Check for updates

Object-based attention generalizes to multisurface objects

Gennady Erlikhman[1] · Taissa Lytchenko[2] · Nathan H. Heller[3] · Marvin R. Maechler[3] · Gideon P. Caplovitz[2]

Abstract

When a part of an object is cued, targets presented in other locations on the same object are detected more rapidly and accurately than are targets on other objects. Often in object-based attention experiments, cues and targets appear not only on the same object but also on the same surface. In four psychophysical experiments, we examined whether the "object" of attentional selection was the entire object or one of its surfaces. In Experiment 1, facilitation effects were found for targets on uncued, adjacent surfaces on the same object, even when the cued and uncued surfaces were oriented differently in depth. This suggests that the "object-based" benefits of attention are not restricted to individual surfaces. Experiments 2a and 2b examined the interaction of perceptual grouping and object-based attention. In both experiments, cuing benefits extended across objects when the surfaces of those objects could be grouped, but the effects were not as strong as in Experiment 1, where the surfaces belonged to the same object. The cuing effect was strengthened in Experiment 3 by connecting the cued and target surfaces with an intermediate surface, making them appear to all belong to the same object. Together, the experiments suggest that the objects of attention do not necessarily map onto discrete physical objects defined by bounded surfaces. Instead, attentional selection can be allocated to perceptual groups of surfaces and objects in the same way as it can to a location or to groups of features that define a single object.

Keywords Object-based attention · Cue facilitation · Attention spreading

Due to capacity limits of our visual system, we can process only a small portion of the immense amount of information that reaches us at any given moment. Visual attention is the process that selects subsets of the immense amount of visual information that reaches us at any given moment for subsequent processing. For decades, visual attention has been a focus of modern cognitive scientists who have worked to empirically define the scope, function, and limits of this selective process. The study of attention has led to a classification scheme based largely on what is being attended. For example, attention can be deployed to a particular location

in the visual scene (i.e., a "spotlight"; Cave & Bichot, 1999; Eriksen & Eriksen, 1974; Eriksen & Yeh, 1985; Posner, 1980; Posner, Snyder, & Davidson, 1980; Treisman, 1982), to specific features across a scene such as color, orientation, or motion (i.e., "feature-based" attention; for reviews see Carrasco, 2011; Maunsell & Treue, 2006; Theeuwes, 2013), to surfaces or depth planes (He & Nakayama, 1992, 1995), to objects (i.e., "object-based" attention; Chen, 2012; Duncan, 1984; Egly, Driver, & Rafal, 1994; Scholl, 2001; Treisman, Kahneman, & Burkell, 1983), or to object parts (Vecera, Behrmann, & Filapek, 2001; Vecera, Behrmann, & McGoldrick, 2000). In this paper, we focus on these latter surface-based and object-based aspects of visual attention with a specific goal of determining how configural properties of objects and their surfaces interact with the allocation of visual attention. The effects of object-based attention can be observed when endogenous or exogenous cues are used to guide attention toward an object, in which case multiple locations on or features of the object can be attended at once (Duncan, 1984; He & Nakayama, 1992; Rafal, Egly & Driver, 1994; see Chen, 2012, for a review). For example, in an overlapping objects paradigm, two outline objects are shown in the same spatial location (Duncan, 1984). Participants are better at reporting two different properties of the same object compared with a single property from

✉ Gennady Erlikhman
gennady@ucla.edu

[1] Department of Psychology, University of California, Los Angeles, 2349 Franz Hall, 405 Hilgard Ave, Los Angeles, CA, USA

