Research Article

Figure-Ground Segmentation Can Occur Without Attention

Ruth Kimchi[1] **and Mary A. Peterson**[2]

[1]*University of Haifa and* [2]*University of Arizona*

ABSTRACT—*The question of whether or not figure-ground segmentation can occur without attention is unresolved. Early theorists assumed it can, but the evidence is scant and open to alternative interpretations. Recent research indicating that attention can influence figure-ground segmentation raises the question anew. We examined this issue by asking participants to perform a demanding change-detection task on a small matrix presented on a task-irrelevant scene of alternating regions organized into figures and grounds by convexity. Independently of any change in the matrix, the figure-ground organization of the scene changed or remained the same. Changes in scene organization produced congruency effects on target-change judgments, even though, when probed with surprise questions, participants could report neither the figure-ground status of the region on which the matrix appeared nor any change in that status. When attending to the scene, participants reported figure-ground status and changes to it highly accurately. These results clearly demonstrate that figure-ground segmentation can occur without focal attention.*

Figure-ground segmentation is the process by which the visual system organizes a visual scene into figures and their backgrounds. This is one of the most important visual processes because figure-ground distinctions are fundamental to the visual perception of objects and to visuomotor behavior.

Gestalt psychologists, who were the first to recognize the importance of figure-ground segmentation, distinguished figures and grounds in terms of their phenomenal appearance (Koffka, 1935; Rubin, 1915/1958). Figures appear to have a definite shape, so that their bounding contours are assigned as belonging to them. Grounds are shapeless near the contours they share with figures and appear to continue behind the figures near those contours. Much of the research on figure-ground perception has been concerned with identifying the properties that determine which regions will appear as figures. For example, smaller regions are likely to be perceived as figures (Rubin, 1915/1958), as are symmetrical regions (Bahnsen, 1928), convex regions (Hoffman & Singh, 1997; Kanizsa & Gerbino, 1976), regions with higher spatial frequency (Klymenko & Weisstein, 1986), lower regions (Vecera, Vogel, & Woodman, 2002), regions with a wide base (Hulleman & Humphreys, 2004), and regions depicting familiar objects (Peterson & Gibson, 1994a, 1994b).

An important, yet unresolved, issue concerns the relation between figure-ground segmentation and attention. The study reported in this article addressed one aspect of this issue: Can figure-ground segmentation occur without attention?

Many modern theories of perception have assumed that figure-ground segmentation operates preattentively to deliver the perceptual units to which focal attention is allocated for further processing (e.g., Julesz, 1984; Marr, 1982; Neisser, 1967; Treisman, 1986). Although this view has been widely accepted, researchers have also suggested that deliberate attention (Koffka, 1935; Rubin, 1915/1958) and the location of fixation or spatial attention (Hochberg, 1971; Sejnowski & Hinton, 1987) can influence figure-ground organization. Few studies, however, have directly examined the relation between figure-ground segmentation and visual attention. Peterson and Gibson (1994b) showed that fixation location can contribute to figure-ground segmentation. Baylis and Driver (1995; Driver & Baylis, 1996) examined performance on a contour-matching task with ambiguous displays and showed that endogenous attention influenced figure-ground assignment; their experiments suggested that exogenous attention did not influence figure-ground perception. However, recent research by Vecera, Flevaris, and Filapek (2004), using the same contour-matching task with similar ambiguous displays, demonstrated that exogenous attention can influence figure-ground assignment, provided that the exogenous cues are located inside the figure-ground display. These results demonstrate that exogenous spatial attention can act as a cue for figure-ground assignment, but do not speak to the

Address correspondence to Ruth Kimchi, Department of Psychology, University of Haifa, Haifa 31905, Israel, e-mail: rkimchi@research.haifa.ac.il.

question of whether or not attention is required for figure-ground segmentation to occur.

Recently, Nelson and Palmer (2007) examined the effects of figural cues (i.e., familiarity) on attention. They used bipartite displays in which the central contour sketched a familiar shape on one side but not the other, and presented detection-discrimination targets equally often on the two sides of the central contour. They found a perceptual advantage for targets presented on the figure, a result suggesting that a figural cue can attract attention to the region that is biased to be perceived as figure. However, it cannot be determined from Nelson and Palmer's experiments whether the figural advantage was due to direct influence of the figural cue on attention or to attraction of attention to the region perceived as figure (see also Weisstein & Wong, 1987). Nevertheless, the possibility that figural cues per se can attract attention and that exogenous attention can influence figure-ground assignment (Vecera et al., 2004) suggests a potential reinterpretation of research by Driver, Baylis, and Rafal (1992) that is often cited (e.g., Barenholtz & Feldman, 2006; Mazza, Turatto, & Umiltà, 2005) as providing evidence that figure-ground segmentation is preattentive.

Driver et al. (1992) studied a patient with right-hemisphere damage and severe left neglect. The patient was presented with a display divided by a contour into a small, bright, green section and a larger, dimmer, red section; the green section could be on the far left or the far right of the display. His task was to decide whether the dividing contour matched a probe line. The patient performed well above chance when the green section appeared at the far left, that is, when the dividing contour fell to the right of the green section, although the contour appeared in his contralesional field, but he performed at chance when the green section appeared at the far right, that is, when the dividing contour fell to the left of the green section, although the contour appeared in his ipsilesional field. According to Driver et al., these results indicate that the patient retained intact figure-ground segmentation despite his pathological bias in spatial attention and thus imply that figure-ground segmentation is preattentive. Although these results clearly indicate that the patient's neglect was applied to the contralesional side of the green figure rather than to the contralesional field as a whole, the possibility that exogenous attention influenced the figural status of the green section cannot be ruled out. The green section was more salient than the red section because it was brighter. Given that color perception can be preserved in the contralesional field in some patients (Cohen & Rafal, 1991), attention could have been drawn automatically to the green section by virtue of its salience, thereby increasing its likelihood of being perceived as figure, much as exogenous cues can influence figure-ground assignment (Vecera et al., 2004). Thus, despite the ingenuity of this study, it does not provide unequivocal evidence that figure-ground segmentation can occur without attention.

In this article, we report new evidence demonstrating that figure-ground segmentation can occur for unattended stimuli. This evidence was obtained using an inattention paradigm with indirect on-line measures of unattended processing (devised by Russell & Driver, 2005; see also Kimchi & Razpurker-Apfeld, 2004).

Observers were presented with two successive displays, each of which included a small target matrix (made up of random black and white squares) that appeared on a task-irrelevant scene of alternating regions organized into figures and grounds by convexity. The task was to judge whether the matrices in the two displays were the same or different. When the matrices differed, only one black square changed its location, rendering the task sufficiently demanding to absorb attention. The figure-ground organization of the scene backdrop stayed the same or changed across the two displays, independently of whether or not the target matrix changed. The edges in the backdrop always changed from the first to the second display regardless of whether or not the figure-ground organization changed, to control for the possibility that a change in backdrop organization could be detected from local changes in edges per se. We examined whether the figure-ground organization of the scene backdrop influenced performance on the matrix-change task. We hypothesized that if the unattended backdrop was segmented into figures and grounds, then congruency effects would be obtained; that is, responses to *same* targets would be faster or more accurate when the backdrop organization stayed the same than when it changed, and responses to *different* targets would be faster or more accurate when the backdrop organization changed than when it stayed the same. After the last experimental trial, observers were probed with surprise questions asking whether the region on which the target was presented in the preceding display appeared to be figure or ground and whether the figure-ground status of that region had changed between the two displays on that trial.

In Experiment 1, we found that changes in the figure-ground organization of the backdrop produced congruency effects on performance of the target-change task, even though accuracy in reporting these changes was no better than chance. In Experiment 2, we instructed participants to attend to the scene backdrops and ignore the matrices; in this case, explicit reports about the figure-ground organization of the backdrop were highly accurate.

EXPERIMENT 1

Method

Participants

Forty-six students at the University of Haifa (39 females, 7 males; age range: 19–28 years) participated in this experiment. All had normal or corrected-to-normal vision. Four observers performed at chance level in at least one of the conditions, and they were replaced.

Stimuli

Each display consisted of a small target matrix presented on a scene of alternating regions organized into figures and grounds by local convexity (Fig. 1). The displays were presented on a gray field at a viewing distance of 60 cm. Each target was made up of 12 black and 13 white small (0.19°) squares, randomly located in a 5 × 5 matrix subtending 0.95° × 0.95°.

Each trial consisted of two successive displays. The target sometimes changed from the first to the second display (different-target trials) and sometimes remained the same (same-target trials). A change was made by switching the location of one small black square in the matrix with that of a white one.

