mechanism/s of conjunction search. *NeuroReport* 10:143–148 © 1999 Lippincott Williams & Wilkins.

Key words: Binding; Feature conjunction; Perceptual segregation; Set-size effects; Stereoscopic disparity

Introduction

A yellow vertical bar target is easily recognizable, or pops out, from among blue vertical distractors. In general, a visual search target distinguished from the distractors on the basis of a single visual feature typically pops out [1,2]. By contrast, the search for a target defined by a conjunction of two or more features requires an apparently serial and self-terminating scan through the distractors. Thus, the reaction time in a visual search for a yellow vertical target from among blue vertical and yellow horizontal distractors usually increases with the number of distractors. Such a visual search, in which reaction time increases as a function of the number of image elements, is often referred to as a serial search. Correspondingly, popout is often referred to as parallel search, since popout is believed to result from a parallel search of all the image elements together [1–4]. Reaction time of parallel visual search is independent of the number of distractors [1,2]. We will refer to the above model of conditions under which parallel and serial visual searches occur as the standard model.

Several recent reports [5–8], however, have suggested that the standard model has exceptions: in certain cases the conjunction of pairs of features can be searched for in a parallel fashion, rather than serially as the standard model predicts. For instance, targets defined by conjunctions of either stereoscopic disparity and color or stereoscopic disparity and motion pop out [5]. Similarly, conjunction searches of stereoscopic disparity and vernier offset, stereoscopic disparity and orientation [6], or shape and motion [7] are also parallel (see Table 1). We will collectively refer to such instances of parallel conjunction search as exceptions to the standard model.

The exceptions to the standard model are interesting because they suggest that the visual system processes such features differently compared to those features which conform to the standard model. For example, the brain might contain conjunction detectors for some pairs of features but not others [5,7]. This possibility is especially intriguing given the fact that the visual system appears to process different features separately, often in separate visual areas [9–11].

It is, however, possible that the parallel search in case of one or more exceptions to the standard model is facilitated not by anomalous feature extraction *per se*, but by strong perceptual grouping induced by one or both of the visual features.

Table 1. Some reported exceptions to the standard model.

Features	Reference
Stereoscopic disparity and color	5
Stereoscopic disparity and motion	5
Form and motion	7
Stereoscopic disparity and vernier offset	6
Stereoscopic disparity and orientation	6
Any pairing of: Stereoscopic disparity, spatial frequency, size, color and direction of contrast	3
Highly discriminable size, orientation, shape and color	8

Perceptual grouping, or perceptual segregation, refers to the grouping of smaller image elements into a larger image element, such as a line, a surface or an object. For example, a regular rectangular lattice of dots may be seen as an array of vertical columns, horizontal rows or a textured surface, depending on the proximity of the dots to each other. That is, proximity can induce perceptual grouping. Perceptual grouping can also be induced by a variety of other factors, some of which are described by the Gestalt laws of perceptual organization [12]. The effect of perceptual grouping on visual search in general has been documented extensively [13–16]. However, it is not known whether perceptual grouping can help explain specific cases of parallel conjunction search. Indeed, the notion that the exceptions to the standard model, including the ones addressed in our experiments, represent genuinely anomalous conjunction searches is widely accepted [2,8,17–20].

Of the features mentioned above, either stereopsis or motion can induce strong perceptual grouping [13,14,21,22]. For instance, in a visual search for the conjunction of stereoscopic disparity and color [5], the distractors may be grouped according to their stereoscopic depth (e.g. near and far planes) so that all the distractors in each plane can be searched in parallel, resulting in a reaction time that is independent of the number of distractors. If this scenario were true, reaction time should once again increase as a function of the number of distractors if grouping of image elements into depth planes can somehow be made difficult. To test for this possibility, we studied visual searches under conditions which either allowed or prevented perceptual grouping of bars according to their stereoscopic depth planes.

It is important to note that many previous studies do (or appear to) attempt to degrade grouping by randomizing the spatial placement of distractors [5,6]. However, we found in preliminary experiments that this still allows for a small number of perceptual groups each containing a cluster of similar distractors, effectively reducing the number of distractors to the number of perceptual groups (see Fig. 1), thus producing apparently parallel reaction times proportional to the number of perceptual groups allowed, rather than the actual number of distractors. This suggests that random spatial placement of distractors alone does not degrade the grouping sufficiently to rule out perceptual grouping as a factor in producing the parallel reaction times.

