# Facilitation by exogenous attention for static and dynamic gestalt groups

**Fahrettin F. Gonen · Hamza Hallal · Haluk Ogmen**

**Abstract** Attentional mechanisms allow the brain to selectively allocate its resources to stimuli of interest within the huge amount of information reaching its sensory systems. The voluntary component of attention, endogenous attention, can be allocated in a flexible manner depending on the goals and strategies of the observer. On the other hand, the reflexive component, exogenous attention, is driven by the stimulus. Here, we investigated how exogenous attention is deployed to moving stimuli that form distinct perceptual groups. We showed that exogenous attention is deployed according to a reference frame that moves along with the stimulus. Moreover, in addition to the cued stimulus, exogenous attention is deployed to all elements forming a perceptual group. These properties provide a basis for the efficient deployment of exogenous attention under ecological viewing conditions.

**Keywords** Object-based attention · Space-based attention · Attentional capture

The visual system receives a staggering amount of information. Given its limited resources, it needs to filter out irrelevant information while selecting information of interest for further processing. Attention plays a crucial role in the filtering and selection of information. Attentional processes can be classified into two broad types (e.g., Cheal & Lyon, 1991; Egeth & Yantis, 1997; Jonides, 1981; Müller & Rabbitt, 1989; Nakayama & Mackeben, 1989; Posner, 1980; Weichselgartner & Sperling, 1987): *Endogenous attention* refers to a relatively slow component, whereby observers direct their attention by voluntary control to a stimulus or

location of choice. *Exogenous attention* is a relatively fast component through which attention is directed as a reflexive response to the stimulus. For example, the onset of a flash at a peripheral location reflexively summons attention to this location. Since endogenous attention is under voluntary control, it can be flexibly directed, depending of how targets of interest are defined—for example, by color, location, shape, semantics, and so forth (Barrett, Bradshaw, & Rose, 2003; Müller & Rabbitt, 1989). On the other hand, because exogenous attention is driven by the stimulus, rather than by voluntary control, where it is directed to may not be as flexible as with endogenous attention.

Various lines of evidence indicate that exogenous attention can be directed to retinotopic and spatiotopic locations as well as to "objects" (e.g., Boi, Vergeer, Ogmen, & Herzog, 2011; Brown, Breitmeyer, Leighty, & Denney, 2006; Egly, Driver, & Rafal, 1994; Egly, Rafal, Driver, & Starrveveld, 1994; Iani, Nicoletti, Rubichi, & Umiltà, 2001; Lamy & Egeth, 2002; Lamy & Tsal, 2000; Moore, Yantis, & Vaughan, 1998; Reppa, Schmidt, & Leek, 2012; Theeuwes, Mathôt, & Grainger, 2013; Vecera, 1994). To demonstrate the latter effect, Egly, Driver, and Rafal (1994) presented two adjacent rectangles and cued one edge of one of the rectangles. When Reaction Times (RTs) were compared between two equidistant targets, one within the cued rectangle and the other within the uncued rectangle, it was found that they were significantly shorter to the target within the cued rectangle.[1] This was interpreted as exogenous attention being summoned to the whole cued object. According to an alternative explanation, once exogenous attention is attracted to a location within an object, it automatically spreads to the entire object.

F. F. Gonen · H. Hallal · H. Ogmen (✉)
Department of Electrical and Computer Engineering, Center for Neuro-Engineering and Cognitive Science, University of Houston, Houston, TX 77204-4005, USA
e-mail: ogmen@uh.edu

---

[1] Note that cueing can also generate inhibitory effects, known as inhibition of return (for a review, see Reppa et al., 2012). Our study is focused on the facilitatory effects of the cue.

Both of these interpretations are based on the concept of an "object"; however, the exact definition of the "object" concept remains difficult (e.g., Humphreys & Riddoch, 2007; Kasai, Moriya, & Hirano, 2011; Marr, 1982; Pinna, 2014; Scholl, 2001). For example, one may define an object by contour closure. According to this definition, the allocation or spreading of attention would be limited by the contours of the stimulus. However, it has been shown that "object"-based attentional benefits can be also observed for stimuli with open contours (e.g., Avrahami, 1999; Marino & Scholl, 2005; Marrara & Moore, 2003), as well as for gestalt groups (Wagemans et al., 2012; Wertheimer, 1923) without contours (Marrara & Moore, 2003). Boi et al. (2011) showed that cueing effects follow spatio-temporal grouping relations. These findings suggest an important role for perceptual grouping in the allocation of attention. In fact, one way to operationalize the concept of "object" is to use gestalt principles such as figure–ground segregation and perceptual grouping. One goal of our study was to examine the relationship between the allocation of exogenous attention and perceptual grouping.

Most studies of attention have focused on static stimuli. However, our natural environment is dynamic, and many objects of interest are in motion. Using moving stimuli, Lamy and Tsal (2000) and Soto and Blanco (2004) showed that the facilitatory effect of attention can be found, not just on the basis of the spatial location of the cue, but also on the basis of the moving position of the cued object. These studies used predictive cues, and thereby engaged both exogenous and endogenous components of attention. Ro and Rafal (1999) used two moving squares, one of which was cued. After the end of motion, a target appeared with equal probabilities either in the cued or the uncued square. With this uninformative cue, they showed that one can observe either facilitatory or inhibitory effects of the cue, depending on the stimulus parameters. Their results suggest that attentional effects can be observed for moving objects; however, they did not compare these effects to spatial effects of cueing. Furthermore, the spatial distance between the cue and the final position of the cued object was shorter than the spatial distance between the cue and the final position of the uncued object, raising the possibility of a distance-based spatial effect of the cue. A second goal of our study was thus to extend Ro and Rafal's finding for the facilitatory effect of exogenous cues by comparing it to purely spatial (retinotopic[2]) as well as equidistant within- and between-object effects. Moreover, we examined the relationship between the allocation of exogenous attention and perceptual grouping for both static and dynamic stimuli using similar configurations.

In the first experiment, we modified Egly, Driver, and Rafal's (1994) paradigm to introduce motion and to keep the

motion at a fixed eccentricity. In the second experiment, we introduced color grouping to this design, to investigate whether attention is allocated exclusively to the cued item or to all items that are perceptually grouped with the cued item. Finally, in the third experiment we used another grouping principle, grouping by common fate, to investigate whether the findings from color grouping would extend to other grouping types.

## Experiment 1: allocation of exogenous attention to dynamic stimuli

### Method

All experiments reported in this article were conducted according to a protocol approved by the University of Houston Committee for the Protection of Human Subjects, in accordance with the federal regulations (45 CFR 46), the ethical principles established by the Belmont Report, and the principles expressed in the Declaration of Helsinki. Fifteen students from the University of Houston participated in Experiment 1. All had normal or corrected-to-normal vision and were naïve to the purpose of the experiment. Subjects provided written informed voluntary consent approved by the University of Houston Committee for the Protection of Human Subjects.

