Task-specific engagement of object-based and space-based attention with spatiotemporally defined objects

Qingzi Zheng¹ · Cathleen M. Moore¹

Accepted: 2 November 2020 / Published online: 4 January 2021
© The Psychonomic Society, Inc. 2021

Abstract

We used a form of ambiguous apparent motion known as *Ternus* motion to isolate the effects of object-based and space-based attention, and to explore functional differences between them. Two frames of horizontally aligned disks that were shifted by one position between frames were temporally separated by either a short or a long inter-stimulus interval (ISI). Short ISI displays were perceived as *element motion* where one disk appeared to jump across the other two. Long ISI displays were perceived as *group motion* where all three disks appeared to move together. Because element and group motion imply mutually exclusive object structures, adding stimuli (e.g., a small gap) to one disk in each frame created conditions of orthogonal object and location status (same or different), depending on ISI. We used two tasks with different functional demands, an identification task (Experiments 1 and 3a) in which observers responded to a single attribute of the final stimulus, and a comparison task (Experiments 2 and 3b) in which observers compared two attributes across two stimuli. Reliable object-specific effects occurred only with the comparison task, whereas location-specific effects occurred with both tasks. These results confirm that attention can be directed to objects separately from spatial locations and vice versa, and, moreover, that object-based and space-based attention are engaged differently depending on the processing demands of the task.

Keywords Attention: object-based · Attention: space-based · Perceptual organization

Introduction

At any given moment, we are surrounded by an overwhelming amount of visual information, only some of which is relevant to our current needs and goals. The set of mechanisms by which relevant information is prioritized and processed differently from irrelevant information is collectively referred to as *visual attention*. When driving, for example, we focus our attention on the road in the front of us as well as on cars that pass us or switch into our lane. At the same time, we try to avoid distraction from cars and other objects that are not in our path. These examples highlight that the focus of attention can be on particular regions of space (e.g., in the road vs. off to the side of the road) or on objects in the world (e.g., another car in the oncoming lane of traffic vs. a pedestrian on the sidewalk). Few methods for studying visual attention are designed to

measure both types of selection, and those that are designed to are often limited in the conclusions that can be drawn from them. In this study, we sought to simultaneously measure effects of space-based and object-based attention in the context of dynamic displays in which, over time, multiple objects could occupy a single location and individual objects could occupy multiple locations.

Perhaps the most frequently used method of measuring space-based and object-based attention is the two-rectangles paradigm of Egly, Driver, and Rafal (1994). Building on earlier spatial-cueing paradigms (e.g., Posner, 1980; Posner, Nissen, & Ogden, 1978), Egly et al. presented displays with two identical rectangles that were aligned with each other such that their respective ends were separated by the same distance as their length. At the beginning of each trial, one end of one rectangle brightened as a cue, after which a stimulus square filled one of the ends of one of the rectangles and the task was to push a button as quickly as possible following the presentation of the square. The square could appear in one of three locations: the cued end of the cued rectangle (*valid* trials), the uncued end of the cued rectangle (*invalid same-object* trials), or the equidistant end of the uncued rectangle (*invalid different-object* trials). The typical finding is that responses tend to

✉ Qingzi Zheng
qingzi-zheng@uiowa.edu

¹ Department of Psychological and Brain Sciences, G60 PBSB,
University of Iowa, Iowa City, IA 52242, USA

be faster in the valid condition than in the invalid same-object condition, an effect that is attributed to space-based attention because both locations are within the cued object and therefore differ only with regard to their distance from the cue. In addition, responses tend to be faster in the invalid same-object condition than in the invalid different-object condition, an effect that is attributed to object-based attention because the two uncued locations are equidistant from the cue but one is in the cued object while the other is not (for reviews, see Chen, 2012; Reppa, Schmidt, & Leek, 2012). However, the interpretation of results from the two-rectangles paradigm has been complicated by an accumulation of findings. First, the object-specific effect varies with the distance between the rectangles, indicating that it is not spatially invariant (Vecera, 1994). Second, the pattern of effects varies with the orientation of the rectangles (horizontal vs. vertical) and whether an object crosses over one of the meridians (horizontal or vertical) of the visual field (Al-Janabi & Greenberg, 2016; Barnas & Greenberg, 2019; Hein, Blaschke, & Rolke, 2017; Pilz, Roggeveen, Creighton, Bennett, & Sekuler, 2012). A recent extensive examination of these spatial dependencies by Chen and Cave (2019) suggests that object-specific effects in the two-rectangles paradigm are caused by object structure modifying existing biases of how spatial selection mechanisms are prioritized, rather than object-based attention that is distinct from space-based attention.

An alternative approach of isolating object-based attention from space-based attention is to present multiple objects that overlap in space, thus occupying more-or-less the same spatial location (Duncan, 1984). Performance on tasks that require reporting information from one or both objects is then compared. Accuracy tends to be higher when reporting two attributes from a single object than when reporting two attributes from two different objects. Moreover, there is sometimes no additional cost in reporting two attributes of a single object compared to reporting only a single attribute of that object. Given that the two objects were at the same location, these results were offered as evidence that attention can be allocated to individual objects separate from their spatial locations. A concern regarding Duncan's original study is that spatial differences were not fully controlled because the overlapping stimuli differed in their spatial-frequency content (Watt, 1988; but see Baylis & Driver, 1993; Ernst, Palmer, & Boynton, 2012; Valdes-Sosa, Cobo, & Pinilla, 1998). More critical for the purpose of the present study, however, is that although logically the design can provide a measure of object-based attention isolated from space-based attention, there is no opportunity to assess whether space-based attention has also been engaged, because stimuli only appeared at a single location.

In the present study, we used a form of ambiguous apparent motion, known as *Ternus motion* (Pikler, 1917; Ternus, 1926), to separately measure effects of space-based and object-based attention simultaneously within physically

identical stimuli other than their timing. As illustrated in Fig. 1A, Ternus motion consists of two cycling frames that are separated by a blank inter-stimulus interval (ISI). In each frame, there are three horizontally aligned disks, which are shifted by one position across frames (Movie 1). When the ISI is short, people tend to perceive *element motion*, which looks like one disk jumping back and forth as an element across two other disks, which appear to remain stationary. In contrast, when the ISI is long, people tend to perceive *group motion*, which looks like all three disks moving back and forth together as a group. Ternus motion is especially useful for separating the effects of space-based and object-based attention because even though the disks and their spatial locations within the displays are identical, the two different motion perceptions imply different, and mutually exclusive, object structures, as defined by the correspondence of stimuli across time. As shown in Fig. 1B, we can identify the three disks as objects A, B, and C. A', B', and C' indicate the correspondence between stimuli across the two frames, given the different motion perceptions. For element motion, object B, for example, is at the same location in frames 1 and 2, whereas for group motion, it is at different locations across the two frames. In this way, multiple objects can be tested at a single location and a single object can be tested at multiple locations, all within the context of identical stimuli.