Here we report that, for three different pairs of visual features, the conjunction search is parallel when perceptual grouping is allowed, and serial when grouping is effectively prevented. Our results suggest that perceptual grouping is sufficient to explain the parallel searches for these three pairs of features and that it is not necessary to invoke novel mechanisms of feature extraction.

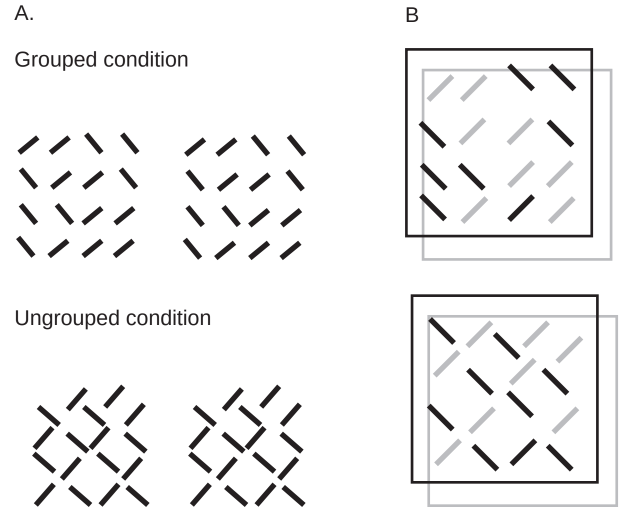


FIG. 1. Schematic illustration of grouped and ungrouped conditions. (**A**) Stereograms similar to those used in the experimental conditions. The images can be fused either convergently or divergently. The percepts of divergent viewers will be reversed in depth from those of convergent viewers, and *vice versa*. The target, a bar which represents a unique conjunction of stereoscopic depth and orientation, is the third bar from left in the bottom row in both the conditions shown. (**B**) Perspective rendering of the stereoscopic percepts resulting from divergent viewing of images in (A) (the near and far planes are drawn in black and gray, respectively). Note that in the grouped percept, the distractors in the near and the far plane perceptually segregate into spatially distinct groups in each plane despite the random depth plane placement of the distractors. The target is easy to find regardless of which of the groups it is located in, since it can be distinguished from its neighbors on the basis of either its depth plane or orientation. By contrast, the two classes of distractors appear intermingled with each other in the ungrouped condition, making the detection of the target harder. See text for additional details.

Materials and Methods

Subjects: Five volunteer subjects, CK, DF, JH, MA and RJ (two females and three males), participated in this study. All but JH (one of the authors) were naive subjects. All the subjects had normal (or corrected to normal) vision. All aspects of a given experiment were identical from subject to subject, except for the stereoscopic hardware, which was adjusted for each subject to allow for the most comfortable binocular fusion. All subjects performed at $>95\%$ in all experiments, except CK, who performed at 78% in one experiment. The results obtained from CK were otherwise indistinguishable from those of the other subjects. We however included an additional subject in the experiment in which CK participated (experiment 3). None of the subjects preferred to receive an auditory feedback about the trial result, although this was offered to them at the beginning of each experiment.

The task: The task consisted of a visual search of an array of bars for a single, known target representing a unique conjunction of a pair of features (see below). The target was embedded among 21, 35, 51 or 65 distractors, depending on the trial. Three experiments were conducted, each involving conjunction search for a different pair of features. In experiment 1, the target was defined by a unique conjunction of a given orientation (45° or 135°) and a given stereoscopic depth (near or far). The distractors shared either the target's orientation or its stereoscopic depth plane, but not both. For instance, in a trial where the target was a vertical bar in the near plane, the distractors were either horizontal bars in the near plane or vertical bars in the far plane. Experiments 2 and 3 were similar to experiment 1 except in that experiment 2 consisted of conjunction search for stereoscopic disparity and color (red or green), while the conjunction search in experiment 3 involved stereoscopic disparity and vernier offset (offset to the left or to the right). The bars were equiluminant with each other (15 cd/m^2) against a dark (< 0.5 cd/m^2) background. They subtended $1.2 \times 0.2°$ of visual arc each and were spaced at a center-to-center distance of 1.6° from each other. We used highly discriminatable visual features in all three of our experiments. Thus, the bars in all the experiments were readily discriminatable from each other and from the background for all subjects.