*Apparatus* All stimuli were presented on a 20-in. NANAO FlexScan color monitor. The resolution of the display was set to 656 × 492 pixels with a 100-Hz frame rate. Generation of the stimuli was made possible by a video card (Visual Stimulus Generator; VSG 2/3) manufactured by Cambridge Research Systems. A fixed head- and chinrest was set to a distance of 1 m away from the display monitor. The screen size was approximately 23° × 17.5° and each pixel corresponded to 1.7 arcmin. Reaction times (RTs) were measured by a joystick device interfaced to the VSG board.

*Stimulus and procedure* The stimuli used in Egly, Driver, and Rafal (1994) were modified in order to introduce motion. The stimulus (Fig. 1) consisted of four arcs, either stationary or rotating around a fixation point, which was a white plus sign, placed in the center of the screen. To keep eccentricity constant, the stimuli rotated along a virtual circle of fixed radius. The dimensions of the stimulus are shown in the upper panel of Fig. 1.

The angular extent of each arc in polar coordinates was set to 52.5°, and the spacing between them was 37.5°. The height of each arc from the fixation point to their edge was 7° of visual angle. The angular extent of each arc and the spacing between the arcs added up to 90°, so that the four arcs divided the circle equally. The sizes of the cue and the target in polar coordinates were 7.5° (i.e., 1° of visual angle).

---

[2] In our study, eyes were kept steady, and thus spatiotopic and retinotopic references to the locations of the cue and the target were equivalent.
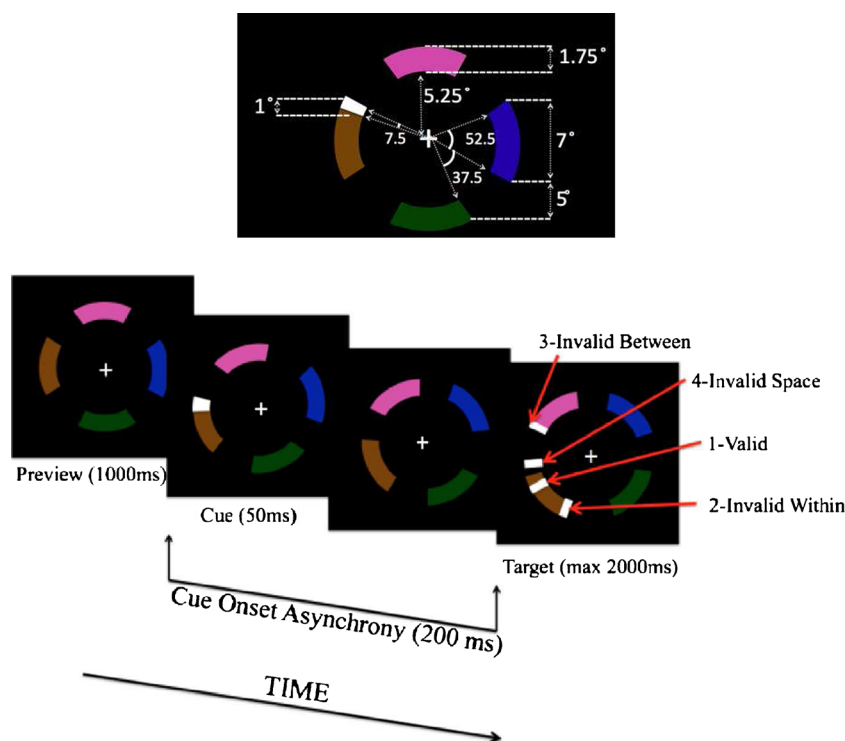
**Fig. 1** Stimulus configuration for Experiment 1. Upper panel: Spatial dimensions. Lower panel: Temporal sequence

The coordinates of the stimuli in CIE color space were as follows: The background was dark; the target and the cue were white (0.2044, 0.48085, 20). The four arcs were blue (0.3044, 0.6541, 2), green (0.0312, 0.5808, 2), brown (0.1044, 0.3076, 2), and dark pink (0.3776, 0.3808, 2). The arcs were displayed with different colors to make each arc visually distinct from the others. This minimized the possibility of confusing the arcs with each other when they were rotating.

Each trial started with the preview of four arcs, which lasted 1,000 ms. In the static condition, the arcs remained stationary during the preview period. In the dynamic condition, the arcs rotated either clockwise or counterclockwise, selected randomly on each trial. Half of the dynamic trials were clockwise, and the other half had counterclockwise rotation. The rotation speed was 40°/s in polar coordinates. In other words, one complete turn of the arcs required 9 s. The starting positions of the arcs were such that the midpoints of the arcs were on the cardinal axes, as is depicted in Fig. 1, upper panel. In each trial, the assignments of the four colors to the arcs were random. After the preview, the cue was presented for 50 ms in one of the two edges of a randomly selected arc. Thus, the cue could appear in eight possible locations. Following the cue onset asynchrony (COA) period of 200 ms, the target was presented. In the static case, three conditions were possible for the target location: In the *valid* condition, the target was at the same end of the same arc as the cue, but slightly shifted (Fig. 1, lower panel), to make the valid target and the invalid-space targets equidistant from the location of the corner of the arc in which the cue appeared. In the *invalid-within* condition, the target and the cue were presented inside the same arc, but at different ends. In the *invalid-between* condition, the target was displayed in an arc adjacent to the one in which the cue was presented (see Fig. 1). The center-to-center angular distances from the cue to the invalid-within target and from the cue to the invalid-between target were equal and set to 45°. The dynamic case added one more target presentation condition—namely, *invalid-space*—in which the target was presented at the location of the cue, but after the cued arc had moved from that position (Fig. 1).

Observers were asked to keep their gaze steady on the fixation stimulus and to press a joystick button as soon as the target appeared. RTs were the dependent variable. Catch trials without target presentation were also included. In the absence of target presentation, observers were asked not to press the joystick button. Observers were asked to respond as quickly as possible while keeping the highest possible level of accuracy in catch trials. When the observer pressed the joystick button in a catch trial, feedback in the form of an auditory signal was given. The dynamic and static stimuli were blocked separately. Within each block, all target conditions were presented in random order. Each block was separately repeated six times. In each block, the static stimuli consisted of 16 valid, 16 invalid-within, 16 invalid-between, and 12 catch trials, making a total of 60 trials. For the dynamic stimuli, there were eight valid, eight invalid-within, eight invalid-between, eight invalid-space, and eight catch trials, making a total of 40 trials. Thus, the cue was not predictive with respect to target conditions.

Before recording the data, each subject had 540 trials for training purposes. RT data were analyzed by two-factor repeated measures analyses of variance (ANOVAs; with Huynh–Feldt correction for sphericity, as necessary) and with preplanned paired $t$ tests. Pairwise comparisons of the RTs for different target locations were planned as follows (see Fig. 1 for the various target locations): In the static case, the valid location corresponded to the location of the cue according to retinotopic/spatiotopic as well as object-based reference frames. The invalid-within and invalid-between locations correspond to equidistant locations from the cue within the cued-object and within an uncued object, respectively. Comparison of the invalid-within and invalid-between RTs would reveal whether there would be any object-based advantage of the cue. Comparison of invalid-within and valid RTs would reveal whether there is a distant-dependent advantage of the cue inside the cued object. In the dynamic case, the invalid-space location corresponds to the cued location *according to a retinotopic/spatiotopic* reference frame. Valid location corresponds to the cued location *according to a reference frame that moves with the object*. Pairwise comparison of these two RT values allows us to establish which of these two reference frames is more effective for exogenous attention. Invalid-within and invalid-between locations correspond to locations equidistant from the cue *according to a reference frame that moves with the object.* Comparison of these two RTs would allow us to establish whether there was any object-based advantage for moving objects. Finally, comparison of invalid-within and valid RTs was done to assess whether there was any distant-dependent advantage of the cue inside the cued object.