Figure 2 illustrates how adding stimuli to one disk in each of the two frames in Ternus displays sets up orthogonal comparisons of object and location relationships between those stimuli. The top row shows conditions in which two stimuli ("1" and "2" in Fig. 2) are at the same location, and the bottom row illustrates conditions in which they are at different locations. Comparing performance in conditions across rows provides a measure of location-specific effects. The left column illustrates conditions in which the two stimuli appear on the same object, and the right column illustrates conditions in which they appear on different objects. Comparing performance across columns provides a measure of object-specific effects. Notice that there is a short-ISI condition and a long-ISI condition for each of the comparisons. This is a feature of the design in that any differences due to ISI alone are balanced. However, it also means that the design does not allow for an assessment of possible interactive effects of locations and objects, because they would be confounded with any main effect of ISI there may be.

We used this design to investigate the effects of space-based and object-based attention in two different tasks. One was a simple identification task and the other was a comparison task. Specifically, Experiment 1 used an identification task in which the first stimulus was essentially a cue (one of the disks suddenly enlarging) and the second stimulus was a small gap on one side of a disk (see Exp. 1 in Fig. 3). This cue is uninformative, because it does not predict which disk the gap would appear on, but as an abrupt looming stimulus it was

Fig. 1 (A) *Illustration of Ternus motion.* Two frames, each showing three disks but shifted by one position across frames, are shown in cycling displays. When the inter-stimulus interval (ISI) between frames is short, element motion is perceived: one disk appears to hop over the other two (left). When the ISI is long, group motion is perceived: all three disks appear to move as a group (right). See movie 1 for a demonstration. (B) *Illustration of the different object correspondences that are implied by perceived element and group motion.* The three disks are identified as

objects A, B, and C, where A', B', and C' indicate the correspondence between stimuli across frames defining object structure. Consider object B. For element motion, object B is at the same location in frames 1 and 2, whereas for group motion, object B is at different locations across the two frames. Thus, despite being derived from identical frames of stimuli, element and group motion perceptions imply the perceptions of two mutually exclusive object structures

Fig. 2 *Illustration of the experimental design.* Adding a stimulus to a single disk in each frame of the Ternus display allowed for separate measures of location-specific and object-specific effects. Conditions were defined by the relations between stimulus 1 and stimulus 2. Specifically, the two stimuli could appear to be parts of a single object at a single location, a single object at different locations, different objects at the same location, or different objects at different locations. Object-specific effects were calculated by comparing conditions from the left column (Same Object) with those from the right column (Different Object). Location-

specific effects were calculated by comparing conditions from the top row (Same Location) with those from the bottom row (Different Location). Notice that the same-object and different-object conditions, and the same-location and different-location conditions, each includes a long-inter-stimulus interval (ISI) and a short-ISI version. Thus, any object-specific or location-specific effects that occur cannot be attributed to ISI. A limitation of the design is that the interaction between object status and location status cannot be measured because it is confounded with the main effect of ISI

Fig. 3 Illustration of trial events for each of the four experiments. Timing was identical across all experiments. After the 4- to 6-cycle entrainment period in which the display cycled with either a 0- or 200-ms inter-

stimulus interval (ISI), stimulus1 and stimulus 2 were presented. Notice that Experiments 3a and 3b used identical stimuli; they differed only in the tasks

expected to capture attention (e.g., Franconeri & Simons, 2003). The task was to identify whether the gap was on the left or right side of the disk. The question addressed in this experiment was would the speed and/or accuracy of identification be affected by whether the cue was at the same location, same object, or both as the to-be-identified gap stimulus? Experiment 2 used a comparison task, in which both stimuli were gaps on the sides of disks, and the task was to compare whether they were on the same side (i.e., both on the left or both on the right) or different sides (i.e., one on the left and one on the right, or vice versa) of their respective disks (see Exp. 2 in Fig. 3). The question addressed in this experiment was would the speed and/or accuracy of the comparison be affected by whether the disks in which the gaps appeared were at the same location, were perceived as the same object, or both? Experiments 3a and 3b were similar to Experiments 1 and 2, except that they used identical displays and differed only in the task (see Exp. 3a & 3b in Fig. 3).

We measured the effects of space-based and object-based attention with two different tasks, because we hypothesized that if there are separate types of attention, then it is likely that they serve different functions. Based on their processing requirements, the two tasks were expected to engage those attention systems differently and yield different patterns of effects. Specifically, the comparison task was inspired by the object file framework of Kahneman, Treisman, and Gibbs

(1992). The world consists of objects, and objects (as well as viewers) move. Therefore, multiple objects can occupy a single location over time, and single objects can occupy multiple locations over time. Highlighting these qualities, Kahneman, Treisman, and Gibbs (1992) introduced the construct of an *object file* as a spatiotemporally defined representation of an entity in the world with which information, such as identity and spatial history, can be associated and updated over time. Under this view, attending to an object triggers a reviewing process such that information that was previously associated with that object file is retrieved and, through an impletion process, is integrated with current information, so that a continuous object that changes over time is perceived. The comparison task, unlike the identification task, requires comparing information across time. Given this, we hypothesized that the comparison task would be more likely to benefit from object structure than the identification task would, and therefore would yield larger object-specific effects. Contrasting these two tasks, therefore, was an opportunity to reveal differential engagement of the two attention systems, based on functional task demands, within the context of identical displays.

To preview the results, the overall pattern across the set of experiments was that while location-specific effects were observed with both the identification task and the comparison task, object-specific effects were observed only with the comparison task.

General method

Participants Participants were students at the University of Iowa who received credit for a requirement in an introductory Psychology course. All reported normal or corrected-to-normal visual acuity and were naïve with respect to the purpose of the experiment. Overall, 25 (17 female and eight male; 18–22 years old), 25 (12 female and 13 male; 18–22 years old), 36 (32 female and four male; 18–26 years old), and 36 (32 female and four male; 18–19 years old) individuals participated in Experiments 1, 2, 3a, and 3b, respectively. No individual participated in more than one experiment. The numbers of participants were chosen based on estimates of effect sizes from a prior experiment that was similar to those reported here but had a reduced design. A power analysis based on those estimates indicated that a sample of 12 participants would provide .80 power for the smallest expected effects. To be conservative, we ran 25 participants in Experiments 1 and 2, respectively. We then increased our sample size for Experiments 3a and 3b because we anticipated wanting to compare results between experiments. After testing the predetermined number of participants for a given experiment, we calculated an overall error rate for each, and replaced any that exceeded more than twice the average across all participants in that experiment. This resulted in the replacement of one participant in each of Experiments 2, 3a, and 3b. In addition, we replaced one participant in each of Experiments 1 and 3b because they did not complete the session.