The bars were arranged in the two depth planes (both normal to the line of gaze) so that perceptual grouping of bars according to their depth planes was either allowed (grouped condition) or degraded (ungrouped condition). In the grouped condition, the distractors were randomly allocated to either depth plane. While this random placement of bars, presumably similar to those used by Nakayama and Silverman [5], degrades the grouping to some extent, it nonetheless allows the formation of a small number of perceptual groups of bars sharing the same depth plane (see Fig. 1). In the ungrouped condition, we alternated the depth planes of bars such that the bars adjacent to a given bar in a given depth plane tended to be in the opposite depth plane. This alternation of depth planes from one bar to the next tended to substantially degrade the perceptual grouping. To further degrade perceptual grouping, we added a slight spatial jitter to ungrouped stimuli (up to ~10% of the interbar distance). Corresponding pairs of grouped and ungrouped conditions were identical in all other respects. Trials were randomly interleaved. In order to avoid lateral masking effects, we allowed the subjects to move their eyes freely throughout the experiment.

Prior to each experiment, the subjects were told the features of the search target, following which they underwent practice trials to familiarize themselves with the experiment. The results from, and the stimuli used in, the practice trials were discarded. Each trial began with the onset of a prestimulus stereogram. After the subjects achieved stereoscopic fusion using this stereogram, they pressed a button to indicate readiness for the trial. The prestimulus stereogram was then turned off and the search stimulus presented at the same stereoscopic disparity as the prestimulus stereogram. The subjects were required to search for the target and to press a button when it was found. The subjects pressed one button with one hand if they found the target in one depth plane and a different button with the other hand to report the target in the opposite plane. The experiments and the data collection were computer controlled.

Twenty percent of the trials were null (catch) trials, where the stimulus did not include the target. Half the catch trials represented the grouped condition, while the other half represented the ungrouped condition.

Results

The experiments which originally reported the parallel reaction times in the three cases we investigate differed slightly from each other [5,6] and from our experiments. We therefore sought to reproduce the parallel searches in our experimental setup, so that the parallel reaction times obtained under grouped conditions can be compared directly with the reaction times under ungrouped conditions. The reaction times in grouped conditions showed that when grouping was allowed (i.e. not fully degraded), the searches were indeed parallel. In all three experiments for all subjects, the reaction times stay essentially constant as the number of distractors increases when perceptual grouping is allowed (grouped conditions, Fig. 2). The average slope of reaction times in grouped conditions averaged over all the experiments and all the subjects is 0.97 ms/distractor. Thus, we were able to reliably reproduce the three reported parallel searches given our experimental setup and trial design. However, when perceptual grouping is degraded, i.e. in the ungrouped condition, reaction times steadily increase as a function of the number of distractors, even though all other experimental parameters are identical to the grouped condition. Reaction times increase at an average rate of 76.81 ms/distractor (70.79 ms/distractor in experiment 1, 74.60 ms/distractor in experiment 2 and 82.98 ms/distractor in experiment 3; see Table 2 for details).

Comparing reaction time slopes, however, is not a particularly reliable method of comparing reaction