Results and discussion

RTs less than 150 ms and greater than 1,000 ms were excluded from all analyses. These excluded trials constituted 0.7 % of all trials. The accuracy on catch trials was 95.9 % or higher. Figure 2 and Table 1 show RTs averaged across observers, along with standard errors of the means.
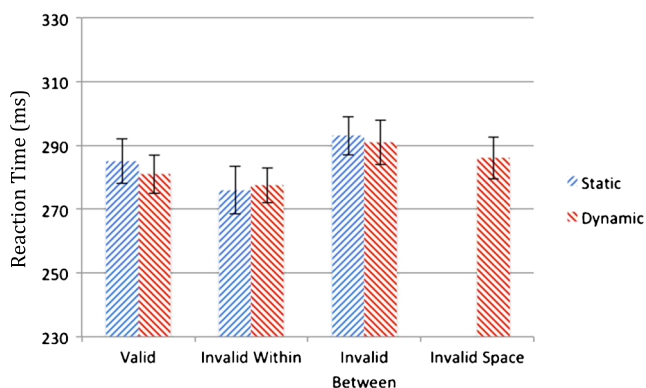


**Fig. 2** Mean reaction times (±standard errors) for the dynamic and static stimulus conditions in Experiment 1

Considering the valid, invalid-within, and invalid-between conditions, we observed no significant difference between moving and static stimuli [$F(1, 14) = 0.110$, $p = .745$, $\eta_p^2 = .008$]. The cue–target relationship had a significant effect [$F(2, 28) = 26.925$, $p < .001$, $\eta_p^2 = .658$]. The interaction between these two main factors was not significant [$F(2, 28) = 1.922$, $p = .165$, $\eta_p^2 = .121$].

For the static condition, in preplanned comparisons, a significant effect of object was found (invalid-within RT = 275 ms vs. invalid-between RT = 292 ms, with cueing effect = 17 ms), $t(14) = -7.055$, $p < .001$, $d = 1.822$. This result replicated previous findings. A significant effect of location within the cued object was also found (valid RT = 285 ms vs. invalid-within RT = 275 ms), $t(14) = 4.172$; $p = .001$, $d = 1.077$; surprisingly, RTs for invalid-within trials were shorter than those for valid trials (cueing effect = 10 ms). The reason for this was not clear; one may speculate that it may be due to the contributions of two factors: (1) the relative probabilities of valid and invalid-within conditions and (2) the possibility of an inhibitory effect exerted by the cue at its immediate neighborhood.

Regarding the first factor, let us note that previous studies had used higher probabilities for the valid condition than for the invalid-within condition (e.g., Avrahami, 1999; Brown et al., 2006; Brown & Denney, 2007; Egly, Driver, & Rafal, 1994; Marino & Scholl, 2005; Marrara & Moore, 2003). Since our goal was to study exogenous attention, we used the same probabilities for all stimulus options. Cueing effects have been shown to depend not only on validity but also on the probabilities of different response options (Shomstein & Yantis, 2004); the difference between valid and invalid-within RTs was much smaller when the probability difference between the two conditions was smaller. For example, Shomstein and Yantis's COA = 200 ms condition yielded a difference between invalid-within and valid RTs of 80 ms when the probabilities of invalid-within and valid were 8.3 % and 50 %, respectively. When the difference between these probabilities was reduced (41.7 % invalid-within and 50 % valid), the invalid-within and valid RT difference was reduced from 80 to 21 ms. Regarding the second factor, other studies have typically highlighted only the edges of the rectangle as the cue stimulus (e.g., Egly, Driver, & Rafal, 1994; Iani et al., 2001), whereas we presented a filled square cue (the reason for our choice of a filled square was to provide an effective cue for the invalid-space condition). It might be possible that a putative inhibitory effect of a filled cue at the valid location, combined with an already small difference between the valid and invalid-within conditions due to probabilities, yielded the shorter RTs for the invalid-within than for the valid condition.

For the dynamic condition, in preplanned comparisons, a significant effect of object was found (invalid-within RT = 278 ms vs. invalid-between RT = 292 ms, cueing effect = 14 ms), $t(14) = -4.249$, $p = .001$, $d = 1.097$. No significant

**Table 1** Mean reaction times (in milliseconds) and standard errors for Experiments 1–3

| | Static | | | | | Dynamic | | | | | I.Space |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Valid | I.Within | I.Between | | | Valid | I.Within | I.Between | | | |
| Exp. 1 | 285 ± 6.7 | 275 ± 6.2 | 292 ± 6.5 | | | 280 ± 6 | 278 ± 5.8 | 292 ± 6.9 | | | 285 ± 6.4 |
| Exp. 2 | 253 ± 8.8 | 241 ± 9.3 | All 260 ± 8.2 | S.C. 234 ± 8.1 | D.C. 291 ± 8.7 | 280 ± 11.2 | 280 ± 10.2 | All 283 ± 7.7 | S.C. 264 ± 8.6 | D.C. 315 ± 7.6 | 275 ± 8.8 |
| Exp. 3 | | | | | | 256 ± 3.6 | 261 ± 3.2 | 269 ± 2.9 | | | 271 ± 2.8 |

effect of location within the cued object was found (valid RT = 280 ms vs. invalid-within 278 ms), $t(14) = 0.779$, $p = .449$, $d = 0.201$. To analyze space- versus object-based cueing, we compared the valid and invalid-space conditions. The difference was small but significant (valid RT = 280 ms, invalid-space RT = 285 ms, cueing effect = 5 ms), $t(14) = -2.244$, $p = .042$, $d = 0.579$, indicating that, for a moving stimulus, exogenous attention is summoned more effectively *according to the motion of the stimulus*, as compared to the static retinotopic/spatiotopic location of the cue.

Overall, the results of this experiment indicated similar properties of exogenous attention for static and moving stimuli; in both cases, exogenous attention is summoned to the whole arc (object) rather than exclusively to the specific location of the cue within the object. Furthermore, dynamic stimuli indicate that this allocation occurs according to a nonretinotopic reference frame that moves following the motion of the stimulus, as opposed to a purely retinotopic reference frame. In fact, as is shown in Fig. 3, if the cueing effect were based on the retinotopic location of the cue, a cue presented in an object would have no effect on this object when the object moved and occupied a different retinotopic location. On the other hand, if the cue is represented according to a reference frame that moves with the stimulus, its location would remain invariant with respect to the moving stimulus, leading to an object-based effect for stimuli in motion, as we have found here.