Apparatus All experiments were run on an Apple Mac Mini computer (OS X, Version 10.9.5). Stimuli were presented on an EIZO FS2434 LCD monitor with a resolution of $1,920 \times 1,080$ pixels and a frame rate of 16.67 ms. Viewing distance was fixed at 57 cm, using a chin rest. Participants responded by pressing keys on a standard keyboard. Stimulus events were controlled, and responses were recorded using MATLAB (version R2013b, 8.2, The Mathworks Inc., Natick, MA, USA) with the Psychophysics toolbox extension (version 3.0.11; Brainard, 1997; Pelli, 1997).

Stimuli Stimuli were cycling displays, with each cycle consisting of two frames separated by a variable ISI. Each frame included three horizontally aligned black (1.3 cd/m^2) disks presented on a gray (64.3 cd/m^2) background (see Fig. 3). Disks were 1.1° in diameter and were separated center-to-center by 1.6° . In one of the two frames, the center disk appeared 1° above a black $0.3^\circ \times 0.3^\circ$ central fixation cross. In the other frame, the entire set of disks was shifted one position to the left or to the right. The shift was to the left or right of the center equally often across trials, but was always the same within a trial. The ISI was either 0 or 200 ms. These were

selected as values that we expected to reliably induce element motion and group motion, respectively.¹ Cycles began with the centered display, and continued for 4–6 cycles (randomly selected) to entrain a given motion percept. Either on the second frame of the final cycle of the entrainment period or on the first frame of the cycle immediately following the entrainment period, “stimulus 1” was presented, and in the following frame, “stimulus 2” was presented. In Experiment 1, stimulus 1 was one of the three disks (equally often) being enlarged from 1.1° to 1.4° in diameter. Stimulus 2 was a small 0.1° gap that appeared on either the left or right edge of one of the three disks (equally often). Right-side gaps appeared at the 2 o’clock (upper) or 4 o’clock (lower) position, equally often. Left-side gaps appeared at the 10 o’clock (upper) or 8 o’clock (lower) position, equally often. In Experiment 2, both stimuli 1 and 2 were small 0.1° gaps that appeared on the left or the right side, respectively. No specific gap location was repeated, so that if gap 1 was on the same side of a disk as gap 2, then it would be presented at the other gap location on that side of the disk (e.g., gap 1 at the 8 o’clock position and gap 2 at the 10 o’clock position). In Experiments 3a and 3b, stimulus 1 was an enlarged (1.4°) disk that also had a gap (0.15°) on one side, and stimulus 2 was a small (0.1°) gap on the left or right side (equally often) of one of the three disks (equally often).

Task The task in Experiments 1 and 3a was an *identification task*, which was to identify whether the final gap was on the left or right side of the disk in which it appeared by pressing the “F” or “J” key with the left or right index fingers, respectively. For Experiments 2 and 3b, the task was a *comparison task*, which was to compare the locations of the two gaps and report whether they were on the same side of their respective disks (i.e., both on the left side of a disk or both on the right side of a disk) or on different sides (i.e., one on the left side of a disk and the other on the right side of a disk, or vice versa) by pressing the “F” or “J” key with the left or right index fingers, respectively. Participants were asked to make their responses as quickly as possible while not making too many errors.

Procedure Protocols were approved by the University of Iowa Institutional Review Board. Each participant completed a single session that lasted approximately 1 h. Following an informed consent process, participants were provided with a set of written instructions, presented on the computer screen,

¹ To confirm that our displays elicited the expected Ternus motion perceptions, we conducted a separate experiment with 20 different participants. The same displays as those in Experiments 3a and 3b (see Fig. 3) were presented, and participants reported whether a given display looked like a single disk moving across the others as “element motion,” whether all the disks appeared to move together as “group motion,” or “neither.” For 0-ms ISI trials, $90 \pm 3.6\%$ responses were “element motion,” whereas $7 \pm 2.1\%$ were “group motion” and $3 \pm 1.7\%$ were “neither.” In contrast, for 200-ms ISI trials, $92 \pm 3.5\%$ responses were “group motion,” whereas $5 \pm 1.6\%$ were “element motion,” and $3 \pm 2.4\%$ were “neither.”

which the experimenter read aloud while the participant followed along. Participants then completed two blocks of 20 practice trials each. For the second practice block, trial response time (RT) was presented after each trial to motivate the speeded aspect of the task. This feedback was not provided in any other block of trials. Following practice, participants completed seven blocks of 74 trials each, the first two trials in every block were warm-up trials, and the first full block of trials was considered practice. Neither warm-up nor practice trials were included in subsequent analysis. Participants could rest as much as they liked between blocks and initiate the next block by pressing the space bar. Feedback was provided between blocks in the form of mean RT and accuracy for the preceding block and how many blocks remained in the session. In addition, if accuracy was higher than 99% across two consecutive blocks, an additional message was provided to encourage faster responding: “You are doing great, and you have made very few errors over the past two blocks. Please try going a bit faster. It is okay to make a few errors. Thank you!” If accuracy was less than 95% across two consecutive blocks, an additional message was provided to encourage more accurate responding: “You are doing great, but you made a fair number of errors over the last two blocks. Please try to slow down a little bit to make fewer errors. Thank you!” At the end of the final block, a written message indicated that the experiment was complete and that the participant could go out to the main lab. The experimenter then provided a brief explanation of the experiment and answered any questions.

The timing of trial events was identical across all four experiments, as illustrated in Fig. 3. Trials began with the presentation of a black fixation cross on a gray background (not shown) for 500 ms, which was followed by the first frame of the Ternus display. After the 4- to 6-cycle entrainment period in which the display cycled with either a 0- or a 200-ms ISI, stimulus 1 and stimulus 2 were presented. The disks were removed immediately after stimulus 2, leaving just the fixation cross, which was removed when a key press was made. If the response was correct, it was followed by a 500-ms intertrial interval (ITI), after which a fixation cross for the next trial was shown. If the response was incorrect, a 200-ms 500-Hz tone was played and a written error message was presented for 2,000 ms before the 500-ms ITI.