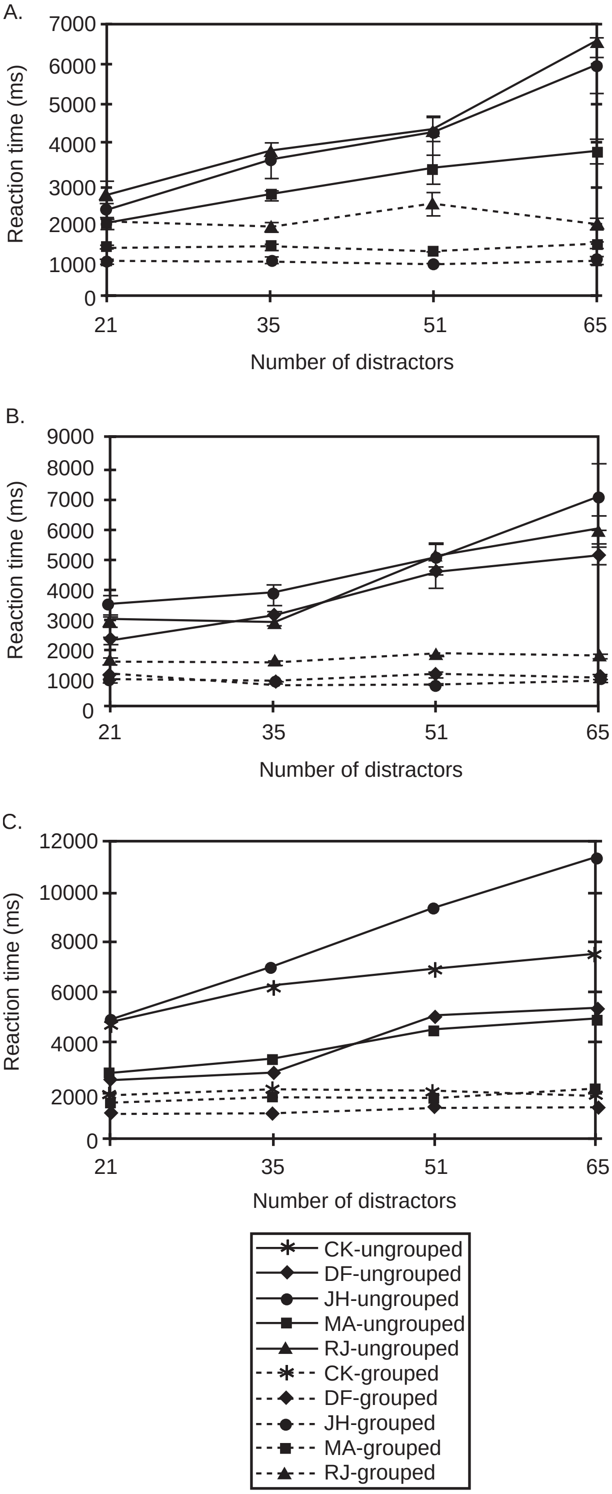


FIG. 2. Reaction times as a function the number of distractors in grouped and ungrouped conditions. Average reaction times ($\pm$s.e.) in visual searches for a conjunction of stereoscopic disparity and orientation (**A**), stereoscopic disparity and color (**B**) and stereoscopic disparity and vernier offset (**C**) are shown. The search stimuli were designed so that perceptual grouping of bars according to their stereoscopic depth planes was either allowed (grouped condition) or prevented (ungrouped condition). Each data point represents the reaction time averaged over 12 correct trials. Reaction times in the ungrouped condition are significantly greater than those in the corresponding ungrouped condition ($p < 0.01$; except for subject RJ in panel A with 21 distractors, where $0.01 < p < 0.05$).

times. Reaction time in visual search tasks can vary from one experimental setup to another for a variety of reasons not directly related to feature conjunction, such as the discriminatability of the features themselves or the overall difficulty of the task [1,15]. We therefore compared reaction times in grouped and ungrouped conditions using a two-way ANOVA (number of distractors $\times$ perceptual grouping). In each of the three experiments, the pooled reaction times from all subjects in grouped conditions were significantly different from those in ungrouped conditions (The p values for the grouping factor were 8.67^{-11}, 2.68^{-10} and 3.12^{-10} in experiments 1, 2 and 3, respectively. ANOVA results for the grouping factor for individual subjects in each experiment are listed in Table 2).

Reaction times can sometimes increase as a function of the number of distractors even in searches that are otherwise parallel [1,14,15]. The largest and the fewest number of distractors we used were 65 and 21, respectively. While these numbers are comparable to those used in the specific earlier reports revisited here [5,6], they are larger than some others [1,3]. Thus, it is possible the reaction times in our experiments were an artefact of the number of distractors we used. However, we believe this is not the case for two reasons. First, in grouped conditions, the reaction times stay more or less unchanged regardless of the number of distractors for all subjects in all experiments (see slopes for grouped conditions in Table 2). Second, the reaction time in every given grouped condition is lower than that in the corresponding ungrouped condition for every subject, experiment and the number of distractors used (one tailed t-tests, $p < 0.01$), with the exception of subject RJ in experiment 1 with 21 distractors ($0.01 < p < 0.05$ in this case). We therefore conclude that the size of the distractor set did not by itself cause an increase in the reaction times in ungrouped conditions.