[2] Department of Psychology, University of Nevada, Reno, Reno, NV, USA

[3] Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH, USA

two objects (see also Kahneman & Henick, 1981). By attending to one feature of an object, all other features of the object were also processed. These effects cannot be explained by the simple spatial allocation of attention because both objects occupy the same region of space. Another popular framework for revealing object-based attention effects is the "two-rectangle paradigm" (see Fig. 1), in which two vertically or horizontally arranged rectangles are shown side by side (Egly et al., 1994). One of the ends of the rectangles is cued, and then a target appears in either that cued end (valid condition), at the opposite end of the same rectangle (invalid, same object), or at the end of the other rectangle (invalid, other object). Reaction times are fastest for targets that appear in the cued location and slowest for targets appearing in the uncued location on the other object. When the target appears on the opposite end of the same object in which the cue appeared, reaction times are as fast as when the target appears in the same location as the cue (valid condition). The distance from the cue to the targets on the same or other objects are matched, so again, the cue facilitation effects for targets appearing on the same object cannot be explained by spatial allocation of attention. However, the "objects" in a majority of studies investigating object-based attention, and particularly those that employ variants of the two-rectangle paradigm are defined by single two-dimensional surfaces (i.e., a rectangle) residing in a single depth plane. In contrast objects in the real-world are volumetric and are typically composed of multiple surfaces, each of which likely vary in depth. It is therefore unclear from these experiments whether the reported effects of attention truly represent the selection of an entire object or whether the effects reflect the selection of a single surface lying within a single depth plane.

Some evidence in favor of attention selecting an entire object comes from studies using visual search paradigms. Images of cubes pop out among similar objects whose interiors do not allow for a 3-D interpretation, suggesting that some types of junctions that are indicative of surface

orientation in depth may be processed preattentively (Enns & Rensink, 1990, 1991; although see Brown, Weisstein, & May, 1992; Zhang, Huang, Yigit-Elliott, & Rosenholtz, 2015). For facilitation effects, when rectangles in the two-rectangle display are covered by an occluder, target detection at one end of the rectangle can still be facilitated by a cue at the other end, even though there is an intervening surface separating the two (Albrecht, List, & Robertson, 2008; He & Nakayama, 1992, 1995; Law & Abrams, 2002; Moore, Yantis, & Vaughan, 1998; Reppa & Leek, 2006). In these displays, each end of the rectangle is a bounded surface in the image. Additional processing is needed to determine that they are amodally completed behind the occluding surface. However, after completion, the object in these displays is still a two-dimensional surface lying in a single depth plane. Another line of evidence for volumetric objects being the objects of attention comes from the inhibition of return (IOR) paradigm. Similar to cue facilitation, a peripheral location is cued, and, after a pause, a target appears either in the cued or an uncued location. If the interval between the cue and target exceeds approximately 300 ms, then responses to targets presented in the cued, valid location may be *slower* than to uncued locations (Klein, 1988; Maylor & Hockey, 1985; Posner & Cohen, 1984; Posner, Rafal, Choate, & Vaughan, 1985; see Klein, 2000, for a review).

It has been shown that the effects of IOR can extend to multiple surfaces defining a volumetric object; however, the results have been equivocal (Bourke, Partridge, & Pollux, 2006; Gibson & Egeth, 1994; Umiltà, Castiello, Fontana, & Vestri, 1995). For example, Gibson and Egeth (1994) found evidence for IOR for cues and targets on the same surface of a "brick" (parallelepiped) after controlling for location-based attentional effects. However, IOR effects were weakest for cues and targets presented on different surfaces of the same object. In another study, the objects were pairs of rectangles arranged in depth and presented stereoscopically (Bourke
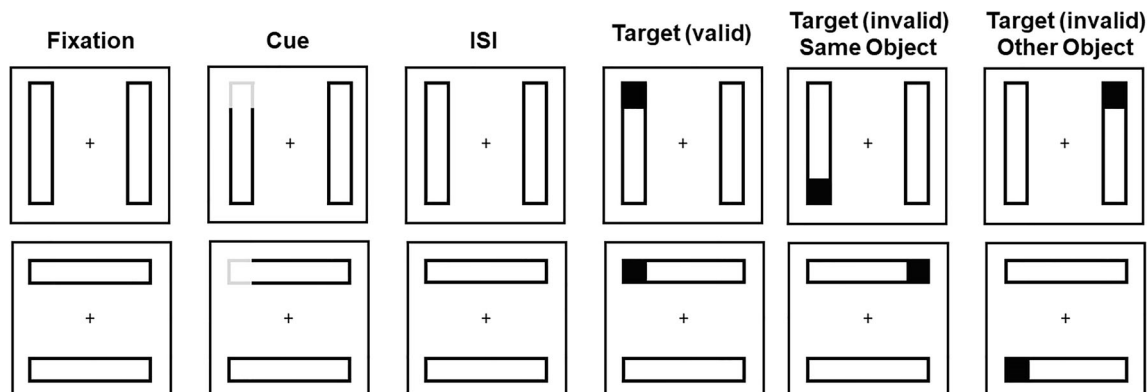