In a recent study, Theeuwes et al. ([2013](#)) moved a single object (either a rectangle or a cross) and reported an advantage for both the object-centered location and the retinotopic location of the cue within the object, as compared to the uncued locations within the object. The retinotopic and object-centered validity advantages were similar (Theeuwes et al., [2013](#)). In our study, instead of analyzing the effect of cueing within a single object, we examined the effect of cueing on multiple objects. We also had an "invalid-space" condition that probed the spatial effect of the cue *outside* of the object. Although both studies agree that exogenous attention operates according to a reference frame that moves with the object, retinotopic cueing was weaker in our study (valid vs. invalid-space). This may be due to the fact that, since in the Theeuwes et al. study the retinotopic cue was still inside the object, their retinotopic condition was likely to reflect a combination of retinotopic and object effects. In our study, with the invalid-space target being outside the cued object, we did not expect any object-level effect for this target. Therefore, comparing purely retinotopic to purely object-based effects of exogenous attention, we demonstrated an advantage for object-based exogenous attention as compared to the retinotopic case. Let us note, however, that this difference, although significant, was relatively small.

As we discussed in the introduction, although many studies have reported "object"-based attention, what constitutes an object has not been well defined. In the next experiment, we
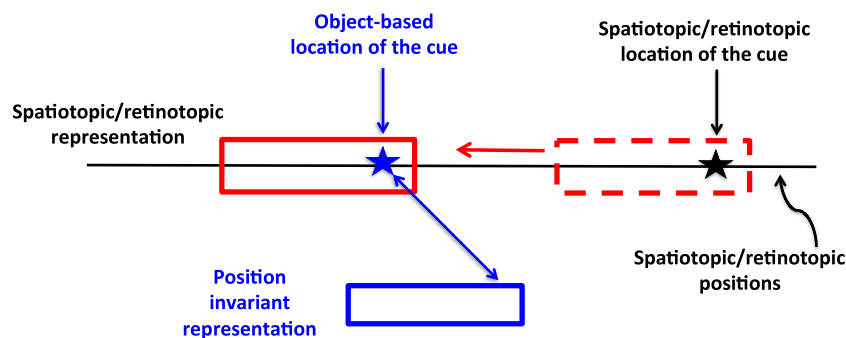


**Fig. 3** An illustration of how an exogenous cue would exert its effect according to spatiotopic, retinotopic, and object-based coordinates. In our experiments, since the eyes were stationary, the predictions of retinotopic and spatiotopic cueing would be the same. A rectangular object is moving from right to left, and the cue is applied to the right end of the rectangle. The star symbols indicate where the cueing would exhibit its effect according to different coordinate systems

modified the stimulus of Experiment 1 to investigate the role of perceptual grouping in the allocation of exogenous attention.

## Experiment 2: the effect of grouping by color on the allocation of exogenous attention

Method

The methods were as in Experiment 1, with the following exceptions. Twelve subjects from the University of Houston population participated in this experiment. Ten of the subjects were new, and two of the subjects were from Experiment 1 subject population.

All arcs had a luminance of 2 cd/m$^2$. In order to study grouping by color, the colors assigned to the arcs were such that, in each trial, two randomly selected arcs had the same color, while the other two had different colors (Fig. 4). Colors were generated on a color wheel, defined on the CIE XYZ coordinates as $X = 0.2044 + .20*\cos(2\pi\alpha/180)$, $Y = 0.48085 + .20*\sin(2\pi\alpha/180)$, $Z = 2.0$. For color randomization, in each trial, three random numbers between 0 and 180 were generated for the angle $\alpha$. They were not allowed to be equal or to have a difference less than 30. The first generated number was used for two arcs chosen randomly from among the four arcs. The remaining two numbers were used for the 3rd and 4th arcs.

If two consecutive arcs had the same color, the invalid-between option became "invalid-between same-color." When the two consecutive arcs had different colors, the invalid-between option became "invalid-between different-color" (Fig. 4). Half of the invalid-between trials were invalid-between same-color, and the remaining half were invalid-between different-color. The static and dynamic conditions were blocked separately, and their order was randomized for each subject. A bright white cue (CIE coordinates 0.2044, 0.4808, 20) was used. The cue was not predictive of target condition: In the static condition, the valid and invalid-within cases each had 97 trials. The invalid-between case had 98 trials. There were also 68 catch trials. In the dynamic condition, the valid, invalid-within, invalid-between, and invalid space conditions each had 48 trials, along with 48 catch trials. As we mentioned before, invalid-between trials were divided equally between invalid-between same-color and invalid-between different-color (49 each in the static case and 24 each in the dynamic case). The aforementioned trials constituted one session, and each subject completed five sessions, giving totals of 1,800 and 1,200 trials for the static and dynamic conditions, respectively. Before recording the data, each subject performed 540 trials for training purposes.
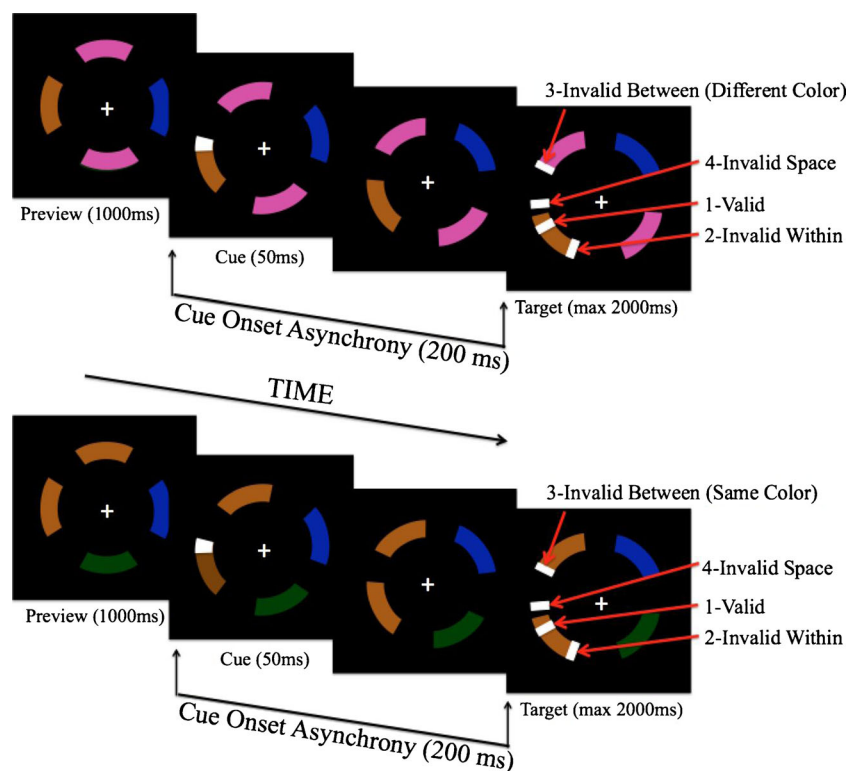


**Fig. 4** Stimuli for Experiment 2. The invalid-between different-color condition is shown in the upper panel, and the invalid-between same-color condition in the lower panel

## Results and discussion

RTs less than 150 ms and greater than 1,000 ms were excluded from all analyses. These excluded trials constituted 0.6 % of all trials. The accuracy in catch trials was 96 % or higher; mean RTs are shown in Fig. 5 and Table 1. A two-way repeated measures ANOVA with the factors Dynamic–Static and Validity (valid, invalid-within, or invalid-between) showed a significant difference between the static and dynamic stimuli [$F(1, 11) = 31.322$, $p < .001$, $\eta_p^2 = .740$]. Cue validity and the interaction between the two factors were also significant [$F(2, 22) = 15.501$, $p < .001$, $\eta_p^2 = .585$, and $F(2, 22) = 10.805$, $p = .002$, $\eta_p^2 = .496$, respectively]. Interestingly, for the dynamic case, RTs in Experiment 2 were similar to those in Experiment 1, whereas the stimuli in Experiment 2 generated overall faster RTs in the static case.