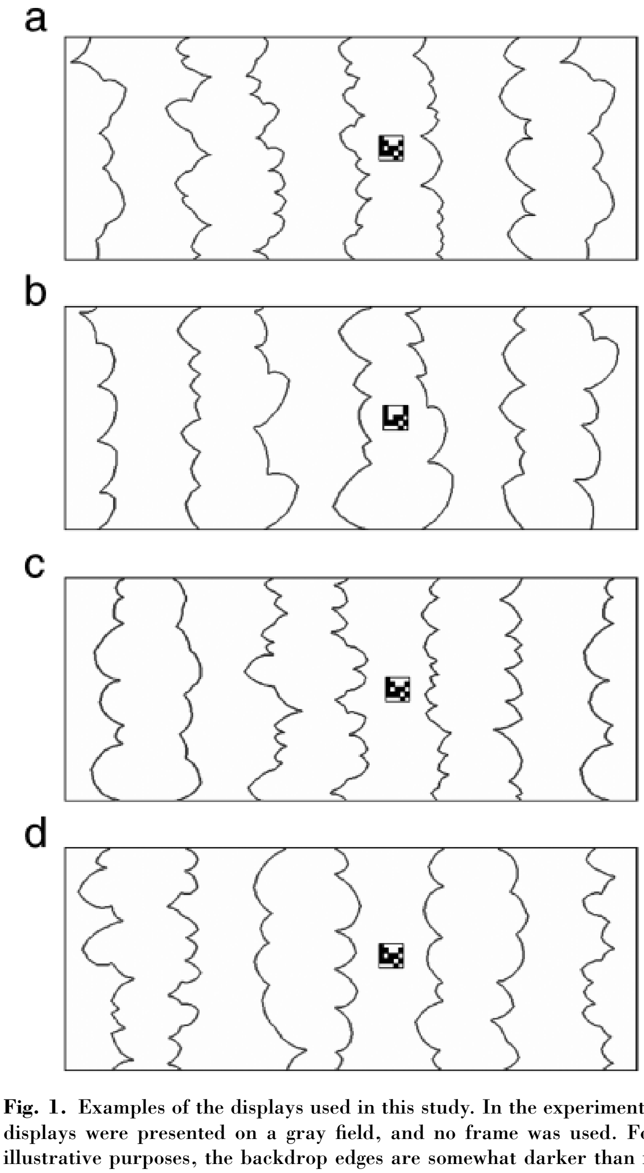

Fig. 1. Examples of the displays used in this study. In the experiments, displays were presented on a gray field, and no frame was used. For illustrative purposes, the backdrop edges are somewhat darker than in the actual stimuli. The matrix always appeared on the backdrop region to the right of the central edge (i.e., the fifth region from the left). This region could be convex (figure, or F) or concave (ground, or G), and the number of parts in this region could be small or large. The examples here illustrate (a) the F type with a large part number, (b) the F type with a small part number, (c) the G type with a large part number, and (d) the G type with a small part number. The matrices in (a) and (b) depict an example of a change in matrix (a change in the location of one small black square).

The backdrop stimuli were chosen from the set of eight-region outline displays created by Kim and Peterson (2002; for details, see Peterson & Salvagio, 2008). Each backdrop stimulus subtended 7.31° in height and 18.43° in width and consisted of four locally convex regions (i.e., regions with multiple convex parts) alternating with four locally concave regions (i.e., regions with multiple concave parts). The convex and concave regions in a display were equal in area; no two regions, whether within or across displays, were the same shape. The matrix in each display was presented on the backdrop region to the right of the central edge (i.e., the fifth region from the left). We included 20 different backdrop stimuli in which the fifth region was convex (figure-type backdrops, or F; see Figs. 1a and 1b) and 20 different backdrop stimuli in which this region was concave (ground-type backdrops, or G; see Figs. 1c and 1d). In half of the backdrops of each type, the fifth region had a relatively large number of parts (8–13; Figs. 1a and 1c), and in the other half, this region had a relatively small number of parts (3–7; Figs. 1b and 1d).

These 40 backdrop stimuli were randomly paired (with the constraint that a backdrop with a small number of parts was paired with a backdrop with a large number of parts) so as to produce 20 pairs of each of four types (the first letter denotes the first backdrop type, and the second letter denotes the second backdrop type): FF, GG, FG, and GF. Eighty additional pairs were produced by reversing the order of the stimuli in each pair (reversed FG and GF pairs turned into GF and FG pairs, respectively). Altogether, there were 160 pairs, 40 pairs of each type; in half of the pairs of each pair type, the first stimulus had a small number of parts, and in the other half, the first stimulus had a large number of parts. Each individual backdrop stimulus was repeated eight times. For the practice trials, 8 additional backdrop stimuli (4 F, 4 G) were paired to produce 16 different pairs, 4 of each type.

Design and Procedure

The participants completed 160 experimental trials in two blocks of 80 trials each, preceded by one practice block of 16 trials. A 2 (target: same, different) × 2 (backdrop organization: same, different) × 2 (starting backdrop organization: F, G) within-subjects design was used, producing eight different conditions. Half of the trials were same-target trials, and half were different-target trials. Independently of whether the target changed or remained the same on each trial, the figure-ground organization of the scene backdrop also changed or remained the same. Half of the same-backdrop trials were FF trials, and half were GG trials; half of the different-backdrop trials were FG trials, and half were GF trials. The order of the trials within each block was randomly permuted. The fixation cross was always aligned with the center of the fifth region of the backdrop to be presented, and the target was always centered in the place where the fixation cross had been.

The event structure of each trial is shown in Figure 2. Each trial started with a fixation cross that appeared for 750 ms. After a 250-ms interval, the first display appeared for 200 ms. It was followed by a 150-ms interval, and then the second display appeared for 200 ms. At this point, participants had to decide, as rapidly and as accurately as possible, whether the two successive targets were the same or different. They indicated their decision by pressing one of two response keys. An auditory tone provided immediate feedback after an incorrect response. The intertrial interval was 1,000 ms.

Immediately after participants completed the last experimental trial, they were asked two forced-choice questions. The first question asked, "Did the matrix in the previous display appear to lie on a shape bounded by black borders or on a space between shapes?" An example was presented for each alternative ("shape" or "space between shapes"), and participants indicated their choice by pressing one of two response keys. The second question asked, "Was there a change in the region the matrix appeared to lie on across the two displays in the previous trial (from a shape to a space between shapes or from a space between shapes to a shape)?" The two alternatives were "change" and "no change," and participants indicated their choice by pressing one of two response keys.

Results and Discussion

On-Line Performance on the Matrix Task

All reaction time (RT) summaries and analyses are based on participants' mean RTs for correct responses. RTs less than 150 ms and greater than 1,500 ms were discarded (2.04% of all trials). Table 1 presents mean RT, mean percentage correct, and mean inverse-efficiency (IE) score for each type of trial. The IE score was determined for each condition for each participant by dividing the mean RT by the proportion correct for that condition (Townsend & Ashby, 1983). Because analyses showed that the backdrop significantly influenced RTs for different-target trials

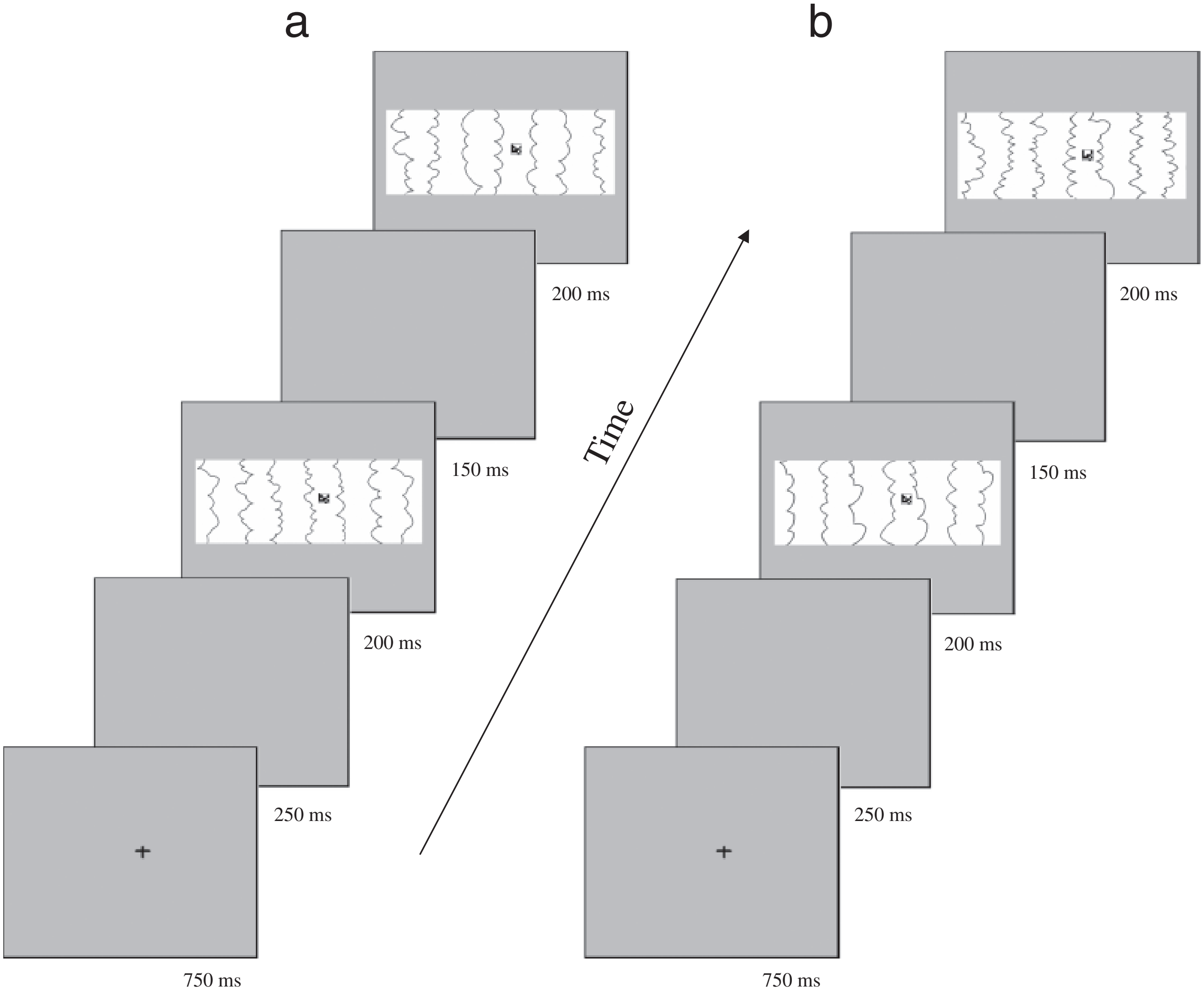

Fig. 2. Sequence of events in a trial. The illustration depicts two examples: (a) a same-target trial (matrix is unchanged) on a backdrop that changes from figure to ground and (b) a different-target trial (matrix changes) on a backdrop that stays figure.