Fig. 1 "Two-rectangle" paradigm from Egly et al. (1994). Rectangles can be arranged either vertically (top row) or horizontally (bottom row). A cue highlights the end of one rectangle. A target (black square) then appears

either in the same location (cue is valid), or in a different location (cue is invalid). The cue-to-target distance is the same for when the target appears at the end of the same object or on the other object

et al., 2006; see also Theeuwes & Pratt, 2003). In one condition, the corners of each pair of rectangles were connected in depth to make a Necker-cube-like object. Reaction times were the same to targets presented in the same location as the cue as they were to targets presented on a nearer or farther surface, but only when the two surfaces were connected. A similar effect has been observed for 2-D surfaces that can be made to appear to belong to the same (2-D) object by the addition of contextual cues (Müller & O'Grady, 2009).

Taken together, the evidence for the selection of volumetric objects by attention is equivocal: Sometimes, IOR effects are comparable for when cues and targets appear on the same surface as when they appear on different surfaces of the same object, but not always. Perhaps these effects depend on whether a single rotating object is used, as in Gibson and Egeth (1994), or whether there are several stationary objects (Bourke et al., 2006), or perhaps volumetric-object-based attentional effects only occur for stereoscopically defined surfaces. In the following experiments, we set out to examine whether cue facilitation effects could be found for different surfaces of the same object. In Experiment 1, we modified the two-rectangle paradigm to use cubes instead of 2-D rectangles, either vertically or horizontally arranged (see Fig. 2). If facilitation effects are limited to individual surfaces, then reaction times to targets presented on adjacent surfaces of a volumetric solid should be similar to those of equidistant targets on the surface of another object. In contrast, if attention to a location on one surface facilitated detection of targets anywhere on the entire object—if the object of attention was the entire object and not just the individual surface—then reaction times to targets on the same object should be faster than those on other objects. In Experiments 2a and 2b, we sought to modulate facilitation effects by adding cues to the display that would break the link between the surfaces (i.e., would make them appear to belong to different objects). In Experiment 3, surface grouping was restored by addition of yet another cue. Across the experiments, we find that attentional facilitation can extend across multiple surfaces of the same object as well as across groups of noncontiguous surfaces.

Experiment 1

In most studies on object-based attention using the Egly et al. (1994) paradigm, objects are horizontally or vertically arranged rectangles. The goal of Experiment 1 was to test whether cueing a location on one surface would confer facilitation effects to targets on adjacent, bounded surfaces if the two surfaces were perceived to belong to the same object. The objects in this experiment were vertically or horizontally arranged cubes. A location on one of the surfaces of one of the cubes was cued and then a target appeared in the same location (valid), a different location on the same surface (invalid, same surface), on an adjacent surface still on the same cube (invalid, other surface), or on the other object (invalid, other object). These different conditions are illustrated in Fig. 3.