For the static condition, in preplanned comparisons, a significant effect of object was found (invalid-within RT = 241 ms vs. invalid-between RT = 260 ms, cueing effect = 19 ms), $t(11) = -4.15$, $p = .002$, $d = 1.19$. The effect of the location within the cued object was also significant (valid RT = 253 ms vs. invalid-within RT = 241 ms, cueing effect = 12 ms), $t(11) = 1.823$, $p = .009$, $d = 0.52622$, with shorter RTs for invalid-within than for valid trials, in agreement with the findings of Experiment 1.

For the dynamic condition, a significant effect of object was not found (invalid-within RT = 280 ms vs. invalid-between RT = 283 ms), $t(11) = -0.524$, $p = .610$, $d = 0.1513$. This null finding was based on an analysis in which the same- versus different-color cases were lumped together. However, if grouping by color affects exogenous cueing, a significant effect of object should be found by separating invalid-between same-color cases from invalid-between different-color cases. If the effect of the cue applies, not only to the cued arc but also to all arcs that are perceptually grouped with the cued arc, then the invalid-between same-color condition should generate faster RTs than would the invalid-between different-color condition, as would the invalid-within condition as compared to the invalid-between condition.
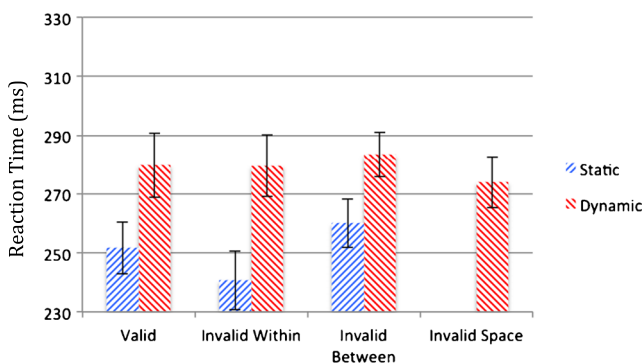
Figure 6 and Table 1 show the RT data according to color-grouping analysis. As was mentioned in the previous paragraph, the effect of object was already significant for the static condition with the combined analysis, and it is also significant when the same-color and different-color cases are separated (invalid-between same-color RT = 234 ms vs. invalid-between different-color RT = 291 ms, cueing effect = 57 ms), $t(11) = -11.899$, $p < .001$, $d = 3.4352$. With this color-grouping analysis, we now find that, for the moving stimulus, too, RTs for the invalid-between same-color condition (264 ms) are significantly shorter than those for the invalid-between different-color condition (315 ms; cueing effect = 51 ms), $t(11) = -5.248$, $p < .001$, $d = 1.515$. These results indicate that exogenous attention is deployed to all stimuli forming a perceptual group, both for static and moving stimuli. Within the group, the effect of the cue was stronger for the uncued element of the group than for the cued element for the static stimulus (valid RT = 253 ms vs. invalid-between-same-color RT = 234 ms, cueing effect = 19 ms), $t(11) = 3.058$, $p = .011$, $d = 0.883$, but not for the moving stimulus (valid RT = 280 ms vs. invalid-between-same-color RT = 264 ms), $t(11) = 1.894$, $p = .085$, $d = 0.5468$.

The effectiveness of the retinotopic/spatiotopic cue was not different from that of the cued element of the group (invalid-space RT = 275 ms vs. valid RT = 280 ms), $t(11) = 1.192$, $p = .26$, $d = 0.344$, nor of the other element in the cued group (invalid-space RT 275 ms vs. invalid-between same-color RT = 264 ms), $t(11) = 1.330$, $p = .210$, $d = 0.3839$. Thus, the small advantage of the object-based cue in comparison to the retinotopic/spatiotopic cue found in Experiment 1 was not observed here.

The results of this experiment show that exogenous attention is drawn, not just to the cued element, but also to other elements that are perceptually grouped with the cued element. These results also confirm the finding from the previous experiment that the reference frame of exogenous attention follows the motion of the stimuli.
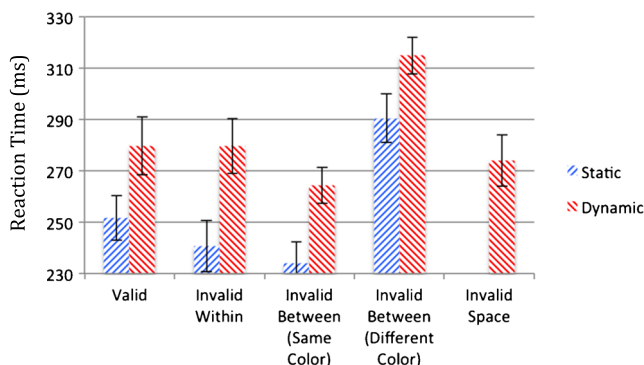


Fig. 5 Mean reaction times (±standard errors) for Experiment 2



Fig. 6 Mean reaction times (±standard errors) for Experiment 2. These are the same data as in Fig. 5, except that the invalid-between condition is separated into invalid-between same-color and invalid-between different-color conditions

222222222

2222222222

Atten Percept Psychophys (2014) 76:1709–1720

# Experiment 3: the effect of grouping by motion on the allocation of exogenous attention

The goal of this experiment was to test further the finding that exogenous attention is summoned to perceptual groups, by using another grouping principle—namely, grouping by common fate. In Experiment 2, all of the arcs had the same motion. Thus, in addition to the grouping of the two arcs with identical color as a unit, all four arcs may have also been grouped by common motion. In fact, the percept was that of four arcs rotating, with two of them forming a group by color. In this experiment, we changed the roles of color and motion by using identical color for all elements in the stimulus, while separating them into two distinct groups on the basis of the direction of motion.

## Method

The apparatus was identical to that described in Experiment 1. The stimuli consisted of six disks, each disk with a diameter of 0.8° of visual angle and a luminance of 4 cd/m². All disks were of a blue color (CIE coordinates: 0.3044, 0.6541, 4) on a black background. As in the previous experiments, the fixation point was a white plus sign (+) placed in the center of the monitor. All disks moved with linear trajectories at a speed of 5°/s. The initial positions of the disks were selected inside an invisible inner circle, not allowing them to reach the edge of the screen for their whole motion duration. The inner circle's diameter was 5°. The stimuli were divided into two groups of randomly chosen three disks. The disks belonging to the same group had the same direction of motion, and the two groups had different directions of motion (Fig. 7). When different disk trajectories crossed, there was no interference, in that the disks continued to move with the same velocity. The cue and

the target were white disks (CIE coordinates: 0.2044, 0.48085, 20) with a smaller diameter (0.6°) than the moving disks. They appeared inside the moving disks (Fig. 7).