TABLE 1

Mean Reaction Time (RT) on Correct Trials (in Milliseconds), Mean Percentage of Correct Responses, and Mean Inverse Efficiency in Experiment 1

	Backdrop organization			
	Same		Different	
Target	FF trials	GG trials	FG trials	GF trials
	Mean correct RT			
Same	589	591	585	604
Different	595	584	563	583
	Mean percentage correct			
Same	90	92	90	89
Different	89	88	90	89
	Mean inverse efficiency			
Same	665	646	662	695
Different	681	679	636	666

Note. Inverse efficiency was calculated by dividing the mean RT by the proportion correct for each participant for each condition. The abbreviations for the trial types indicate the types of backdrops (F = figure type; G = ground type) and their order across the two successive displays.

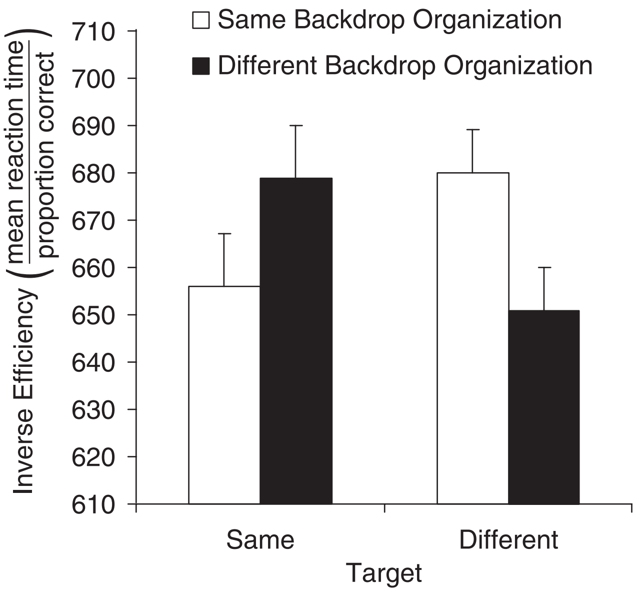

Fig. 3. Results from Experiment 1: inverse-efficiency scores for same and different targets as a function of the backdrop's organization (same, different). Error bars indicate standard errors of the means.

and accuracy for same-target trials, we focused on IE, which combines speed and accuracy, and therefore allowed us to evaluate congruency effects with a single measure.[1]

Figure 3 depicts mean IE scores for same and different targets as a function of backdrop organization (same, different). These results show congruency effects arising from changes in the figure-ground organization of the backdrops: On different-target trials, judgments were more efficient (i.e., IE scores were lower) when backdrop organization changed across the two displays than when it remained the same, and on same-target trials, judgments were more efficient when backdrop organization stayed the same than when it changed. A 2 (target) × 2 (backdrop organization) × 2 (starting backdrop organization) analysis of variance confirmed these results. The interaction between target and backdrop organization was significant, $F(1, 45) = 13.90$, $p < .0005$, $\eta_p^2 = .24$, and did not vary with starting organization, $F < 1$. Analysis of simple effects showed that responses to different targets were significantly more efficient when backdrop organization was changed than when it was unchanged, $F(1, 45) = 10.47$, $p < .005$, $\eta_p^2 = .19$, and responses to same targets were significantly more efficient when backdrop organization was unchanged than when it was changed, $F(1, 45) = 4.47$, $p < .05$, $\eta_p^2 = .09$.

The only other significant result was an interaction between starting organization and backdrop organization, $F(1, 45) = 6.81$, $p < .02$, $\eta_p^2 = .13$. Analysis of simple effects revealed two important results. First, performance was more efficient on FG trials (IE = 649) than on GF trials (IE = 681), $F(1, 45) = 9.13$, $p < .005$, $\eta_p^2 = .17$. The FG and GF trials differed in an important way: On the GF trials, the backdrop region on which the matrix appeared changed from ground to figure, so that a new figure (a "new object") appeared in the target's backdrop region; no new figure appeared in this region on FG trials. Presumably, the implicit processing of a new figure on the GF trials produced less efficient responses to the target. This result indicates that changes in figure-ground organization, and not simply changes in convexity versus concavity of the backdrop regions, were registered by the visual system. Changes in convexity/concavity per se would not predict a difference between these two types of trials, because in both types convex and concave regions changed their location across successive displays. The direction of change would have mattered only if the convex regions were designated as figures, such that a new figure appeared when the backdrop region on which the target appeared changed from concave to convex (GF trials), but not when it changed from convex to concave (FG trials).

Second, performance was equally efficient on FF (IE = 673) and GG (IE = 663) trials, $F < 1$. This result indicates that the congruency effects produced by backdrop organization could not have been due to implicit capturing of attention by backdrop convexity, as per Nelson and Palmer's (2007) suggestion that a figural cue can attract attention automatically. Had attention

---

[1] Analyses of variance (Target × Backdrop Organization × Starting Organization) conducted on correct RT and accuracy yielded a significant Target × Backdrop Organization interaction, $F(1, 45) = 7.03$, $p < .02$, $\eta_p^2 = .14$, for RT and $F(1, 45) = 5.96$, $p < .02$, $\eta_p^2 = .12$, for accuracy; there was no significant three-way interaction, $F < 1$ for RT and $F(1, 45) = 1.06$, $p > .30$, for accuracy. Responses on different-target trials were significantly faster when backdrop organization was changed than when it was unchanged, $F(1, 45) = 8.52$, $p < .01$, $\eta_p^2 = .16$, and responses on same-target trials were significantly more accurate when backdrop organization was unchanged than when it was changed, $F(1, 45) = 4.73$, $p < .05$, $\eta_p^2 = .10$.

TABLE 2

Percentage of Participants Who Responded Correctly to Each Forced-Choice Question in Experiment 1

	Backdrop organization in last trial			
Forced-choice question	FF	GG	FG	GF
Type of region	73 (8/11)	25 (3/12)	25 (3/12)	64 (7/11)
Change in region	45 (5/11)	58 (7/12)	42 (5/12)	55 (6/11)

Note. The abbreviations for the trial types indicate the types of backdrops (F = figure type; G = ground type) and their order across the two successive displays. The numbers in parentheses indicate the number of participants who responded correctly.

been captured by backdrop convexity, then performance efficiency on the matrix task should have differed between trials on which the matrix appeared on convex regions (FF trials) and trials on which it appeared on concave regions (GG trials), but no such difference was observed.[2]

Response to Surprise Questions

Table 2 presents the percentage of participants who responded correctly to each surprise question. Overall, only 21 participants (46%) correctly reported the region on which the target appeared in the preceding display; this percentage was not different from chance. Of the 22 participants who were presented with a figure on the last display (FF or GF), 15 (68%) reported seeing a "shape," $\chi^2(1) = 2.9$, n.s. Of the 24 participants who were presented with a ground (GG or FG), only 6 (25%) reported seeing a "space between shapes," $\chi^2(1) = 6.0$, $p < .025$; this finding suggests some bias to respond "shape."

Only 23 participants (50%) correctly reported whether or not the figure-ground status of the region had changed on the preceding trial; no bias was detected in change responses.

These results show that participants performed at chance in reporting the figure-ground status of the region on which the target appeared in the preceding display and in reporting whether its figure-ground status had changed during the trial. An informal postexperimental debriefing further revealed that participants in this experiment were unaware of the nature of the backdrop scene and of changes in its organization.

The observed congruency effects arising from changes in the backdrop's figure-ground organization suggest that figure-ground segmentation can occur under conditions that satisfy criteria for inattention (Mack & Rock, 1998; Moore, Grosjean, & Lleras, 2003). The target-change task was sufficiently demanding to absorb attention (mean accuracy = 89.5%), the backdrop's figure-ground organization was irrelevant to the task, and participants performed poorly in reporting the figure-ground status of the backdrop or any change in it; these considerations strongly suggest inattentional blindness to the scene backdrop. In the next experiment, we examined whether the figure-ground organization of the backdrop stimuli could be perceived when attention was allocated to them.