Although we describe the cubes here as a collection of bounded surfaces that are distinct from each other and from the background on which they appear, it is not clear what features of the display constitute "objects" for the attentional selection system. For example, each surface in a cube is a bounded region. Closure of a region can sometimes strengthen objecthood cues (Marino & Scholl, 2005), perhaps increasing the likelihood that each surface would be treated as a separate object preventing facilitation effects from applying to targets presented on adjacent surfaces of the same object. At the same time, closure may not be necessary to produce facilitation effects (Avrahami, 1999; Ben-Shahar, Scholl, & Zucker,
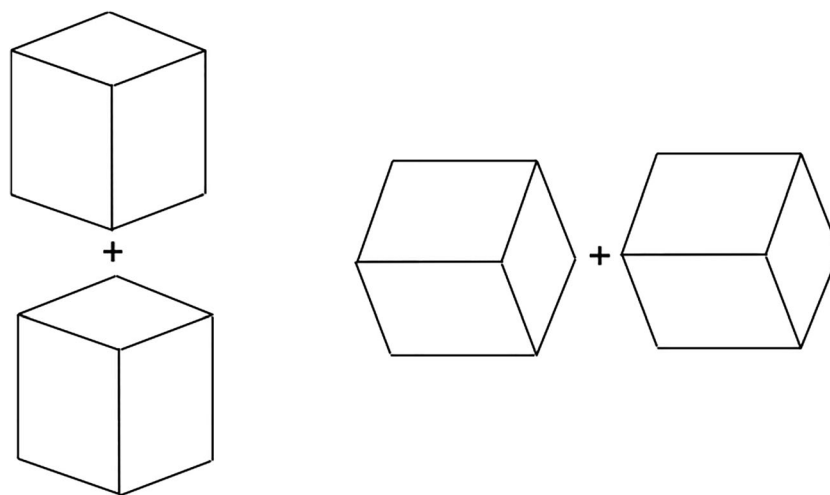

Fig. 2 Example stimuli used in Experiment 1. Instead of two vertically or horizontally arranged 2-D surfaces, the objects were 3-D cubes that were either vertically (left) or horizontally (right) arranged
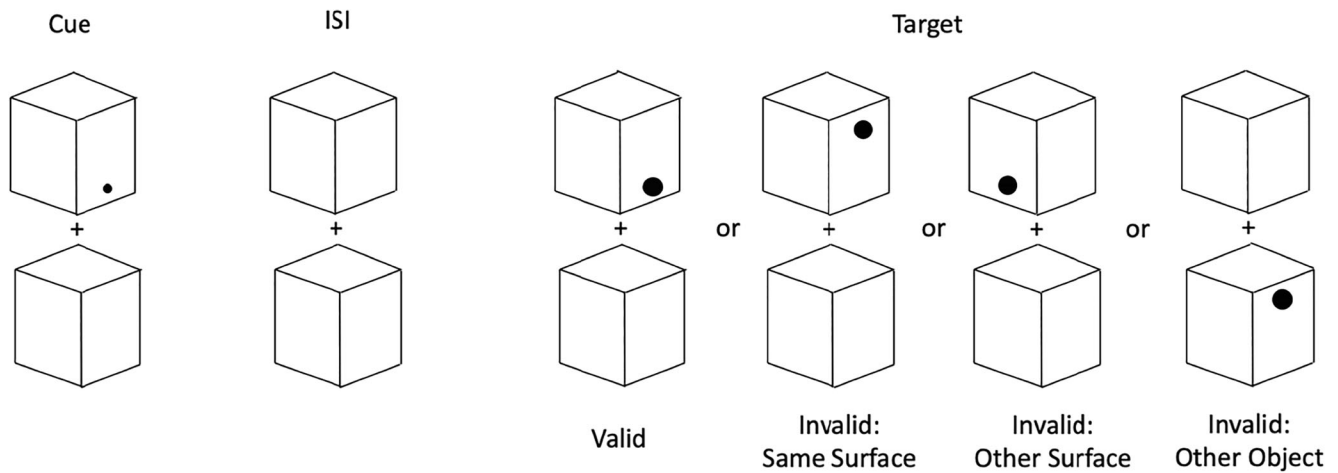
Fig. 3 Examples of the sequence of stimuli in Experiment 1 when the objects were vertically arranged. The small black dot represents the cue and the large black dot the target. The cue could appear either at the top or bottom of either forward-facing surface of each cube. Relative to the cue position, the target could appear in either the same position (valid) or one of three invalid positions. The horizontal arrangement of the objects was created by rotating each display clockwise by 90°. ISI = interstimulus interval