Design and analyses A 2 (Object: same, different) $\times$ 2 (Location: same, different) within-subjects design was used in each experiment. Individual trial types were created through a set of factorially manipulated display factors and were then assigned to one of the conditions of the design based on the relation between stimuli 1 and 2 across frames. Specifically, ISI was either short or long. The direction of the shift from the center was either leftward or rightward. The disk on which stimulus 1 appeared could be the left, center, or right disk in

the array, and the disk on which stimulus 2 appeared could also be the left, center, or right disk in the array. Finally, the two frames in which stimulus 1 and stimulus 2 were presented could be either the first and second frames of a cycle, or the second and first frames of two consecutive cycles. In this way, nothing about stimulus 1 predicted anything task-relevant about stimulus 2. Given all of these conditions, there were 72 unique trial types. Only some of them met the criteria for the four conditions of the 2×2 design. “Same Location” trials were defined as those in which stimuli 1 and 2 appeared on disks at exactly the same image location across frames, and “Different Location” trials were defined as those in which stimuli 1 and 2 appeared on disks that were one position away from each other (see Fig. 2). None of the other distances were included in the analysis because although they were different-location conditions, they were also always different-object conditions, and therefore could not be matched with same-object conditions for comparison. Another critical step in assigning trials to conditions was to ensure that an equal number of short-ISI trials and long-ISI trials contributed to each condition mean. This was important because if there was any effect of ISI itself, then over-representation of one ISI condition over the other would contaminate estimates of that particular experimental condition. Given the full factorial design, balancing the number of observations across conditions required an elimination of a third of the different-location trials. This is because there are a third more ways for two disks to be adjacent to each other across frames (i.e., different location) than there are to be at exactly the same location across frames (i.e., same location). In Experiment 1, the average number of observations per condition was 44 trials (with a minimum of 34 trials and a maximum of 57 trials) for each participant. In Experiment 2, the average number of observations per condition was 36 trials (with a minimum of 26 trials and a maximum of 52 trials) for each participant. For Experiments 3a and 3b, there were 48 observations per condition for each participant. The number of observations differed across experiments because trial types were selected from the set of 72 possible types randomly with replacement for Experiments 1 and 2, whereas they were selected randomly without replacement (i.e., pseudo-randomly) for Experiments 3a and 3b.

The main dependent measure was RT. RTs that were shorter than 250 ms or longer than 3,000 ms were excluded from the analysis, resulting in the exclusion of 0.54% of trials in Experiment 1, 0.72% in Experiment 2, 0.40% in Experiment 3a, and 0.26% in Experiment 3b. Error rates were also evaluated to confirm that speed-accuracy trade-offs did not contribute to the patterns of observed effects. Analyses of error rates were conducted on their arcsin transformations. Alpha was set at .05 throughout. Effect sizes are reported as *adjusted partial eta-squared* ($adj \hat{\eta}_p^2$), which is an estimate of partial eta-squared that adjusts for its positive bias (Mordkoff, 2019).

Results and discussion

Table 1 shows the mean RTs and error rates (ERs) for the main comparisons for each of the four experiments. These correspond to the marginal means of the design illustrated in Fig. 2 (i.e., same object, different object, same location, different location). Figure 4 summarizes the results in terms of the object-specific effect (different object – same object) and the location-specific effect (different location – same location) for each experiment. The individual cell means of the design that is illustrated in Fig. 2 are given in the Appendix. Note, however, that caution is needed in interpreting the individual means in terms of location versus object status. Because of the way objects are defined using ISI to induce element versus group motion perceptions, the interaction between location and object status is confounded with ISI. For this reason, we focus on the marginal means that are summarized in Table 1 and the main effects shown in Fig. 4.

Mean RTs and the arcsin transformations of the ERs for each experiment were submitted to separate 2 (Object: same, different) $\times$ 2 (Location: same, different) repeated-measures analyses of variance (ANOVA). As noted, the interaction between Object and Location also reflects the main effect of ISI.

For Experiment 1, which used the identification task, there was a reliable main effect of Location, $F(1,24) = 12.63$, $p = .002$, $adj \hat{\eta}_p^2 = .317$, confirming that RTs were shorter when the gap appeared in the disk that was at the same location in the image as the first stimulus (i.e., the enlarged disk, which in this task can be thought of as a cue), but no main effect of Object, $F(1,24) = 2.41$, $p = .134$, $adj \hat{\eta}_p^2 = .053$. The interaction term was not significant, $F(1,24) = 2.07$, $p = .163$, $adj \hat{\eta}_p^2 = .041$. The same analysis on the ERs revealed the same pattern, except that the interaction term was significant, $F(1,24) = 5.26$, $p = .031$, $adj \hat{\eta}_p^2 = .146$, likely indicating that more errors were made on short-ISI trials than long-ISI trials (mean difference = $1.6 \pm 0.72\%$).

Fig. 4 Summary of object-specific and location-specific effects across the four experiments. Object effects (Different Object – Same Object) and Location effects (Different Location – Same Location) are shown separately for both the response times (**top**) and the error rates (**bottom**). Error bars are within-subject standard errors. Asterisks indicate statistical significance

For Experiment 2, which used the comparison task, there was a reliable main effect of Location, $F(1,24) = 11.74$, $p = .002$, $adj \hat{\eta}_p^2 = .301$, confirming that RTs were shorter when the second gap appeared in the disk at the same location as the disk with the first gap, and a reliable main effect of Object, $F(1,24) = 17.78$, $p < .001$, $adj \hat{\eta}_p^2 = .402$, confirming that RTs were shorter when the disk with the second gap was perceived to be the same object as the disk with the first gap. The interaction term was not significant, $F(1,24) = 3.85$, $p = .062$, $adj \hat{\eta}_p^2 = .102$. The same analysis conducted on the ERs again revealed the same pattern as the RTs except that the

Table 1 Mean response times (RTs) and error rates (ERs) with standard errors for the marginal means of the design shown in Fig. 2