EXPERIMENT 2

Method

Participants

Twenty-three new individuals (18 females, 5 males; age range: 19–27 years) participated in this experiment.

Stimuli, Design, and Procedure

The stimuli, design, and procedure were the same as in Experiment 1, except that participants were instructed to attend to the region on which the matrix appeared while ignoring the matrix itself, and to answer two forced-choice questions immediately after the second display in each trial disappeared. The two questions were identical to the two surprise questions in Experiment 1. The complete questions were presented and explained to the participants in the beginning of the experiment. Immediately following the second display in each trial, the two alternatives for the region question ("shape," "space between shapes") appeared on the screen. Following participants' response, the two alternatives for the change-detection question ("change," "no change") appeared on a new screen. Responses were made by pressing one of two keys.

Results and Discussion

Table 3 presents the mean percentage of correct choices for each question. Overall, participants correctly reported whether the region on which the matrix appeared in the preceding display was a "shape" (figure) or a "space between shapes" (ground) on 96% of all trials. An analysis of variance showed no difference in accuracy between trial types, $F < 1$. Thus, when attention was allocated to the backdrop region on which the matrix appeared, observers were aware of its figure-ground status, perceiving the

TABLE 3

Mean Percentage of Correct Region Detection and Change Detection for Each Trial Type in Experiment 2

	Backdrop organization			
Forced-choice question	FF	GG	FG	GF
Type of region	96	97	96	96
Change in region	94	94	95	88

Note. The abbreviations for the trial types indicate the types of backdrops (F = figure type; G = ground type) and their order across the two successive displays.

---

[2] Similar results were obtained for correct RT: The Backdrop Organization × Starting Organization interaction was significant, $F(1, 45) = 6.85$, $p < .02$, $\eta_p^2 = .13$, and analysis of simple effects showed a significant difference between FG (574 ms) and GF (594 ms) trials, $F(1, 45) = 8.49$, $p < .01$, $\eta_p^2 = .16$, and no difference between FF (592 ms) and GG (588 ms) trials, $F < 1$.

convex region as figure and the concave region as ground. These results are consistent with previous results, which demonstrated that when observers were presented with the same eight-region displays and asked to report the figural status of a probed region, they were likely to see the convex regions as figures (Kim & Peterson, 2002; Peterson & Salvagio, 2008).

Overall, participants were also accurate in reporting whether or not the figure-ground status of the region on which the matrix appeared had changed during the preceding trial, correctly responding on 93% of all trials. This result, indicating that participants were aware of figure-ground changes in the attended backdrop region, is compatible with previous findings demonstrating that when presented with foreground items lying on a background display, observers could detect background changes only when attention was allocated to the background (Mazza et al., 2005; Turatto, Angrilli, Mazza, Umiltà, & Driver, 2002).

Significant trial-dependent differences in change-detection accuracy were observed, however, $F(3, 66) = 4.52, p < .02$. Post hoc Tukey HSD (honestly significant difference) comparisons revealed that change detection was significantly lower on GF trials (88%) than on the other trials (94–95%). Thus, the appearance of a new figure in the attended location on GF trials interfered to some extent with change-detection accuracy. This result, obtained under conditions of attention, is consistent with the proposal that a change from figure to ground is not equivalent to a change from ground to figure, because only the latter involves processing of a new object, which can interfere with task performance even when the task is change detection (cf. Landman, Spekreijse, & Lamme, 2004). In Experiment 1, in which the backdrop was unattended, the direction of change in the backdrop region on which the target matrix appeared had a similar effect: Performance on the target-change task was significantly less efficient on GF than on FG trials.

Taken together, the results of the two experiments strongly suggest that figure-ground organization occurred in the scene backdrop when it was outside the focus of attention.

GENERAL DISCUSSION

Our results provide clear evidence that figure-ground segmentation can occur for unattended stimuli. When observers performed a demanding change-detection task on a small matrix presented on a task-irrelevant scene of alternating regions organized into figures and grounds by convexity, changes in the scene's figure-ground organization produced reliable congruency effects on performance. As noted earlier, these results cannot be due to implicit capturing of attention by convexity in the scene backdrop, nor to the backdrop's changes in convexity/concavity per se. These congruency effects arose despite inattentional blindness to the scene backdrop. When probed with surprise questions, participants could report neither the figure-ground status of the region on which the target appeared in the preceding display nor whether the figure-ground status of the region had changed during the preceding trial, but when participants attended to the scene backdrop, their answers were highly accurate.

The finding that figure-ground segmentation can occur without attention, together with previous findings indicating that some grouping (e.g., Kimchi & Razpurker-Apfeld, 2004; Russell & Driver, 2005) and surface completion (Moore et al., 2003) occur under inattention, supports the view that at least some perceptual organization processes are preattentive.

Note that our finding does not imply that figure-ground segmentation must always precede the deployment of focal attention, as many models of perception have assumed (e.g., Julesz, 1984; Treisman, 1986). The backdrop stimuli in our study contained eight alternating regions that were equated for all other stimulus factors (such as size, contrast, symmetry, familiarity, and orientation) but convexity. Convexity is a powerful cue for figural assignment in such displays (e.g., Hoffman & Singh, 1997; Kanizsa & Gerbino, 1976; Peterson & Salagio, 2008) and can, for example, override symmetry (Kanizsa & Gerbino, 1976). It is possible that when other, perhaps less potent, figural cues are involved, segmentation requires the scrutiny of focal attention.

Furthermore, in natural scenes, adjacent regions are likely to have multiple competing cues. Figure-ground assignment in this case requires the resolution of cross-edge competition (Peterson & Kim, 2001; Peterson & Skow, 2008), which may demand focal attention. Research on perceptual grouping suggests that attentional demands of organizational processes and the time course of these processes may be related: Grouping that took place without attention was achieved rapidly, whereas grouping that required attention was likely to consume time (Kimchi & Razpurker-Apfeld, 2004; Razpurker-Apfeld & Kimchi, 2007). For example, grouping into columns or rows by common color occurred rapidly and was accomplished without attention (see also Russell & Driver, 2005), but grouping into a shape by common color consumed time and did not occur under inattention. Peterson and Lampignano (2003; Peterson & Enns, 2005) have already demonstrated that when competing figural cues are present, figure-ground assignment is time-consuming. It may well be the case, then, that under such conditions it is also attention demanding (Kimchi & Razpurker-Apfeld, 2004; Trujillo, Allen, & Peterson, 2008).

Evidence that spatial attention can act as a cue for figure-ground assignment (Peterson & Gibson, 1994b; Vecera et al., 2004) also casts serious doubt on the assumption that figure-ground segmentation must necessarily be completed prior to the deployment of focal attention.

In sum, the present results provide strong evidence that some figure-ground segmentation can occur for stimuli that are unattended. Furthermore, the results suggest the exciting possibility that the relationship between attention and figure-ground perception is complex and multifaceted.

Acknowledgments—This research was supported partly by the Max Wertheimer Minerva Center for Cognitive Processes and Human Performance, University of Haifa (R.K.), and partly by National Science Foundation Grant BCS0425650 (M.A.P.). We thank Yafa Lev for programming assistance, and Dannah G. Raz and Allegra Dan for assistance in stimulus preparation and data collection. We are indebted to two anonymous reviewers for their helpful comments.

REFERENCES

Bahnsen, P. (1928). Eine untersuchung über symmetrie und asymmetrie bei visuellen wahrnehmungen. *Zeitschrift für Psychologie*, *108*, 355–361.

Barenholtz, E., & Feldman, J. (2006). Determination of visual figure and ground in dynamically deforming shapes. *Cognition*, *101*, 530–544.

Baylis, G.C., & Driver, J. (1995). One-sided edge assignment in vision: 1. Figure-ground segmentation and attention to objects. *Current Directions in Psychological Science*, *4*, 140–146.

Cohen, A., & Rafal, R.D. (1991). Attention and feature integration: Illusory conjunctions in a patient with a parietal lobe lesion. *Psychological Science*, *2*, 106–110.

Driver, J., & Baylis, G. (1996). Edge-assignment and figure-ground segmentation in short-term visual matching. *Cognitive Psychology*, *31*, 248–306.

Driver, J., Baylis, G.C., & Rafal, R.D. (1992). Preserved figure-ground segregation and symmetry perception in visual neglect. *Nature*, *360*, 73–75.

Hochberg, J. (1971). Perception I: Color and shape. In J.W. Kling & L.A. Riggs (Eds.), *Woodworth and Schlossberg's experimental psychology* (pp. 395–474). New York: Holt, Rinehart & Winston.

Hoffman, D.D., & Singh, M. (1997). Salience of visual parts. *Cognition*, *63*, 29–78.

Hulleman, J., & Humphreys, G.W. (2004). A new cue to figure-ground coding: Top-bottom polarity. *Vision Research*, *44*, 2779–2791.

Julesz, B. (1984). Textons, the elements of texture perception, and their interactions. *Nature*, *290*, 91–97.

Kanizsa, G., & Gerbino, W. (1976). Convexity and symmetry in figure-ground organization. In M. Henle (Ed.), *Vision and artifact* (pp. 25–32). New York: Springer.

Kim, J.H., & Peterson, M.A. (2002). Contextual modulation of the strength of Gestalt configural cues [Abstract]. *Journal of Vision*, *2*, 481.