Each trial started with a preview: Six disks with a randomly chosen six starting positions, having two different randomly chosen linear trajectories, started their motion. After a preview period of 500 ms, the cue, which was a smaller white disk, appeared in one of the disks and traveled with that disk for 100 ms. The COA interval was 200 ms, during which disks continued their linear motion. After the COA, the target was presented inside one randomly chosen disk. Observers were asked to keep their gaze steady on the fixation cross and to press a joystick button as soon as the target appeared. On catch trials, no target was presented, and subjects were required not to press the joystick button. The subjects' task was to respond to targets as quickly as possible while keeping the highest accuracy possible on catch trials. Feedback was given for incorrect responses on catch trials. The maximum duration of the target was set to 1,000 ms, during which the subject had to press the joystick button.

The valid target appeared in the same disk as the cue. The invalid-within target appeared in a disk moving in the same direction as the cued disk—that is, in a disk that belonged to the same perceptual group as the cued disk. The invalid-between target appeared in a disk that moved with a different direction than the cued disk. In order to control for distance effects, the average distance across trials between the cue and the invalid-within target was the same as the average distance across trials between the cue and the invalid-between target. The last target option was invalid-space, in which the target did not appear in any disk, but appeared in the first retinotopic/spatiotopic location where the cue had appeared. This stimulus bore some similarity to stimuli in the multiple object tracking (MOT) paradigm (Pylyshyn & Storm, 1988).
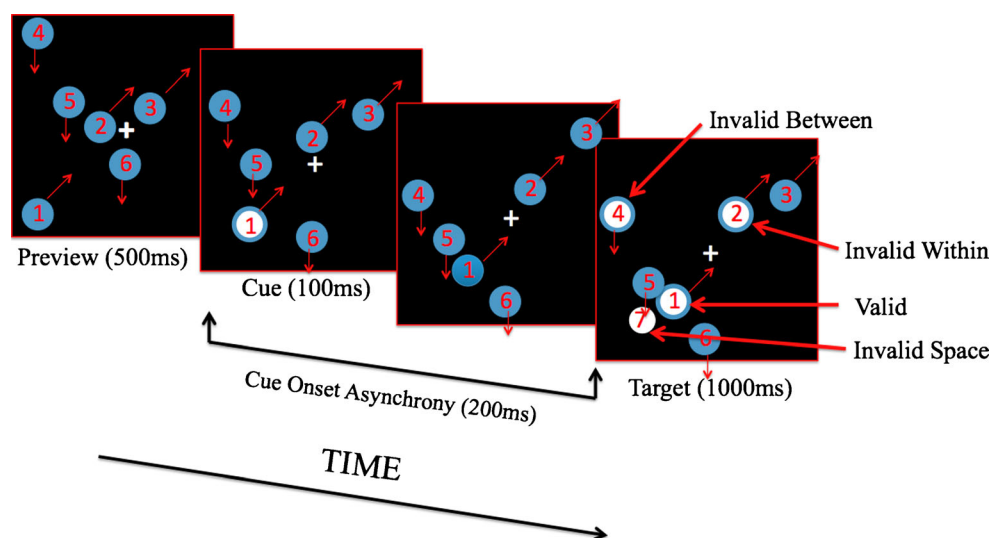


**Fig. 7** Schematic representation of the stimuli in Experiment 3

Typically, in MOT, observers track the identities of a set of preselected moving targets among distractors. Identities are defined at the beginning of a trial, and thus MOT is likely to strongly engage endogenous attention. However, our stimulus design was tailored to primarily engage *exogenous* attention. Moreover, unlike MOT, in which stimuli move in independent directions, we used stimuli that moved with the same direction so as to generate perceptual grouping, to assess how exogenous attention is summoned to perceptual groups.

Trials of all target options were presented in random order, and the cue was not predictive of target location. Each session had 48 trials each for the valid, invalid-within, invalid-between, invalid-space, and catch trials. Each subject completed four sessions, yielding a total of 960 trials. Before data collection, each subject performed 300 trials for training purposes.

Twelve subjects from the University of Houston student population participated in the experiment. Four of the subjects were new, whereas the rest were from the population that had participated in Experiment 2.

Results and discussion

RTs less than 150 ms and greater than 1,000 ms were excluded from all analyses. These excluded trials constituted 1.7 % of all trials. Accuracy on catch trials was 95 % or higher. Mean RTs are shown in Fig. 8 and Table 1. A one-way repeated measures ANOVA showed that the effect of cue validity was significant [$F(3, 33) = 12.725, p < .001, \eta_p^2 = .536$]. The most effective cueing happened for the cued disk, and this was significantly better than for the disks within the same group (valid RT = 256 ms vs. invalid within RT = 261 ms, cueing effect = 5 ms), $t(11) = -3.231, p = .008, d = 0.9328$. We found a significant effect of motion grouping (invalid-within RT = 261 ms vs. invalid-between RT = 269 ms, cueing effect = 8 ms), $t(11) = -5.268, p < .001, d = 1.5207$. Retinotopic/spatiotopic cueing was not effective (invalid space RT = 271 ms vs. invalid-between RT = 269 ms), $t(11) = 0.769, p = .458, d = 0.2218$.
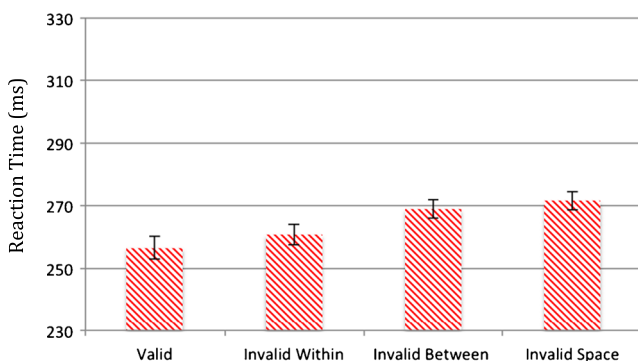


**Fig. 8** Mean reaction times (±standard errors) for Experiment 3

Overall, this experiment also showed that attention is deployed according to perceptual grouping of the stimuli in a nonretinotopic manner: The cued disk received the most attentional resources, even though, due to its motion, it was at a different retinotopic/spatiotopic location than when the cue was presented. Disks that belong to the same perceptual group as the cued disk also attract exogenous attention.

General discussion

Theories of object-based attention

Several theoretical accounts have been offered to explain the operation of object-based attention. For example, according to the *spreading-of-attention* view (equivalently, called *sensory enhancement*; Roelfsema, Lamme, & Spekreijse, 1998), when spatial attention is summoned to a part of an object, it spreads from its original focus to all parts of the object (LaBerge & Brown, 1989). The *shifting-of-attention* hypothesis stipulates that attentional shifts are faster when executed within the same object, as compared to shifts across different objects (Lamy & Egeth, 2002). The *engaging/disengaging* explanation associates costs when attention is engaged with an object and when it is disengaged from an object to shift elsewhere, and suggests that the cost of the disengaging operation is the major source of the object-based advantage (Brown & Denney, 2007). The *attentional prioritization* account explains object-based effects by a process whereby higher priority is given to locations within objects that are already under the focus of attention. The goal of our study was not to evaluate these alternatives, but rather to examine a common concept in all explanations—namely, the concept of object as it pertains to exogenous attention.