2007; Marrara & Moore, 2003, Experiment 5) and so may matter little in this case. Likewise, interior part boundaries sometimes block facilitation or inhibition of return effects (Hecht & Vecera, 2007), and sometimes they do not (Leek, Reppa, & Tipper, 2003; Possin, Filoteo, Song, & Salmon, 2009; Reppa & Leek, 2003, 2006). The edges between surfaces may therefore either help or hurt facilitation. Finally, the fact that the surfaces are oriented in different depth planes may also prevent facilitation effects (Atchley & Kramer, 2001; He & Nakayama, 1992; Reppa, Fougnie, & Schmidt, 2010). We consider these factors in greater detail in the General Discussion.

Method

Participants

Participants were 23 undergraduate students (16 male, seven female, mean age = 25.4 years) at the University of Nevada, Reno, who participated for course credit. This was a similar number of participants to other recent studies using similar cueing paradigms (e.g., Chen & Cave, 2019; Gibson & Egeth, 1994). In all experiments, the goal was to collect approximately 20 participants. All participants had normal or corrected-to-normal vision and were naïve to the purposes of the experiment.

Apparatus

Stimuli were created and shown using Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) in MATLAB. Stimuli were displayed on a 27-inch iMac (late 2013) with a 2.7 GhZ 12-Core Intel Xeon E5 processor and AMD FirePro D700 graphics card. The refresh rate of the monitor was set to 60 Hz. The screen resolution was 2,560 × 1,440 pixels. Viewing

distance was set to 67 cm and participants' heads were stabilized at that distance with a chin rest.

Stimuli

Displays consisted of two cubes arranged one above the other (vertical) or side by side (horizontal). In the following, we describe how the vertical arrangement was constructed; horizontal displays followed the same steps, but with a further rotation of 90° about the z-axis. Each cube was seen edge on and was tilted toward the observer (rotated about the x-axis by 10° (see Fig. 1). The stimuli were shown using parallel projection. Unrotated, the length of each edge would have been 2.38°. After rotation, the sides of the cube projected to parallelograms on the screen, with a width of 1.69° and height of 2.35°. The length of the shortest edge was 1.71°. The cube centers were separated by 3.48°. The size of the gap between the two cubes was 0.55° between the closest points on each cube. The cubes were shown on a black background. The edges of each cube were white and the surfaces black (same color as the background). A gray fixation (0.20° × 0.20°) appeared between the two cubes.

Cues were white circles (0.1° diameter) and could appear near the top or bottom of each frontal surface of each cube, for a total of eight possible cue locations. The locations were 1.69° apart (0.84° above and below the center of each surface). For example, the left surface of the upper cube in the vertical arrangement had two cue positions: one near the bottom and one near the top. The right surface of that cube had two similar locations, with the four locations forming a square with sides of 1.69°. The left surface of the bottom cube also had two target locations near its top and bottom, with the top location being 1.69° away from the bottom location of the left surface of the cube above.

Targets were larger white circles (0.4° diameter) that could appear in the same eight locations as the cues. Cues were either

valid (target appeared in cued location) or invalid (target appeared in an uncued location). There were three invalid cue types: (1) targets could appear on the same surface, but in a different location from the cue. (2) Targets could appear on an adjacent, near surface of the same object (always in the corresponding position, so that, for example, if the cue appeared in the lower portion of a surface, the target would also appear in the lower portion of the adjacent surface). (3) Targets could appear on the closest location of the other object (for example, if a cue appeared at the bottom of left surface of the upper object, the target would appear at the top of the left surface of the bottom object).