Kimchi, R., & Razpurker-Apfeld, I. (2004). Perceptual grouping and attention: Not all groupings are equal. *Psychonomic Bulletin & Review*, *11*, 687–696.

Klymenko, V., & Weisstein, N. (1986). Spatial frequency differences can determine figure-ground organization. *Journal of Experimental Psychology: Human Perception and Performance*, *12*, 324–330.

Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt Brace Jovanovich.

Landman, R., Spekreijse, H., & Lamme, V.A. (2004). The role of figure-ground segregation in change blindness. *Psychonomic Bulletin & Review*, *11*, 254–261.

Mack, A., & Rock, I. (1998). *Inattentional blindness*. Cambridge, MA: MIT Press.

Marr, D. (1982). *Vision*. San Francisco: W.H. Freeman.

Mazza, V., Turatto, M., & Umiltà, C. (2005). Foreground-background segmentation and attention: A change blindness study. *Psychological Research*, *69*, 201–210.

Moore, C.M., Grosjean, M., & Lleras, A.E. (2003). Using inattentional blindness as an operational definition of unattended: The case of surface completion. *Visual Cognition*, *10*, 299–318.

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

Nelson, R.A., & Palmer, S.E. (2007). Familiar shapes attract attention in figure-ground displays. *Perception & Psychophysics*, *69*, 382–392.

Peterson, M.A., & Enns, J.T. (2005). The edge complex: Implicit perceptual memory for cross-edge competition leading to figure assignment. *Perception & Psychophysics*, *14*, 727–740.

Peterson, M.A., & Gibson, B.S. (1994a). Must figure–ground organization precede object recognition? An assumption in peril. *Psychological Science*, *5*, 253–259.

Peterson, M.A., & Gibson, B.S. (1994b). Object recognition contributions to figure-ground organization: Operations on outlines and subjective contours. *Perception & Psychophysics*, *56*, 551–564.

Peterson, M.A., & Kim, J.H. (2001). On what is bound in figures and grounds. *Visual Cognition*, *8*, 329–348.

Peterson, M.A., & Lampignano, D.L. (2003). Implicit memory for novel figure-ground displays includes a history of border competition. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 808–822.

Peterson, M.A., & Salvagio, E. (2008). *Is convexity a strong figural cue? It depends on the context*. Manuscript submitted for publication.

Peterson, M.A., & Skow, E. (2008). Inhibitory competition between shape properties in figure-ground perception. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 251–267.

Razpurker-Apfeld, I., & Kimchi, R. (2007). The time course of perceptual grouping: The role of segregation and shape formation. *Perception & Psychophysics*, *69*, 732–743.

Rubin, E. (1958). Figure and ground. In D.C. Beardslee & M. Wertheimer (Eds.), *Readings in perception* (pp. 194–203). Princeton, NJ: Van Nostrand. (Original work published 1915)

Russell, C., & Driver, J. (2005). New indirect measures of "inattentive" visual grouping in a change-detection task. *Perception & Psychophysics*, *67*, 606–623.

Sejnowski, T.J., & Hinton, G.E. (1987). Separating figure from ground with a Boltzmann machine. In M.A. Arbib & A.R. Hanson (Eds.), *Vision, brain, and cooperative computation* (pp. 703–724). Cambridge, MA: MIT Press.

Treisman, A. (1986). Features and objects in visual processing. *Scientific American*, *255*, 114–125.

Townsend, J.T., & Ashby, F.G. (1983). *Stochastic modelling of elementary psychological processes*. New York: Cambridge University Press.

Trujillo, L.T., Allen, J.J.B., & Peterson, M.A. (2008). *Electrophysiological evidence for inhibitory competition in figure-ground perception*. Manuscript submitted for publication.

Turatto, M., Angrilli, A., Mazza, V., Umiltà, C., & Driver, J. (2002). Looking without seeing the background change: Electrophysiological correlates of change detection versus change blindness. *Cognition*, *84*, B1–B10.

Vecera, S.P., Flevaris, A.V., & Filapek, J.C. (2004). Exogenous spatial attention influences figure-ground assignment. *Psychological Science*, *15*, 20–26.

Vecera, S.P., Vogel, E.K., & Woodman, G.F. (2002). Lower region: A new cue for figure-ground assignment. *Journal of Experimental Psychology: General*, *131*, 194–205.

Weisstein, N., & Wong, E. (1987). Figure-ground organization affects the early processing of information. In M.A. Arbib & A.R. Hanson (Eds.), *Vision, brain, and cooperative computation* (pp. 209–230). Cambridge, MA: MIT Press.

(RECEIVED 9/25/07; REVISION ACCEPTED 12/2/07)

Psychonomic Bulletin & Review
2007, 14 (1), 166-172

Automatic, stimulus-driven attentional capture by objecthood

Ruth Kimchi, Yaffa Yeshurun, and Aliza Cohen-Savransky
University of Haifa, Haifa, Israel

In three experiments, we investigated whether the mere organization of some elements in the visual field by Gestalt factors into a coherent unit (an object), with no abrupt onset or any other unique transient, could attract attention automatically. Participants viewed a display of nine red and green elements, one of which was the target, and had to identify the target's color. On some trials, a subset of the elements was grouped by Gestalt factors (collinearity, closure, and symmetry) into an object. The object was task-irrelevant and unpredictive of the target. Performance on trials with an object present in the display was faster than performance on trials with no object for targets in the object area (a benefit) but slower for targets in a nonobject area (a cost). These findings demonstrate that an object by itself can capture attention automatically in a stimulus-driven manner, much as exogenous cues can.

Perceptual organization—the visual processes structuring disparate visual information into the coherent units that we experience as environmental objects—and visual attention—the mechanisms by which some visual information in a scene is selected—have traditionally been separate fields of study. However, recent research (e.g., Driver, Davis, Russell, Turatto, & Freeman, 2001; Scholl, 2001; Vecera, 2000) points to important connections between the processes of organization and attention.

Many studies have demonstrated that perceptual organization constrains attentional selectivity. For example, responding to two features is easier when they belong to the same object than when they belong to two separate objects (e.g., Duncan, 1984), and interference from distractor stimuli in selective attention tasks is greater when the target and distractors are strongly grouped (e.g., Baylis & Driver, 1992; Kramer & Jacobson, 1991). Recent studies suggest that attention can also constrain perceptual organization. For example, attention can influence flanker–target integration (Freeman, Driver, Sagi, & Zhaoping, 2003) and figure–ground assignment (Vecera, Flevaris, & Filapek, 2004). Also, some forms of grouping can take place without attention, whereas others require controlled attentional processing (e.g., Kimchi & Razpurker-Apfeld, 2004). These findings suggest that perceptual organization and visual attention constrain each other.

The critical role of perceptual organization in designating potential objects raises an important issue concerning the relations between perceptual organization and attention: When some elements in the visual scene are organized by Gestalt factors into a coherent perceptual unit (an object),[1] is visual attention automatically deployed to that object?

Deployment of attention can be goal directed, based on the current behavioral goals of the observer (e.g., Desimone & Duncan, 1995; Posner, 1980). If we know, for example, where the most probable target location is, we can use this information to voluntarily (endogenously) direct our attention to this location. Deployment of attention can also be stimulus driven, in which case attention is captured involuntarily (exogenously) by certain stimulus events, such as an abrupt onset of a new perceptual object and some types of simple luminance and motion transients (e.g., Abrams & Christ, 2003; Jonides, 1981; Yantis & Hillstrom, 1994), or a salient singleton (e.g., Theeuwes, de Vries, & Godijn, 2003; but see Folk, Remington, & Johnston, 1992).

The purpose of this study is to examine whether the mere organization of some elements into an object, with no abrupt onset or any other unique transient, can capture attention automatically in a stimulus-driven manner, much as exogenous cues capture spatial attention automatically.

It is important to note that the question is not whether attention can be deployed to an object; it has already been documented that attentional selection can be object based. In the typical object-based studies, attention is directed to a part or an attribute of the object either exogenously (e.g., by briefly brightening the contour on one end of one of two rectangles; see Egly, Driver, & Rafal, 1994) or endogenously (e.g., by instructions to attend the size of a box and the side of its gap; see Duncan, 1984). These studies imply that the entire object is selected, but they do not show unequivocally that the object per se is the factor that attracts attention, because there are always other factors that do so. Consider, for example, Kramer and Jacobson's (1991) study. Observers responded to a centrally located target while ignoring adjacent distractors. Interference

R. Kimchi, rkimchi@research.haifa.ac.il

from response-incompatible distractors and facilitation from response-compatible distractors were greater when target and distractors were grouped into an object. Although a possible account of this finding is that the entire object captured attention automatically despite the instruction to focus exclusively on the target, an alternative interpretation is that attention was directed to the target and "spread" to the entire object. The latter is especially probable, given that the target location was fixed across the entire experiment.