Reference frame of exogenous attention

Moving stimuli showed that exogenous attention is allocated according to a nonretinotopic reference frame that follows the motion of stimuli. Whereas we found a consistent nonretinotopic benefit for the cue, the retinotopic/spatiotopic effect of the cue depended on the stimulus configuration.

Our data from Experiment 3 can also be analyzed to examine how the reference frame is established in terms of its origin. In the literature, the term "object-*centered*" is often used (e.g., Peterson et al., 1998; Theeuwes et al., 2013; Tipper, Driver, & Weaver, 1991), implying explicitly or implicitly that the center of the reference frame—that is, its origin—is located on the object, possibly in the center of the object itself. If we consider that the attentional effect declines by distance, we can then plot our data from the third experiment according to the distance in each trial between the putative location of the center of the reference frame and the location of the target.
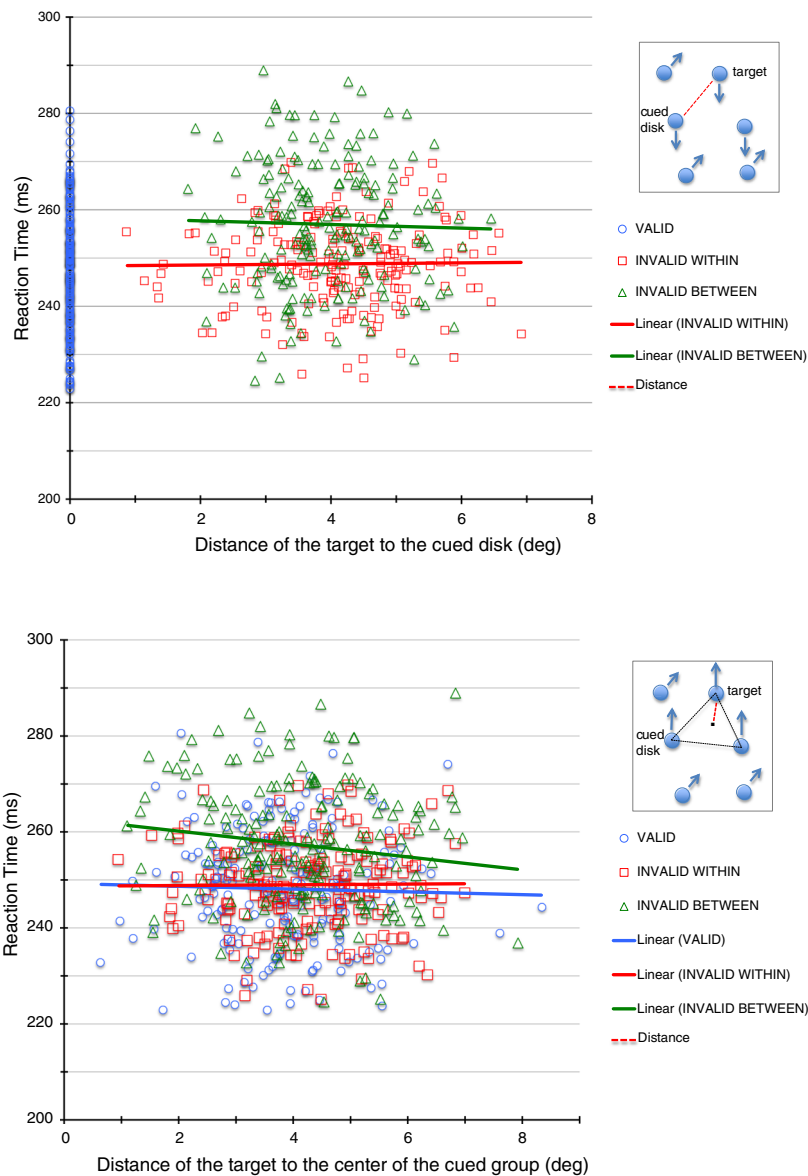
**Fig. 9** Top panel: RTs for one subject, plotted as a function of the distance between the target disk and the cued disk. Each point represents a datum from a single trial. Bottom panel: Similar to the top panel, but plotted as a function of the distance between the target and the center of the cued group. The center of the cued group in each trial was calculated as the mean of the ordinates and abscissae of the disks that had the same direction of motion as the cued disk. The straight lines are linear regression fits. The insets show how the distances on the x-axis were calculated

As an example, the top panel of Fig. 9 shows data from one subject, whose RTs are plotted against the distance of the target disk from the cued disk. Each point in the plot represents a datum from a single trial. If the reference frame is *centered on the cued item*, and if the cueing advantage exhibited a distant-dependent decline, one would expect an increase in RTs as a function of distance. The bottom panel of Fig. 9 shows the same data plotted with respect to the geometric center of the cued group. If the reference frame is *centered on the center of the cued group*, and if the cueing

**Table 2** Average slopes and the results of *t* tests comparing these slopes to 0

|  | Average Slope ± *SEM* and *t*-Test Results, When Distance Is Calculated With Respect to the Cued Disk | Average Slope ± *SEM* and *t*-Test Results, When Distance Is Calculated With Respect to the Center of the Cued Group |
|---|---|---|
| Valid |  | $-0.3327 \pm 0.3231$, $t(11) = -1.03$, $p = .325$ |
| Invalid-within | $0.0486 \pm 0.1427$, $t(11) = 0.341$, $p = .740$ | $-0.0384 \pm 0.1368$, $t(11) = -0.281$, $p = .784$ |
| Invalid-between | $-0.287 \pm 0.1802$, $t(11) = -1.593$, $p = .139$ | $0.1086 \pm 0.2855$, $t(11) = 0.381$, $p = .711$ |

advantage exhibits a distant-dependent decline, one would expect to find an increase in RTs as a function of distance. To assess these hypotheses, we fitted lines to the data of each subject and calculated the slopes. Table 2 gives the average slopes across observers. One can see that none of the slopes is statistically different from zero. Thus, our data do not support the notion that the reference frame is *object-centered*, which would imply a special status for the center of the object, whether it was calculated with respect to the cued element of the group or to the geometric center of the entire group.

## Temporal dynamics of attention and perceptual grouping

According to Gestalt psychologists, attention and grouping constitute two distinct but functionally interdependent processes (e.g., Koffka, 1922, p. 561)—a view supported by several lines of evidence from healthy subjects and from patients with neurophysiological deficits (for a review, see Gillebert & Humphreys, 2014). Our finding that exogenous attention is allocated to perceptual groups is in accordance with this view. An interesting question would be to determine how the dynamics of the two processes can influence their interactions. In our experiment, the cue was presented *after* the presentation of the stimuli. Hence, at the time that the cue was presented, perceptual groups were likely to be already formed, allowing attention to be allocated to perceptual groups. In our future work, we will investigate the relative dynamics of attention and perceptual grouping by varying the cue delay with respect to the onset of different stimulus properties that lead to the establishment of perceptual groups.