Experiments		Object		Location	
		Same	Different	Same	Different
1: Identification	RT (ms)	571 $\pm$ 18.7	579 $\pm$ 17.2	565 $\pm$ 17.4	585 $\pm$ 18.6
	ER (%)	4.0 $\pm$ 0.56	4.7 $\pm$ 0.63	3.8 $\pm$ 0.54	4.9 $\pm$ 0.58
2: Comparison	RT (ms)	798 $\pm$ 34.2	847 $\pm$ 35.4	806 $\pm$ 34.8	839 $\pm$ 34.5
	ER (%)	8.1 $\pm$ 1.02	12.5 $\pm$ 1.29	8.4 $\pm$ 1.12	12.2 $\pm$ 1.27
3a: Identification	RT (ms)	589 $\pm$ 13.6	586 $\pm$ 13.5	583 $\pm$ 13.2	591 $\pm$ 14.0
	ER (%)	5.5 $\pm$ 0.54	5.9 $\pm$ 0.41	5.4 $\pm$ 0.54	5.9 $\pm$ 0.45
3b: Comparison	RT (ms)	793 $\pm$ 19.9	820 $\pm$ 22.1	801 $\pm$ 21.7	812 $\pm$ 20.6
	ER (%)	7.6 $\pm$ 0.80	9.4 $\pm$ 0.95	8.1 $\pm$ 0.78	8.9 $\pm$ 0.95

interaction term was significant, $F(1,24) = 26.43, p < .001$, $adj \hat{\eta}_p^2 = .504$, again likely indicating that more errors were made on short-ISI trials than on long-ISI trials (mean difference = $4.5 \pm 0.90\%$).

The pattern of results across Experiments 1 and 2 is consistent with the hypothesis that the comparison task, which required relating two pieces of information across time, would benefit more from object structure through which a reviewing process could occur, than the identification task, which required access to only a single piece of information. Both tasks can be completed without regard to object structure of the displays, but to the extent that comparisons of within-object information is more efficient than comparisons of between-object information, the object structure was expected to impact performance more in the comparison task of Experiment 2 than in the identification task of Experiment 1, which it did.

Experiments 3a and 3b were conducted to confirm these task differences while matching the details of the stimuli exactly. Unlike Experiments 1 and 2, the stimuli and trial events across Experiments 3a and 3b were identical (see Fig. 3). The only difference between experiments was the task. In Experiment 3a, the task was to identify whether the final gap was on the left or right side of the disk on which it appeared, whereas in Experiment 3b, the task was to compare whether the two gaps were on the same or different sides of their respective disks. As in Experiment 1, which also used an identification task, there was no reliable main effect of Object in Experiment 3a, $F(1,35) = 0.48, p = .493$, $adj \hat{\eta}_p^2 = -.015$.² In contrast, as in Experiment 2, which also used the comparison task, there was a reliable main effect of Object in Experiment 3b, $F(1,35) = 11.13, p = .002$, $adj \hat{\eta}_p^2 = .199$. Neither Experiment 3a nor 3b showed a reliable main effect of Location, 3a: $F(1,35) = 3.45, p = .072$, $adj \hat{\eta}_p^2 = .064$, 3b: $F(1,35) = 1.65, p = .208$, $adj \hat{\eta}_p^2 = .018$. Finally, the interaction term was significant in both experiments, likely indicating that RTs were longer on short-ISI trials than long-ISI trials, Experiment 3a, $F(1,35) = 112.73, p < .001$, $adj \hat{\eta}_p^2 = .756$, mean difference = 65 ± 6.2 ms; Experiment 3b, $F(1,35) = 59.83, p < .001$, $adj \hat{\eta}_p^2 = .620$, mean difference = 80 ± 10.4 ms. Finally, the same analyses conducted on the ERs revealed identical patterns of effects.

The lack of a reliable location-specific effect in Experiments 3a and 3b is different from Experiments 1 and 2. As can be seen in Fig. 4, however, the general trends across these two sets of experiments are the same. For some reason, the magnitudes of all effects were smaller in Experiments 3a and 3b compared to Experiments 1 and 2. Given the similarity of designs, we compared performance across experiments and tasks by conducting a

post hoc 2 (Experiment Set: 1 and 2, 3a and 3b) $\times$ 2 (Task: identification, comparison) $\times$ 2 (Object: same, different) $\times$ 2 (Location: same, different) mixed ANOVA with Experiment Set and Task as between-subject variables and Object and Location as within-subject variables. Unsurprisingly, RTs were reliably longer for the comparison task (809 ± 18.4 ms) than for the identification task (582 ± 10.8 ms), $F(1,118) = 118.36, p < .001$, $adj \hat{\eta}_p^2 = .497$. There was, however, no evidence that this differed across the two sets of experiments, as the interaction between Task and Experiment Set was not significant, $F(1,118) = 0.45, p = .502$, $adj \hat{\eta}_p^2 = -.005$. The main effect of Object was reliable (mean difference = 18 ± 4.0 ms), $F(1,118) = 29.58, p < .001$, $adj \hat{\eta}_p^2 = .194$, as was the main effect of Location (mean difference = 17 ± 3.8 ms), $F(1,118) = 23.07, p < .001$, $adj \hat{\eta}_p^2 = .156$. Thus, in general, two stimuli that appeared on the same object were processed more efficiently than two stimuli that appeared on different objects, regardless of whether they were at the same location or not. And in general, two stimuli that appeared at the same location were processed more efficiently than two stimuli that appeared at different locations, regardless of whether they were on the same object or not. Most relevant to the question of whether these effects depend on task, however, is the finding that the interaction between Object and Task was reliable, $F(1,118) = 22.14, p < .001$, $adj \hat{\eta}_p^2 = .151$, confirming that the effect of object status was larger for the comparison task (mean difference = 35 ± 6.7 ms) than for the identification task (mean difference = 2 ± 3.0 ms). In contrast, the interaction between Location and Task was not reliable, $F(1,118) = 1.12, p = .293$, $adj \hat{\eta}_p^2 = .001$. Therefore, there was no evidence that the impact of location status differed across tasks. Finally, neither the Object $\times$ Task $\times$ Experiment Set interaction, $F(1,118) = 0.62, p = .431$, $adj \hat{\eta}_p^2 = -.003$, nor the Location $\times$ Task $\times$ Experiment Set interaction, $F(1,118) = 0.35, p = .553$, $adj \hat{\eta}_p^2 = -.005$, was reliable, indicating that although the magnitudes of the base effects may have been different (Object $\times$ Experiment Set: $F(1,118) = 4.80, p = .030$, $adj \hat{\eta}_p^2 = .031$; Location $\times$ Experiment Set: $F(1,118) = 4.95, p = .028$, $adj \hat{\eta}_p^2 = .032$), there is no evidence that the differential effects of Object status versus Location status in the two tasks were different in the first (Experiments 1 and 2) and second (Experiments 3a and 3b) sets of experiments. The same analysis was conducted on the ERs, and no effects that were in a different direction or suggested a different overall pattern from that observed in the RTs were found.