To examine whether an object by itself captures attention, it is crucial that the object have no abrupt onset or any other unique transient, and that the object be irrelevant to the task at hand so that there will be no incentive for the observer to deliberately attend the object. To this end, we have modified a paradigm developed by Logan (1995). In this paradigm, observers view a display of nine red and green elements, one of which is the target, and are required to identify the color of the target. The elements form the vertices of four adjacent quadrants that make up a global "diamond." The target is indicated by an asterisk presented in the center of one of the quadrants and an instruction word—*above*, *below*, *right*, or *left*—that precedes the element display and specifies the position of the target relative to the asterisk. For example, if the word is *above*, observers have to identify the color of the element above the asterisk. Thus, performing the task requires locating the asterisk, locating the target relative to the asterisk, and analyzing the target's color. In order to allow manipulation of organization in the display, we have substituted the O elements of Logan's original display with L elements in various orientations (Figure 1). On half the trials, the four Ls of one of the quadrants are rotated, thereby conforming to the Gestalt factors of collinearity, closure, and symmetry and forming a diamond-like object. The asterisk appears in the object quadrant (*inside-object* condition, Figure 1A) on 12.5% of all trials, and in a nonobject quadrant (*outside-object* condition, Figure 1B) on 37.5% of all trials. On 50% of all trials, no object is present in the display (*no-object* condition, Figure 1C). Note that the object is task irrelevant (because the task-relevant feature is the color of a single element) and is unpredictive of the relevant quadrant or the target. Furthermore, no unique onset is associated with the object, because it appears simultaneously with the onset of the entire element display. This is a critical difference from previous research, in which attention was captured by the unique appearance of an object defined by discontinuities in luminance, motion, texture, or depth (e.g., Franconeri, Hollingworth, & Simons, 2005; Yantis & Hillstrom, 1994).

We hypothesize that if attention is automatically drawn to the object, performance will be faster and/or more accurate in the inside-object condition than in the no-object condition (a benefit), because attention is allocated in advance to the object quadrant, and slower and/or less accurate in the outside-object condition than in the no-object condition (a cost), because attention has to be redirected from the object quadrant to the actual relevant quadrant.

We tested this hypothesis in Experiment 1 with 150 msec between the onset of the element display and that of the asterisk. The results showed the expected cost and benefit, demonstrating capture of attention by the object. Experiments 2 and 3 examined the time course of this attentional capture by manipulating the stimulus onset asynchrony (SOA) between the element display and the asterisk.

EXPERIMENT 1

Method

Participants. Fourteen individuals with normal or corrected-to-normal vision participated in this experiment.

Stimuli. The displays consisted of nine red and green Ls in different orientations, one of which was the target. The Ls formed the vertices of four adjacent quadrants that made up a global "diamond" (Figure 1). At a viewing distance of 60 cm, the global display subtended 12.73º × 12.73º, and each L subtended 1.5º × 1.5º. A black asterisk, subtending 0.8º, was presented in the center of one of the quadrants, indicating the relevant quadrant. The asterisk appeared equally often in each of the quadrants. An instruction word (*above*, *below*, *right*, or *left*) preceded the element display and appeared at the center of the screen, subtending about 4.0º. The word specified the position of the target relative to the asterisk. Targets were assigned to positions randomly, with the constraint that each color appeared equally often in each target position (above, below, right, and left of the asterisk) in each quadrant (in the top, bottom, right, and left quadrants). The colors of the nontarget elements were assigned equally often to one of the following color distributions: four green–four red, three green–five red, or five red–three green.

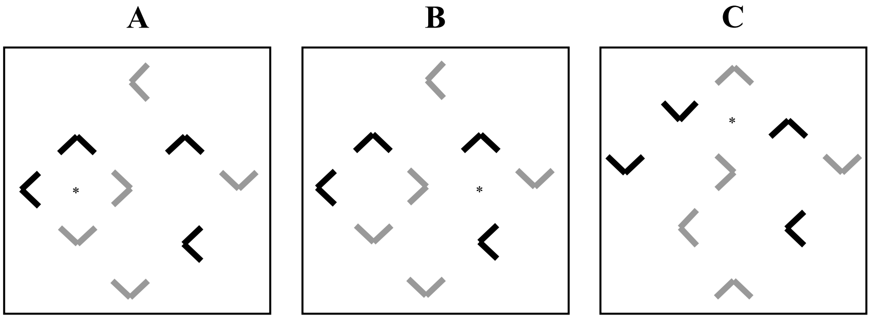

Figure 1. Examples of the displays in the three conditions in all experiments. (A) Inside-object condition: object present in display and asterisk appearing in center of object quadrant; (B) Outside-object condition: object present in display and asterisk appearing in center of nonobject quadrant; and (C) No-object condition: no object present in display. In the experiments, the colors of the elements were red and green. Fifty percent of the trials were no-object trials, 12.5% were inside-object trials, and 37.5% were outside-object trials.

Rotating the four Ls of one of the quadrants to conform to the Gestalt factors of collinearity, symmetry, and closure yielded a diamond-like object (Figures 1A and 1B). The object appeared equally often in each of the four quadrants.

Design and Procedure. The combination of object (present or absent), object position (top, left, bottom, or right quadrant), asterisk position (top, left, bottom, or right quadrant), target position (above, below, left, or right of the asterisk), and target color (red or green) produced 256 different trials that were randomized within a block. There were 4 blocks of experimental trials, for a total of 1,024 trials, preceded by 32 practice trials. The combination of object, object position, and asterisk position produced three critical conditions: *inside-object* (Figure 1A)—an object present in the display and an asterisk appearing in the object quadrant; *outside-object* (Figure 1B)—an object present in the display and an asterisk appearing in a nonobject quadrant; and *no-object* (Figure 1C)—no object present in the display. Fifty percent of the trials within a block were no-object trials, 12.5% were inside-object trials, and 37.5% were outside-object trials.

Each trial (Figure 2) started with a fixation mark that appeared for 500 msec, followed by an instruction word presented for 1,000 msec. Then the element display appeared, and 150 msec later the asterisk was added to the display. The element display and the asterisk stayed on until a response was made. By pressing one of two response keys, participants indicated as rapidly and accurately as possible the color of the target. The intertrial interval was 1,500 msec.

Results and Discussion

Participants were accurate in identifying the color of the target (mean error rate = 5.03%), and there was no indication of speed–accuracy trade-off. All response time (RT) summaries and analyses are based on participants' median RTs for correct responses. Figure 3 shows the mean RTs, collapsed across target color and target position, as a function of condition (inside-object, outside-object, and no-object). Error rate data are presented in Table 1 (Experiment 1). Preliminary analyses indicated that observers were faster and more accurate in responding to targets above or below than right or left of the asterisk, replicating previous findings demonstrating the advantage of the vertical axis over the horizontal axis for deictic relations (e.g., Logan, 1995). Target position, however, did not interact significantly with either condition or target color. The analyses further confirmed that there was no main effect of target color, and it did not interact with condition. This pattern of results concerning target color and target position was true of the next two experiments, as well.

The collapsed RT data were submitted to a one-way repeated measures ANOVA that showed a significant effect

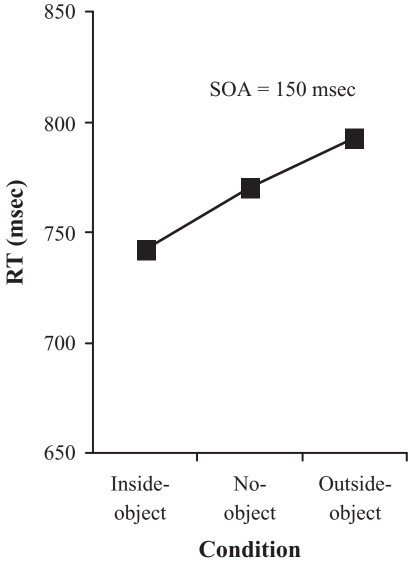

Figure 3. Mean correct response times (RTs) as a function of condition in Experiment 1.

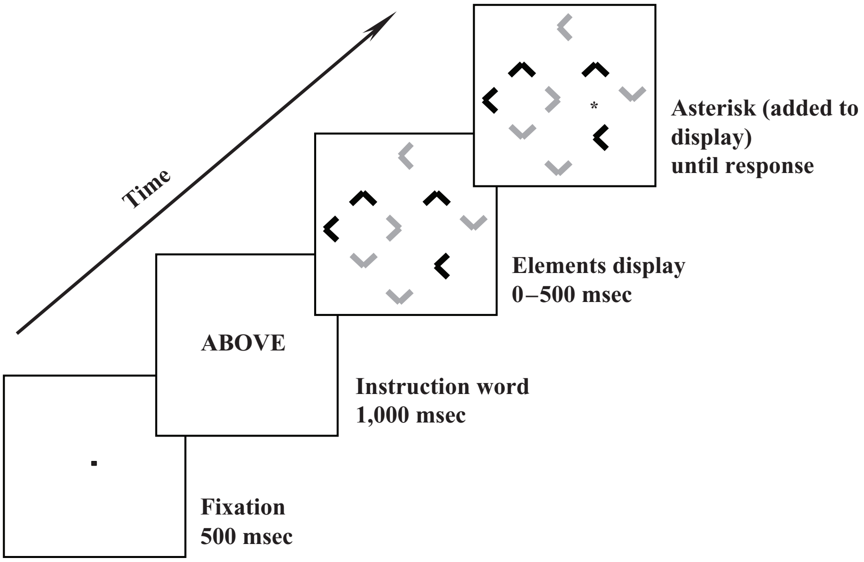

Figure 2. Sequence of events on each trial in all experiments. The stimulus onset asynchrony between the elements display and the asterisk varied. The illustration depicts an outside-object trial with the instruction word *above*. In this trial, the observer had to identify the color of the element above the asterisk.