Discussion

Exceptions to the standard model raise the possibility that certain conjunctions of features may be processed anomalously by the visual system. However, our results provide direct evidence that the parallel reaction times observed in certain conjunction searches are not an intrinsic property of the conjunctions themselves but instead are due to perceptual grouping. While allowing perceptual grouping allows parallel search, degrading perceptual grouping similarly degrades parallel search and restores serial search in otherwise identical stimuli. If the parallel search were a property of the conjunctions themselves, we would not have been able to

Table 2. Comparison of reaction times in grouped and ungrouped conditions. Results for the grouping factor from a two-way ANOVA (number of distractors × perceptual grouping) are also shown.

Experiment (conjunction)	Subject	Slope (ms/item)		ANOVA	
		Grouped	Ungrouped	F(1,3)	*p*
1 (Stereo and orientation)	JH	−0.42	80.04	61.92	1.84^{-08}
	MA	−0.03	42.18	26.43	2.087^{-05}
	RJ	1.13	90.15	26.82	1.883^{-05}
2 (Stereo and color)	DF	−1.95	66.96	24.25	3.729^{-05}
	JH	−1.60	80.21	119.36	2.058^{-11}
	RJ	3.55	76.62	33.86	3.415^{-06}
3 (Stereo and vernier offset)	CK	−3.69	60.86	64.27	1.291^{-08}
	DF	5.23	72.31	24.88	3.142^{-05}
	JH	−0.01	147.41	146.89	1.977^{-12}
	MA	7.49	51.34	29.32	1.002^{-05}

restore serial search by degrading perceptual grouping (while leaving all other attributes of the stimuli intact). The fact that we were able to reproduce the reported parallel reaction times suggests that the serial reaction times in the ungrouped conditions were not an artefact of our experimental setup or trial design. Taken together, these findings suggest that perceptual grouping is necessary and sufficient to explain the reported parallel searches for all the three pairs of features we tested: stereoscopic disparity and color, stereoscopic disparity and orientation and stereoscopic disparity and vernier offset.

Thus, the three 'exceptions' to the standard model are not exceptions at all: they are artefacts of the perceptual grouping in as much as they conform to the standard model when perceptual grouping is well controlled. The susceptibility to perceptual grouping itself is not a distinguishing attribute of these features either; features previously known to conform to the standard model can also be subject to perceptual grouping in visual searches [2,16]. Thus, no basis exists on which the three pairs of features we have investigated can be considered to be violations of the standard model.

Our results also demonstrate the importance of adequately controlling for perceptual grouping in visual search tasks. Grouping induced by some features is harder to degrade than that induced by others. It is especially difficult to prevent grouping on the basis of stereoscopic depth planes; we found that allowing even a small amount of random groupings, presumably similar to those in the displays of Nakayama and Silverman [5], tended to produce parallel searches.

Perceptual grouping and visual search: The role of perceptual grouping in visual searches in general has been examined by several others previously [1,3,14,16], although none addressed the question of whether perceptual grouping can help explain parallel reaction times in specific conjunction searches. Displays that allow parallel conjunction searches are often those in which non-targets 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 are spatially grouped together [16].

Treisman and Sato [3] also found that perceptual grouping affects visual search reaction times. While they did not specifically address the reported exceptions to the standard model, they explored the feasibility of three general explanations for a parallel reaction time in a conjunction search: perceptual segregation, specific conjunction detectors and inhibition controlled separately by two or more distractor features. Since the conjunction detector hypotheses seemed inconsistent with the neurophysiology of single units, they proposed that a combination of perceptual segregation and feature inhibition might explain parallel conjunction searches. However, regardless of the relevance of feature inhibition to visual perception in general and visual search in particular, our results show that perceptual segregation alone is sufficient to explain the reaction times in at least three different cases of parallel conjunction search.