General discussion

In the present study, we assessed the engagement of object-based and space-based attention under different task conditions using Ternus motion (Fig. 1) to manipulate the perceived

² When partial eta squared is very small, adjusted partial eta squared can become negative because it adjusts for a positive bias.

object structure of dynamic displays. Specifically, within a given experiment, we measured the effect of selecting information on the basis of object structure irrespective of spatial location, while separately measuring the effect of selecting information on the basis of location irrespective of object structure (Fig. 2). In addition, we used two different tasks, an identification task (Experiments 1 and 3a), which required only the processing of a single stimulus, and a comparison task (Experiments 2 and 3b), which required comparing two stimuli across time. We predicted that the pattern of object- and location-specific effects would differ across these two tasks such that object-specific effects would be larger and more reliable in the comparison task, whereas location-specific effects would depend less on task. These predictions were derived from the object-file framework of Kahneman et al. (1992), a central component of which is an object-based reviewing process that occurs when a stimulus is perceived as corresponding to the same object as another stimulus, and, as a consequence, all of the information that is already associated with that object, when attended, is automatically retrieved. We hypothesized that an object-based reviewing process of this kind would impact a comparison task more than it would impact an identification task, because unlike an identification task, a comparison task depends on the retrieval of previous information (see also Flombaum & Scholl, 2006). Consistent with these predictions, object-specific effects were observed reliably with the comparison task (Experiments 2 and 3b), but not with the identification task (Experiments 1 and 3a). In contrast, location-specific effects did not differ reliably across tasks.

Results from studies using the object-reviewing paradigm that was developed by Kahneman et al. (1992) to measure object-specific effects while controlling for location are consistent with what we found when we explicitly compared tasks. In the original study (Kahneman et al., 1992), letter stimuli were previewed inside of framed shapes, and after removing the letters, the shapes were shown as moving to new locations. A single letter was then presented inside one of the shapes, and the task was to name the letter as quickly as possible. The main finding, which was cited as evidence of the object-based reviewing process described above, was that RTs were reduced when the final letter was the same as the letter that had appeared in that perceived object during preview compared to when it was the letter from the other perceived object. Notice that this was an identification task, and object effects tended to be fairly small in magnitude. Later studies introduced a comparison task in which participants reported whether or not the final stimulus was the same as any of the stimuli in the preview display (Flombaum & Scholl, 2006; Kruschke & Fragassi, 1996; Mitroff & Alvarez, 2007; Mitroff, Scholl, Wynn, 2004, 2005). This comparison task, especially when done with

non-letter stimuli, which are less likely to be nameable, yielded substantially larger object effects (e.g., Fiedler & Moore, 2015; Kruschke & Fragassi, 1996; Hollingworth & Franconeri, 2009; Moore, Stephens, & Hein, 2010). This trend of larger and more robust object effects in a comparison task in contrast to an identification task is similar to the pattern observed in the present study, in which we explicitly compared these two types of task across both near-identical (Experiments 1 and 2) and identical (Experiments 3a and 3b) stimuli.

Some aspects of the location-specific effects that were observed in the present study were unexpected. Specifically, there was a reliable location-specific effect in both Experiment 1 and Experiment 2, which used an identification task and a comparison task, respectively. There was, however, no location-specific effect in either Experiment 3a or Experiment 3b, which used the same set of tasks. First, we note that despite these differences in significance across individual experiments, the overall analysis across experiments indicated that there was a reliable location-specific effect overall and no evidence that it differed across task. Nonetheless, we offer several post hoc possibilities for the differences in reliability of the location-specific effect across the two sets of experiments. First, it is possible that the small (1.6°) separation between same and different locations supported only relatively unreliable engagement of space-based selection mechanisms, resulting in small or non-existent location-specific effects compared to what larger separations might yield. Although the separation between same and different locations was the same in Experiments 3a and 3b as it was in Experiments 1 and 2, all of the effects were reduced in Experiments 3a and 3b, making any small or weak effect more difficult to detect. Second, it is possible that our method of manipulating object structure, ISI, impacted the location-specific effect by making it noisier than it would have been in static displays. Again, for some reason, the impact of ISI (as suggested by the interaction term of the design) appeared to be greater in Experiments 3a and 3b than in Experiments 1 and 2, despite the fact that the timing of trial events was again the same across all experiments. Finally, it is possible that stimulus 1 guided attention differently in Experiments 3a and 3b. The nature of stimulus 1 was the only methodological difference between Experiments 3a and 3b and Experiments 1 and 2. As shown in Fig. 3, stimulus 1 was an enlarged disk in Experiment 1, and it was a gap on an unchanged disk in Experiment 2. In Experiments 3a and 3b, stimulus 1 was an enlarged disk that also had a gap on it. This combination of an enlargement and a task-relevant attribute (i.e., a gap) might have guided attention differently from a single manipulation of stimulus 1. Given that attention tends to be captured by task-

relevant stimuli (e.g., Folk, Remington, & Johnston, 1992) and gaps were relevant stimuli in all experiments, we expected the gap on an enlarged disk used in Experiments 3a and 3b to capture attention at least as well as the simple enlarged disk did in Experiment 1. Nonetheless, it is possible that combining the two attributes somehow disrupted the location-specific effect, which if relatively fragile would then be difficult to detect. We again, however, note that whatever the reason for the lack of significance in Experiments 3a and 3b, the overall analysis indicates that the location-specific effect did not differ significantly across tasks.

A feature of the present study is that it leveraged dynamic displays to not only isolate location-specific and object-specific effects from each other, but also to define the object structure itself, thereby allowing the different effects to be measured from physically identical displays. Dynamic displays are ideal for isolating object-specific and location-specific effects because they allow for a given object to be at multiple locations, and multiple objects to be at a single location. Behrmann and Tipper (1999) used a similar strategy to test whether an impaired ability to detect stimuli (i.e., neglect) in specific parts of the visual field following unilateral parietal lobe damage occurs within a space-based frame of reference, an object-based frame of reference, or both. They found that patients can suffer impairments in detecting stimuli that are on the opposite side of space relative to the lesion as well as on the opposite side of an object relative to the lesion, even when the object has moved out of the supposedly neglected side of space (see also Tipper, Weaver, Jerreat, & Burak, 1994).