Table 1
Error Rates (Percentages) for Each Experiment As a Function of Stimulus Onset Asynchrony (SOA, in Milliseconds)

SOA	Inside-Object	No-Object	Outside-Object
		Experiment 1	
150	4.9	4.8	5.5
		Experiment 2	
0	3.3	3.4	3.6
75	2.7	3.3	3.7
500	2.8	3.0	3.2
		Experiment 3	
0	5.2	5.8	5.8
50	4.8	5.2	5.2
100	4.7	5.3	5.7

of condition [$F(2,26) = 14.13$, $MS_e = 728$, $p < .0001$]. Planned comparisons revealed that responses in the inside-object condition were 31 msec faster than responses in the no-object condition [$F(1,13) = 6.06$, $MS_e = 1{,}067$, $p < .03$], and responses in the outside-object condition were 24 msec slower than responses in the no-object condition [$F(1,13) = 27.02$, $MS_e = 146$, $p < .0002$]. All of the same analyses were conducted on the error rate data. Although error rates showed trends similar to those of the RTs, none of the effects reached statistical significance.

These results show the expected cost and benefit: When an object was present in the display and the asterisk appeared in the object quadrant, performance was facilitated relative to the no-object condition, and when the asterisk appeared in a nonobject quadrant, performance was hindered relative to the no-object condition. These results suggest that although the object was irrelevant to the task and unpredictive of the relevant quadrant or the target, it captured attention.

EXPERIMENTS 2 AND 3

In these experiments, we examined how early the cost and benefit manifest themselves, and how sustained they are, by varying the SOA between the element display and the asterisk (0–500 msec). If the capture of attention by the "object" is similar to that produced by exogenous cues, a benefit should be evident at 100-msec SOA (because benefit reflects the advanced allocation of attention to the relevant quadrant, which takes about 100 msec; e.g., Posner, 1980). A cost may be evident at shorter SOAs (because even with a 0-msec SOA, attention may be captured by the task-irrelevant object rather than by the task-relevant asterisk, resulting in performance cost). These effects may persist at the 500-msec SOA, because of the temporal overlap between the object and the target (e.g., Collie, Maruff, Yucel, Danckert, & Currie, 2000).

Method

Participants. Twenty-eight individuals with normal or corrected-to-normal vision participated in these experiments (14 in each experiment).

Design and Procedure. The SOA between the element display and the asterisk was 0, 75, or 500 msec in Experiment 2 and 0, 50, or 100 msec in Experiment 3. In each experiment, SOA was combined orthogonally with all other factors (see Experiment 1), producing 768 different trials that were randomized within a block. There were 4 blocks of experimental trials, for a total of 3,072 trials, administered in 2 sessions.

All other aspects of the method were similar to those of Experiment 1.

Results and Discussion

Participants in both experiments were accurate in identifying the target color (mean error rate was 3.1% in Experiment 2 and 5.2% in Experiment 3), and there was no indication of speed–accuracy trade-off (error rate data are presented in Table 1, Experiments 2 and 3). The error rate data showed effects similar to those of the RT data, but none of them reached statistical significance. Therefore, error rates are not discussed further. Mean correct RTs as a function of condition for each SOA are plotted in Figure 4.

In Experiment 2, a 3 (condition) × 3 (SOA) repeated measures ANOVA showed that overall RT decreased when SOA increased [$F(2,26) = 64.68$, $MS_e = 1{,}759$, $p < .0001$]. The effect of condition was significant [$F(2,26) = 27.73$, $MS_e = 289$, $p < .0001$], as was the interaction between condition and SOA [$F(4,52) = 2.95$, $MS_e = 403$, $p < .03$]. Planned comparisons were conducted to assess the benefit and cost at each SOA. No significant benefit was observed at 0- and 75-msec SOA [$F < 1$; $F(1,13) = 1.50$, $MS_e = 487$, $p > .24$, respectively]. Only at 500-msec SOA were responses in the inside-object condition significantly faster (by 29 msec) than in the no-object condition [$F(1,13) = 12.62$, $MS_e = 468$, $p < .004$]. On the other hand, responses in the outside-object condition were significantly slower than responses in the no-object condition at all SOAs. There was a cost of 26 msec at 0-msec SOA [$F(1,13) = 35.07$, $MS_e = 129$, $p < .0001$], a cost of 14 msec at 75-msec SOA [$F(1,13) = 6.03$, $MS_e = 218$, $p < .03$], and a cost of 11 msec at 500-msec SOA [$F(1,13) = 6.55$, $MS_e = 127$, $p < .025$].

In Experiment 3, the ANOVA showed significant effects of SOA [$F(2,26) = 113.05$, $MS_e = 536$, $p < .0001$] and condition [$F(2,26) = 25.18$, $MS_e = 466$, $p < .0001$]. Planned comparisons showed a significant benefit of 21 msec at 100-msec SOA [$F(1,13) = 6.83$, $MS_e = 449$, $p < .02$]. No significant benefit was observed at 0- and 50-msec SOA [$F < 1$; $F(1,13) = 2.39$, $MS_e = 714$, $p > .15$, respectively]. On the other hand, there was a cost of 28 msec at 0-msec SOA [$F(1,13) = 23.66$, $MS_e = 233$, $p < .0003$], a cost of 12 msec at 50-msec SOA [$F(1,13) = 6.64$, $MS_e = 156$, $p < .02$], and a cost of 21 msec at 100-msec SOA [$F(1,13) = 6.97$, $MS_e = 434$, $p < .02$].

These results show that, as with typical exogenous cues, a significant benefit emerged at 100-msec SOA and was still observed at 500-msec SOA, presumably because the object was present until target offset (e.g., Collie et al., 2000; Wascher & Tipper, 2004).[2]

Cost for performance in the outside-object condition was observed at all SOAs, including at 0-msec SOA, suggesting that the task-irrelevant object captured attention, even in a tight competition with the task-relevant aster-

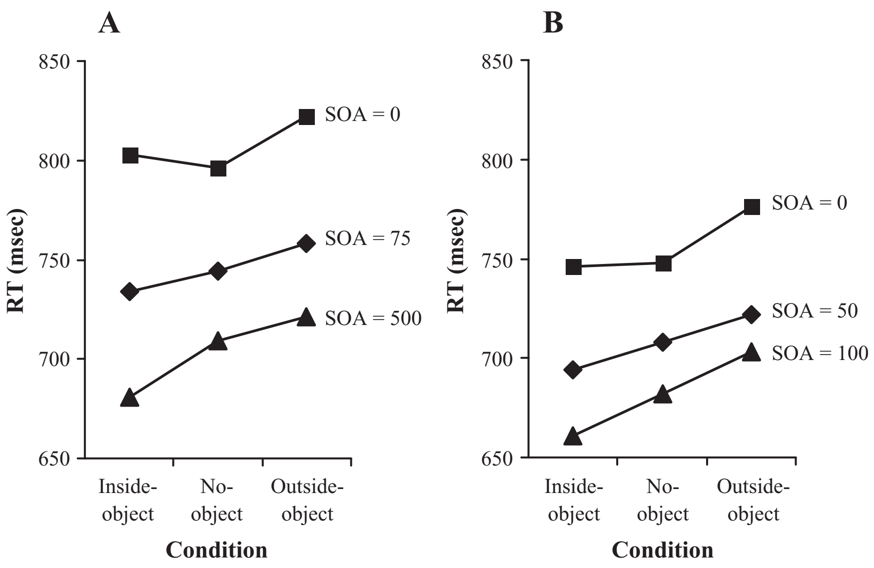

Figure 4. Mean correct response times (RTs) as a function of condition and stimulus onset asynchrony (SOA, in milliseconds) in (A) Experiment 2 and (B) Experiment 3.

isk. It should be also noted that in the present displays some of the targets in the outside-object condition actually "belonged" to the object, whereas others did not. For example, if the object quadrant was the left-hand one and the asterisk appeared in the right-hand quadrant (Figure 1B), the element to the left of the asterisk "belonged" to the object, whereas the other three elements in this quadrant (i.e., the elements above, below, and right of the asterisk) did not. Planned comparisons conducted to assess the cost for each of these target types (Table 2) showed costs for both target types at all SOAs except at 500-msec SOA, at which no significant cost was found for targets that did not "belong" to the object. The cost for targets that "belonged" to the object was somewhat higher than the cost for targets that did not "belong" to the object, suggesting that some of the observed cost may be attributed to difficulty in "extracting" an already-grouped element.

Table 2
Cost (Outside-Object RT Minus No-Object RT, in Milliseconds) for Targets That "Belonged" to the Object and Targets That Did Not "Belong" to the Object As a Function of Stimulus Onset Asynchrony (SOA) in Each Experiment

SOA	Target "Belong" to Object	 Not "Belong" to Object
	Experiment 1	
150	38	13
	Experiment 2	
0	33	19
75	15	10
500	21	3
	Experiment 3	
0	45	16
50	19	6
100	22	18

GENERAL DISCUSSION

The present results show that when some elements in the display were organized by Gestalt factors into an object, with no abrupt onset or any other unique transient, performance for targets in the object area was facilitated relative to performance for targets in the no-object condition (a benefit), and performance for targets in a nonobject area was impeded relative to the no-object condition (a cost), even though the object was irrelevant to the task and unpredictive of the target. A significant benefit emerged when the target's indicator, the asterisk, appeared at least 100 msec after the display's onset, and a significant cost was observed at all SOAs examined, including when the display and the asterisk appeared simultaneously. The time course of these effects is similar to the one obtained with exogenous cues under comparable cue–target temporal relations.