Other exceptions to the standard model: Can perceptual grouping account for other parallel conjunction searches listed in Table 1? While we have not investigated these conjunctions, we believe such a scenario is possible in one or more of the cases. Of particular interest to us is the parallel search for form–motion conjunction [7], since motion is an important visual feature. Motion is known to induce strong perceptual grouping effects [14,21,22]. McLeod *et al.* [14] investigated the possibility that perceptual grouping based on the direction of motion, or common fate, might account for the parallel reaction time, and concluded that such grouping cannot account for the popout. However, the possibility remains that their stimuli might have been subject to perceptual grouping on bases other than the direction of motion. For instance, moving elements may be grouped on the basis of similarity of form [22]. Image elements in motion stimuli can also be segregated on the basis of whether they are moving or stationary. Such grouping according to mobility is the basis of motion transparency [23,24].

We emphasize, however, that it is conceivable that form-motion conjunction as well as other conjunctions listed in Table 1 will turn out not to conform to the standard model even when perceptual grouping can be ruled out as a contributing factor.

Conclusions

Our results provide direct evidence that conjunctions of stereoscopic disparity–color, stereoscopic disparity-orientation, and stereoscopic disparity–vernier offset are likely to be processed in a manner similar to other features by the visual system. The parallel reaction times in visual searches for these conjunctions reported earlier can be accounted for by perceptual grouping. Thus, the standard model holds for at least three pairs of features which were hitherto thought to be exceptions to the model.

References

1. Treisman A. *Q J Exp Psychol* **40A**, 201–237 (1988).
2. Treisman A. The perception of features and objects. In: Baddeley A and Weiskrantz L, eds. *Attention: Selection, Awareness, and Control.* Oxford: Clarendon Press, 1993: 1–34.
3. Treisman A and Sato S. *J Exp Psychol Hum Percept Perform* **16**, 459–478 (1990).
4. Julesz B. *Trends Neurosci* **7**, 41–48 (1984).
5. Nakayama K and Silverman GH. *Nature* **320**, 264–265 (1986).
6. Steinman SB. *Perception* **16**, 389–398 (1987).
7. McLeod P, Driver J and Crisp J. *Nature* **332**, 154–155 (1988).
8. Wolfe JM, Cave KR and Franzel SL. *J Exp Psychol Hum Percept Perform* **15**, 419–433 (1989).
9. Livingstone M and Hubel D. *Science* **240**, 740–749 (1988).
10. Van Essen DC and Gallant JL. *Neuron* **13**, 1–10 (1994).
11. Vaina LM. *Cerebr Cortex* **4**, 555–572 (1994).
12. Bruce V and Green PR. *Visual Perception.* London: Lawrence Erlbaum Associates. 1990: 110–128.
13. Nakayama K. *Proc Natl Acad Sci USA* **93**, 634–639 (1996).
14. McLeod P, Driver J and Crisp J. *J Exp Psychol Hum Percept Perform* **17**, 55–64 (1991).
15. Luck SJ. *Curr Opin Neurobiol* **4**, 183–188 (1994).
16. Treisman A. *J Exp Psychol Hum Percept Perform* **8**, 194–214 (1982).
17. Kinchla RA. *Annu Rev Psychol* **43**, 711–742 (1992).
18. Treisman A. *Curr Opin Neurobiol* **6**, 171–178 (1996) .
19. Humphreys GW and Bruce V. *Visual Cognition.* London: Lawrence Erlbaum Associates. 1989: 183–186.
20. Banks WP and Krajicek D. *Annu Rev Psychol* **42**, 305–331 (1991).
21. Ahlstrom U. *Scand J Psychol* **36**, 343–354 (1995).
22. Kramer P and Yantis S. *Percept Psychophys* **59**, 87–99 (1997).
23. Mulligan JB. *Vision Res* **33**, 2021–2030 (1993).
24. McOwan PW and Johnston A. *Curr Biol* **6**, 1343–1346 (1996).

ACKNOWLEDGEMENTS: This work was supported in part by a Fight for Sight Postdoctoral Fellowship to J.H., by NIH EY 08372 to DJF and by NEI Core Grant 1P30-EY10608. We thank Drs James Makous, John Maunsell, Anne Sereno and Lawrence Snyder for helpful suggestions and Dr Robert Desimone for data collection software.

Received 15 October 1998; accepted 29 October 1998