More recent studies using dynamic displays have focused on facilitative effects of cues in order to address the question of what it is that attention selects, regions of space or objects that can move around in space. Several of these studies introduced probability manipulations such that targets were more likely to appear at a cued location or, alternatively, on a cued object, and found that these manipulations influenced the relative robustness between effects of space-based and object-based attention (Lamy & Tsal, 2000; Soto & Blanco, 2004; see also Moore & Fulton, 2005). These findings were consistent with the view that endogenous attention (i.e., internally controlled selection that is guided by task-relevant information) can be object-based, as well as space-based, leaving the question open as to whether exogenous attention (i.e., externally controlled “capture” like selection) can be object-based or whether it is exclusively space-based. More recent studies used cues that provided no information about where the upcoming target would appear, and therefore provided no reason to strategically use object structure or spatial location to guide selection (Boi, Vergeer, Ogmen, & Herzog, 2011; Gonen, Hallal, & Ogmen,

2014; Lin, 2013; Ro & Rafal, 1999; Theeuwes, Mathôt, & Grainger, 2013). Although the specific displays and tasks differed across these studies, they all found object-specific cueing effects with uninformative cues, and most also found location-specific effects. One limitation of many of these studies, however, is that different objects were defined relative to each other, in part, on featural differences such as being different shapes (Lamy & Tsal, 2000), different colors (Gonen et al., 2014; Soto & Blanco, 2004; Theeuwes et al., 2013) or different sizes (Boi et al., 2011; cf., Lin, 2013). This leaves open the possibility that feature-based selection contributed to the measured object-specific effects.

Similar to the general goal of previous studies, the present study aimed to dissociate object-specific and location-specific effects of attentional selection. We achieved this goal by using dynamic displays to define the object structure, so that different object structures could be perceived with physically identical stimuli other than their dynamics. Specifically, by varying only the time between the presentation of two displays, a given stimulus could be perceived as part of either the same or a different object as another stimulus. Moreover, we compared patterns of these effects under different processing demands all while under identical stimulus conditions. The overall pattern of results provides strong converging evidence that attention can be guided to objects without regard to their location, as well as to locations without regard to the object in those locations, or both depending on the processing demands imposed by the current task.

In summary, the present study separately measured the effects of space-based and object-based attention by using Ternus motion to manipulate the perceived object structure of stimuli in dynamic displays and assessed these effects across two different tasks. Even with identical stimuli, we found robust object-specific effects with the comparison task, which was expected to benefit from the reviewing process hypothesized as part of the object-file framework (Kahneman, Triesman, & Gibbs, 1992), but not in the identification task which was less likely to be impacted as strongly by reviewing. Together with previous studies, the results confirm that there are separable mechanisms of attention, space-based and object-based, that are recruited functionally depending on the specific processing demands given the current task.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.3758/s13414-020-02201-0>.

Acknowledgement The work reported here was supported in part by NIH grant R21 EY029432 to CMM. We thank Andrew Hollingworth, J. Toby Mordkoff, and Shaun Vecera for helpful discussions and the members of the Iowa Attention and Perception Lab for data collection.

Appendix

Table 2 Mean response times (RTs) and error rates (ERs) with standard errors for the individual cell means of the design shown in Fig. 2

Experiments		Short ISI		Long ISI	
		Same Object, Same Location	Different Object, Different Location	Same Object, Different Location	Different Object, Same Location
1: Identification	RT (ms)	556 ± 17.5	585 ± 17.3	585 ± 20.5	573 ± 18.1
	ER (%)	4.2 ± 0.73	6.1 ± 0.96	3.7 ± 0.61	3.4 ± 0.60
2: Comparison	RT (ms)	794 ± 32.9	875 ± 33.8	803 ± 37.1	818 ± 38.7
	ER (%)	8.4 ± 1.15	16.6 ± 1.45	7.7 ± 1.41	8.4 ± 1.45
3a: Identification	RT (ms)	617 ± 13.3	623 ± 14.0	560 ± 15.1	550 ± 13.7
	ER (%)	6.6 ± 0.85	7.5 ± 0.56	4.4 ± 0.65	4.2 ± 0.61
3b: Comparison	RT (ms)	827 ± 20.3	865 ± 21.4	759 ± 21.7	774 ± 24.9
	ER (%)	9.3 ± 0.99	12.0 ± 1.33	5.8 ± 0.78	6.8 ± 0.80

ISI inter-stimulus interval

References

- Al-Janabi, S., & Greenberg, A. S. (2016). Target-object integration, attention distribution, and object orientation interactively modulate object-based selection. *Attention, Perception, and Psychophysics*, 78, 1968–1984. <https://doi.org/10.3758/s13414-016-1126-3>
- Baylis, G. C., & Driver, J. (1993). Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception & Performance*, 19(3), 451–470. <https://doi.org/10.1037/0096-1523.19.3.451>
- Barnas, A. J., & Greenberg, A. S. (2019). Object-based attention shifts are driven by target location, not object placement. *Visual Cognition*, 27:9–10, 768–791. <https://doi.org/10.1080/13506285.2019.1680587>
- Behrmann, M., & Tipper, S. P. (1999). Attention accesses multiple reference frames: Evidence from visual neglect. *Journal of Experimental Psychology: Human Perception & Performance*, 25(1), 83–101. <https://doi.org/10.1037/0096-1523.25.1.83>
- Boi, M., Vergeer, M., Ogmen, H., & Herzog, M. H. (2011). Nonretinotopic exogenous attention. *Current Biology*, 21(20), 1732–1737. <https://doi.org/10.1016/j.cub.2011.08.059>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436. <https://doi.org/10.1163/156856897X00357>
- Chen, Z. (2012). Object-based attention: A tutorial review. *Attention, Perception, and Psychophysics*, 74(5), 784–802. <https://doi.org/10.3758/s13414-012-0322-z>
- Chen, Z., & Cave, K. R. (2019). When is object-based attention not based on objects? *Journal of Experimental Psychology: Human Perception and Psychophysics*, 45(8), 1062–1082. <https://doi.org/10.1037/xhp0000657>
- Duncan, J. (1984). Selective Attention and the Organization of Visual Information. *Journal of Experimental Psychology: General*, 113(4), 501–517. <https://doi.org/10.1037/0096-3445.113.4.501>
- Egley, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, 123(2), 161–177. <https://doi.org/10.1037/0096-3445.123.2.161>
- Ernst, Z. R., Palmer, J., & Boynton, G. M. (2012). Dividing attention between two transparent motion surfaces results in a failure of selective attention. *Journal of Vision*, 12(12), 6, 1–17. <https://doi.org/10.1167/12.12.6>
- Fiedler, A., & Moore, C. M. (2015). Illumination frame of reference in the object-reviewing paradigm: A case of luminance and lightness. *Journal of Experimental Psychology: Human Perception and Performance*, 41(6), 1709–1717. <https://doi.org/10.1037/xhp0000123>
- Flombaum, J. I., & Scholl, B. J. (2006). A temporal same-object advantage in the tunnel effect: Facilitated change detection for persisting objects. *Journal of Experimental Psychology: Human Perception & Performance*, 32(4), 840–853. <https://doi.org/10.1037/0096-1523.32.4.840>
- 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(3), 1030–1044. <https://doi.org/10.1037/0096-1523.18.4.1030>
- Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics*, 65(7), 999–1010. <https://doi.org/10.3758/BF03194829>
- Gonen, F. F., Hallal, H., & Ogmen, H. (2014). Facilitation by exogenous attention for static and dynamic gestalt groups. *Attention, Perception & Psychophysics*, 76(6), 1709–1720. <https://doi.org/10.3758/s13414-014-0679-2>
- Hein, E., Blaschke, S., & Rolke, B. (2017). The influence of object similarity and orientation on object-based cueing. *Attention, Perception, and Psychophysics*, 79(1), 63–77. <https://doi.org/10.3758/s13414-016-1229-x>
- Hollingworth, A., & Franconeri, S. L. (2009). Object correspondence across brief occlusion is established on the basis of both spatiotemporal and surface features. *Cognition*, 113(2), 150–166. <https://doi.org/10.1016/j.cognition.2009.08.004>
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24(2), 175–219. [https://doi.org/10.1016/0010-0285\(92\)90007-0](https://doi.org/10.1016/0010-0285(92)90007-0)
- Kruschke, J. K., & Fragassi, M. M. (1996). The perception of causality: Feature binding in interacting objects. In *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society* (pp. 441–446). Hillsdale, NJ: Erlbaum.
- 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(4), 1387–1400. <https://doi.org/10.1037/0096-1523.26.4.1387>