## Ecological implications

The ecological role of exogenous attention can be viewed as an orienting mechanism toward stimuli of potential interest. From this perspective, the finding that the reference frame of exogenous attention follows the motion of the stimulus makes sense, in that the visual system needs to deploy attention to where the target *is* rather than where it *was*. Moreover, deploying attention to the entire perceptual group can be advantageous, because the group is likely to have stronger behavioral meaning than a part. For example, when the motion of the leg of an approaching animal triggers exogenous attention, it would be more meaningful to assess the entire animal rather than just the leg—to distinguish, for example, a leopard from a deer.

## References

Avrahami, J. (1999). Object of attention, objects of perception. *Perception & Psychophysics, 61,* 1604–1612. doi:10.3758/BF03213121

Barrett, D. J. K., Bradshaw, M. F., & Rose, D. (2003). Endogenous shifts of covert attention operate within multiple coordinate frames: Evidence from a feature priming task. *Perception, 32,* 41–52.

Boi, M., Vergeer, M., Ogmen, H., & Herzog, M. H. (2011). Nonretinotopic exogenous attention. *Current Biology, 21,* 1732–1737. doi:10.1016/j.cub.2011.08.059

Brown, J. M., Breitmeyer, B. G., Leighty, K. A., & Denney, H. I. (2006). The path of visual attention. *Acta Psychologica, 121,* 199–209.

Brown, J. M., & Denney, H. I. (2007). Shifting attention into and out of objects: Evaluating the processes underlying the object advantage. *Perception & Psychophysics, 69,* 606–618. doi:10.3758/BF03193918

Cheal, M., & Lyon, D. R. (1991). Central and peripheral precuing of forced-choice discrimination. *Quarterly Journal of Experimental Psychology, 43A,* 859–880. doi:10.1080/14640749108400960

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology, 48,* 269–297. doi:10.1146/annurev.psych.48.1.269

Egly, R., Driver, J., & Rafal, R. D. (1994a). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General, 123,* 161–177. doi:10.1037/0096-3445.123.2.161

Egly, R., Rafal, R., Driver, J., & Starrveveld, Y. (1994b). Covert orienting in the split brain reveals hemispheric specialization for object-based attention. *Psychological Science, 5,* 380–383. doi:10.1111/j.1467-9280.1994.tb00289.x

Gillebert, C. R., & Humphreys, G. W. (2014). Mutual interplay between perceptual organization and attention: A neuropsychological perspective. In J. Wagemans (Ed.), *Oxford handbook of perceptual organization.* Oxford, UK: Oxford University Press.

Humphreys, G. W., & Riddoch, M. J. (2007). How to define an object: Evidence from the effects of action on perception and attention. *Mind and Language, 22,* 534–547.

Iani, C., Nicoletti, R., Rubichi, S., & Umiltà, C. (2001). Shifting attention between objects. *Cognitive Brain Research, 11,* 157–164.

Jonides, J. (1981). Voluntary vs. Automatic control over the mind's eye's movement. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp. 187–203). Hillsdale: Erlbaum.

Kasai, T., Moriya, H., & Hirano, S. (2011). Are objects the same as groups? ERP correlates of spatial attnetional guidance by irrelevant feature similarity. *Brain Research, 1399,* 49–58.

Koffka, K. (1922). *Principles of gestalt psychology.* New York: Harcourt, Brace & World, Inc.

LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review, 96,* 101–124. doi:10.1037/0033-295X.96.1.101

Lamy, D., & Egeth, H. (2002). Object-based selection: The role of attentional shifts. *Perception & Psychophysics, 64,* 52–66. doi:10.3758/BF03194557

Lamy, D., & Tsal, Y. (2000). Object features, object locations, and object files: Which does selective attention activate and when? *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1387–1400. doi:10.1037/0096-1523.26.4.1387

Marino, A. C., & Scholl, B. J. (2005). The role of closure in defining the "objects" of object-based attention. *Perception & Psychophysics, 67,* 1140–1149. doi:10.3758/BF03193547

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* San Francisco, CA: Freeman.

Marrara, M. T., & Moore, C. M. (2003). Object-based selection in the two-rectangles method is not an artifact of the three-sided directional

cue. *Perception & Psychophysics, 65,* 1103–1109. doi:10.3758/BF03194837

Moore, C. M., Yantis, S., & Vaughan, B. (1998). Object-based visual selection: Evidence from perceptual completion. *Psychological Science, 9,* 104–110. doi:10.1111/1467-9280.00019

Müller, H. J., & Rabbitt, P. M. (1989). Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 315–330. doi:10.1037/0096-1523.15.2.315

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research, 11,* 1631–1647.

Peterson, M. A., Gerhardstein, P. C., Mennemeier, M., & Rapcsak, S. Z. (1998). Object-centered attentional biases and object recognition contributions to scene segmentation in left- and right-hemisphere-damaged patients. *Psychobiology, 26,* 357–370.

Pinna, B. (2014). What is a perceptual object? Beyond the Gestalt theory of perceptual organization. In A. Geremek, M. W. Greenlee, & S. Magnussen (Eds.), *Perception beyond gestalt: Progress in vision research* (pp. 48–66). Hove, UK: Psychology Press.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3–25. doi:10.1080/00335558008248231

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3,* 179–197. doi:10.1163/156856888X00122

Reppa, I., Schmidt, W. C., & Leek, E. C. (2012). Successes and failures in producing attentional object-based cueing effects. *Attention, Perception, & Psychophysics, 74,* 43–69. doi:10.3758/s13414-011-0211-x

Ro, T., & Rafal, R. D. (1999). Components of reflexive visual orienting to moving objects. *Perception & Psychophysics, 61,* 826–836.

Roelfsema, P. R., Lamme, V. A. F., & Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature, 395,* 376–381. doi:10.1038/26475

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition, 80,* 1–46. doi:10.1016/S0010-0277(00)00152-9

Shomstein, S., & Yantis, S. (2004). Configural and contextual prioritization in object-based attention. *Psychonomic Bulletin & Review, 11,* 247–253. doi:10.3758/BF03196566

Soto, D., & Blanco, M. J. (2004). Spatial attention and object-based attention: A comparison within a single task. *Vision Research, 44,* 69–81.

Theeuwes, J., Mathôt, S., & Grainger, J. (2013). Exogenous object-centered attention. *Attention, Perception, & Psychophysics, 75,* 812–818. doi:10.3758/s13414-013-0459-4

Tipper, S. P., Driver, J., & Weaver, B. (1991). Object-centred inhibition of return of visual attention. *Quarterly Journal of Experimental Psychology, 43A,* 289–298. doi:10.1080/14640749108400971

Vecera, S. P. (1994). Grouped locations and object-based attention: Comment on Egly, Driver, & Rafal (1994). *Journal of Experimental Psychology: General, 123,* 316–320. doi:10.1037/0096-3445.123.3.316

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin, 138,* 1172–1217. doi:10.1037/a0029333

Weichselgartner, E., & Sperling, G. (1987). Dynamics of automatic and controlled visual attention. *Science, 238,* 778–780. doi:10.1126/science.3672124

Wertheimer, M. (1923). Laws of organization in perceptual forms. In: W. A. Ellis (Ed.), *Sourcebook of gestalt psychology* (pp. 71–88). London: Routledge & Kegan Paul.