These benefit and cost effects suggest that the object captured attention automatically, in a stimulus-driven manner, much as exogenous cues do. Although it is well documented by now that objects play an important role in visual attention (e.g., Scholl, 2001), our results are the first to demonstrate that an object per se can attract attention automatically.

One may argue that a subgroup of organized elements in a field of unorganized elements functions as a singleton. However, *singleton* usually refers to a single element or item that differs in some attribute from the surrounding elements, which are relatively homogenous in that attribute (e.g., Duncan & Humphreys, 1989). Here, the elements always vary in color and orientation, and the only difference between the object-elements and the other elements resides in the spatial relations among the former. Furthermore, the ability of a singleton to produce stimulus-driven attentional capture depends on current attentional setting (e.g., Folk et al., 1992). In the present study, however, the object is irrelevant, because the task-relevant feature is the target's color. Moreover, grouping some of the elements in the display actually conflicts with the task demand to identify the color of a single element. Thus, the object in our displays captures attention, even though it is incongruent with the current attentional setting. The finding of a significant cost, even when the task-relevant asterisk and the task-irrelevant object appear simultaneously, further indicates that the object is not overridden by a deliberate, goal-directed strategy.

What mechanisms underlie these object effects? We can only speculate. The benefit observed in this study can be accounted for by facilitatory processes at the attended area, such as a mechanism that accelerates the rate at which attended information is processed (e.g., Carrasco & McElree, 2001). The cost may be due to disengagement from the attended object and movement to the task-relevant asterisk (e.g., Posner, Walker, Friedrich, & Rafal, 1984). Alternatively, different mechanisms may account for the observed cost and benefit. In addition to facilitatory processes at the attended area, attention can also operate by inhibiting the information in nonattended areas (e.g., a mechanism of noise reduction; Shiu & Pashler, 1995). The finding that a significant benefit was observed at 500-msec SOA, but that there was no cost for targets that did not "belong" to the object, suggests that the object may have exerted influence by means of different mechanisms. On this account, the benefit can be attributed to facilitatory processes, the cost for targets that did not "belong" to the object can be attributed to inhibitory processes that may decay with time, and the cost for targets that "belonged" to the object may be attributed, at least partly, to the need to "extract" the target from the object.

In sum, our results show that the mere organization of some elements by Gestalt factors into a coherent perceptual unit (an object), with no abrupt onset or any other unique transient, can produce automatic, stimulus-driven attentional capture. This finding suggests that the visual system has a bias for such units. Favoring a perceptual unit that conforms to Gestalt factors is a desirable characteristic for a system whose goal is object identification and recognition, because these factors are likely to imply objects in the environment.

An automatic, stimulus-driven capture of attention by an object may provide a single account for a variety of "object advantage" effects reported in the literature. These include easier detection of four target lines embedded in distractors when the lines are organized into a face-like pattern than when the lines are a meaningless cluster (Gorea & Julesz, 1990); higher sensitivity for a target probe positioned inside rather than outside a circular contour embedded in a random background (Kovács & Julesz, 1993); more accurate discrimination of line segments when flashed on the figure than on the ground (Wong & Weisstein, 1982); better memory for a figure's contour than for the ground's contour (Driver & Baylis, 1996); and greater brain activation when the target appears in a region bounded by an object than in an unbounded region (Arrington, Carr, Mayer, & Rao, 2000).

Finally, once we have demonstrated that an object can capture attention automatically, we can explore which organization factors (e.g., collinearity, closure, or symmetry) are necessary for an object to capture attention. This effort may provide insights into the nature of objecthood.

AUTHOR NOTE

Part of this research formed A.C.-S.'s M.A. thesis, under R.K.'s supervision. This research was supported in part by Israel Science Foundation Grant 94/06. We thank Tomer Carmel for help in data collection, and David Navon, Catherine Arrington, Jim Pomerantz, and an anonymous reviewer for helpful comments. Correspondence concerning this article should be addressed to R. Kimchi, Department of Psychology, University of Haifa, Haifa 31905, Israel (e-mail: rkimchi@research.haifa.ac.il).

REFERENCES

Abrams, R. A., & Christ, S. E. (2003). Motion onset captures attention. *Psychological Science*, **14**, 427-432.

Arrington, C. M., Carr, T. H., Mayer, A. R., & Rao, S. M. (2000). Neural mechanisms of visual attention: Object-based selection of a region in space. *Journal of Cognitive Neuroscience*, **12**, 106-117.

Baylis, G. C., & Driver, J. (1992). Visual parsing and response competition: The effect of grouping factors. *Perception & Psychophysics*, **51**, 145-162.

Carrasco, M., & McElree, B. (2001). Covert attention accelerates the rate of visual information processing. *Proceedings of the National Academy of Sciences*, **98**, 5363-5367.

Collie, A., Maruff, P., Yucel, M., Danckert, J., & Currie, J. (2000). Spatiotemporal distribution of facilitation and inhibition of return arising from the reflexive orienting of covert attention. *Journal of Experimental Psychology: Human Perception & Performance*, **26**, 1733-1745.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, **18**, 193-222.

Driver, J., & Baylis, G. (1996). Edge-assignment and figure-ground segmentation in short-term visual matching. *Cognitive Psychology*, **31**, 248-306.

Driver, J., Davis, G., Russell, C., Turatto, M., & Freeman, E. (2001). Segmentation, attention and phenomenal visual objects. *Cognition*, **80**, 61-95.

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

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

Egly, R., Driver, J., & Rafal, R. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, **123**, 161-177.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 1030-1044.

Franconeri, S. L., Hollingworth, A., & Simons, D. J. (2005). Do new objects capture attention? *Psychological Science*, **16**, 275-281.

Freeman, E., Driver, J., Sagi, D., & Zhaoping, L. (2003). Top-down modulation of lateral interactions in early vision: Does attention affect integration of the whole or just perception of the parts? *Current Biology*, **13**, 985-989.

Gorea, A., & Julesz, B. (1990). Context superiority in a detection task with line-element stimuli: A low-level effect. *Perception*, **19**, 5-16.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye movement. In J. [B.] Long & A. [D.] Baddeley (Eds.), *Attention and performance IX* (pp. 187-203). Hillsdale, NJ: Erlbaum.

Kimchi, R., & Razpurker-Apfeld, I. (2004). Perceptual grouping and attention: Not all groupings are equal. *Psychonomic Bulletin & Review*, **11**, 687-696.

Kovács, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one: Effect of closure in figure–ground segmentation. *Proceedings of the National Academy of Sciences*, **90**, 7495-7497.

Kramer, A. F., & Jacobson, A. (1991). Perceptual organization and focused attention: The role of objects and proximity in visual processing. *Perception & Psychophysics*, **50**, 267-284.

Logan, G. D. (1995). Linguistic and conceptual control of visual spatial attention. *Cognitive Psychology*, **28**, 103-174.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, **32**, 3-25.

Posner, M. I., Walker, J. A., Friedrich, F. J., & Rafal, R. D. (1984). Effects of parietal injury on covert orienting of attention. *Journal of Neuroscience*, **4**, 1863-1874.

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, **80**, 1-46.

Shiu, L.-P., & Pashler, H. (1995). Spatial attention and vernier acuity. *Vision Research*, **35**, 337-343.

Theeuwes, J., de Vries, G.-J., & Godijn, R. (2003). Attentional and oculomotor capture with static singletons. *Perception & Psychophysics*, **65**, 735-746.

Vecera, S. P. (2000). Toward a biased competition account of object-based segregation and attention. *Brain & Mind*, **1**, 353-384.

Vecera, S. P., Flevaris, A. V., & Filapek, J. C. (2004). Exogenous spatial attention influences figure-ground assignment. *Psychological Science*, **15**, 20-26.

Wascher, E., & Tipper, S. P. (2004). Revealing effects of noninformative spatial cues: An EEG study of inhibition of return. *Psychophysiology*, **41**, 716-728.

Wong, E., & Weisstein, N. (1982). A new perceptual context-superiority effect: Line segments are more visible against a figure than against a ground. *Science*, **218**, 587-589.

Yantis, S., & Hillstrom, A. P. (1994). Stimulus-driven attentional capture: Evidence from equiluminant visual objects. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 95-107.

NOTES

1. The difficult question of what constitutes a perceptual object has not yet been answered (see, e.g., Scholl, 2001). Although there may be some debate as to whether an organized group of elements—or rather, any individual element in the visual display—should be considered an object, we use the term *object* to refer to "elements in the visual scene organized by Gestalt factors into a coherent unit."

2. An additional experiment indicated that the object effects persisted also at a 900-msec SOA.

(Manuscript received November 22, 2005;
revision accepted for publication April 28, 2006.)