- Lin, Z. (2013). Object-centered representations support flexible exogenous visual attention across translation and reflection. *Cognition*, 129(2), 221–231. <https://doi.org/10.1016/j.cognition.2013.07.001>
- Mitroff, S. R., & Alvarez, G. A. (2007). Space and time, not surface features, underlie object persistence. *Psychonomic Bulletin & Review*, 14(6), 1199–1204. <https://doi.org/10.3758/BF03193113>
- Mitroff, S. R., Scholl, B. J., & Wynn, K. (2004). Divide and conquer: How object files adapt when a persisting object splits into two. *Psychological Science*, 15(6), 420–425. <https://doi.org/10.1111/j.0956-7976.2004.00695.x>
- Mitroff, S. R., Scholl, B. J., & Wynn, K. (2005). The relationship between object files and conscious perception. *Cognition*, 96(1), 67–92. <https://doi.org/10.1016/j.cognition.2004.03.008>
- Moore, C. M., & Fulton, C. (2005). The spread of attention to hidden portions of occluded surfaces. *Psychonomic Bulletin & Review*, 12(2), 301–306. <https://doi.org/10.3758/BF03196376>
- Moore, C. M., Stephens, T., & Hein, E. (2010). Features, as well as space and time, guide object persistence. *Psychonomic Bulletin & Review*, 17(5), 731–736. <https://doi.org/10.3758/PBR.17.5.731>
- Mordkoff, J. T. (2019). A simple method for removing bias from a popular measure of standardized effect size: Adjusted partial eta squared ($adj \hat{\eta}_p^2$). *Advances in Methods and Practices in Psychological Science*, 2(3), 228–232. <https://doi.org/10.1177/2515245919855053>
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437–442. <https://doi.org/10.1163/156856897X00366>
- Pilz, K. S., Roggeveen, A. B., Creighton, S. E., Bennett, P. J., & Sekuler, A. B. (2012). How prevalent is object-based attention? *PLoS ONE*, 7(2): e30693. <https://doi.org/10.1371/journal.pone.0030693>
- Pikler, J. (1917). Sinnesphysiologische Untersuchungen. Leipzig: J. A. Barth.
- Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978) Attended and unattended processing modes: The role of set for spatial location. In H. L. Pick, & I. J. Saltzman (Eds), *Modes of perceiving and processing information*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32(1), 3–25. <https://doi.org/10.1080/00335558008248231>
- Reppas, I., Schmidt, W. C., & Leek, E. C. (2012). Successes and failures in producing attentional object-based cueing effects. *Attention, Perception, & Psychophysics*, 74(1), 43–69. <https://doi.org/10.3758/s13414-011-0211-x>
- Ro, T., & Rafal, R. D. (1999). Components of reflexive visual orienting to moving objects. *Perception and Psychophysics*, 61(5), 826–836. <https://doi.org/10.3758/BF03206900>
- Soto, D., & Blanco, M. J. (2004). Spatial attention and object-based attention: A comparison within a single task. *Vision Research*, 44(1), 69–81. <https://doi.org/10.1016/j.visres.2003.08.013>
- Ternus, J. (1926). Experimentelle untersuchungen über phänomenale Identität. *Psychologische Forschung (Psychological Research)*, 7(1), 81–136. <https://doi.org/10.1007/BF02424350>
- Theeuwes, J., Mathôt, S., & Grainger, J. (2013). Exogenous object-centered attention. *Attention, Perception & Psychophysics*, 75(5), 812–818. <https://doi.org/10.3758/s13414-013-0459-4>
- Tipper, S. P., Weaver, B., Jerreat, L. M., & Burak, A. L. (1994). Object-based and environment-based inhibition of return of visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, 20(3), 478–499. <https://doi.org/10.1037/0096-1523.20.3.478>
- Valdes-Sosa, M., Cobo, A., & Pinilla, T. (1998). Transparent motion and object-based attention. *Cognition*, 66(2), B13–B23. [https://doi.org/10.1016/S0010-0277\(98\)00012-2](https://doi.org/10.1016/S0010-0277(98)00012-2)
- Vecera, S. P. (1994). Grouped locations and object-based attention: comment on Egly, Driver, and Rafal (1994). *Journal of Experimental Psychology: General*, 123(3), 316–320. <https://doi.org/10.1037/0096-3445.123.3.316>
- Watt, R. J. (1988). *Visual processing: Computational, psychophysical, and cognitive research*. Hillsdale, NJ: Erlbaum.

Open Practices Statement

The data and materials for this study are available upon request to qingzi-zheng@uiowa.edu. This study was not preregistered.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.