Figure 3 shows the four cue–target conditions (one valid, three invalid) for cues that appeared in the bottom right surface of the upper cube in the vertical configuration. No other target locations were tested for that cue location because the cue–target distances would be greater for those other locations. In principle, this means that the target was three times as likely to appear on the same object as on the other object and could be a potential confound. However, the key comparison of interest was between targets appearing on the same surface as the cue (standard object-based attention effect) versus targets appearing on an adjacent surface. Targets were twice as likely to appear on the same surface as the cue (valid and one of the invalid conditions) as on the adjacent surface. If there was any anticipatory effect, it would therefore favor faster reaction times to invalid targets in the same-surface condition compared with the near-surface condition. As discussed in a later section, this was not the case.

In the valid condition, cues appeared in each of the valid locations 30 times over the course of the experiment for the vertical arrangement of the objects and an equal number of times for the horizontal arrangement. Each invalid condition occurred 24 times in each arrangement, three times per cue location. In total, there were 480 valid trials and 48 invalid trials of each of the three kinds described above (144 total). An additional 128 catch trials were included in which no target appeared after the cue. For these trials, cue location was counterbalanced to occur equally often in all eight possible positions. Overall, cues were valid on 63.83% of trials and invalid on 19.15% of trials. The remaining 17.02% of trials were catch trials in which no target appeared following the cue.

Procedure

Participants were given verbal instructions about the nature of the task. Additional written instructions were provided at the beginning of the experiment, followed by 20 practice trials. Each trial started with a blank black screen for 500 ms, after which the fixation cross and the cubes were shown. After 1 s, the cue appeared for 100 ms, after which it disappeared. After another 200 ± 50 ms, the target appeared in one of four locations until a response was made with a key press or until 2 s

had passed. On catch trials, no target was shown, and the trial ended 2 s after the cue disappeared. If a response was made within 150 ms of the target's appearance, a beep was played and a message appeared on the screen asking participants to try not to anticipate the target and to only press a key when they detected the target. A similar message appeared if no response was made after 2 s, warning participants to try to respond more quickly, unless it was a catch trial, in which no target appeared and no response should have been made. If a response was made on a catch trial (false alarm), a beep sound was also played, alerting the participants that they made an error. Every 100 trials, a message appeared on the screen, asking participants to take a break and to resume to the experiment with a key press.

Results

Reaction time (RT) data were grouped into four categories: valid trials in which the cue and target appeared in the same location, invalid trials where the cue and target appeared on the same surface but in different locations, invalid trials where the cue and target appeared on different surfaces of the same object, and invalid trials where the cue and target appeared on different objects. On a subset of the latter type of trial, sometimes cues appeared in the farthest locations on an object (e.g., for vertically arranged cubes, in the upper portion of the left surface of the top cube). For the target to then appear on the other object, the cue-to-target distance would have been greater than if the cue had appeared at the bottom location on the same surface. These trials were excluded from the analysis so that only trials used were those in which the cue-to-target distance was the same. RTs of less than 150 ms or greater than 1,500 ms were also excluded. Across all participants, an average of 3.38% of trials were excluded for falling outside this range. Average accuracy on the catch trials was 89.43%. Average RTs on valid trials were subtracted from the three invalid trial types to create a RT cost score as in Egly et al. (1994). Data were matched across target position and stimulus orientation (e.g., RTs for valid trials where the stimuli were arranged vertically and the target appeared in the lower-right corner of the upper cube were subtracted from all invalid cue conditions where the stimuli were arranged in the same way and the target appeared in that position). A within-subject, repeated-measures 2 (orientation) × 3 (invalid cue condition) × 4 (target locations) ANOVA (Greenhouse–Geisser corrected) found only a main effect of cue condition, $F(1.39, 29.21) = 12.51$, $p < .001$, $\eta_p^2 = 0.37$, no effect of orientation, $F(1, 21) = 2.97$, $p = .10$, $\eta_p^2 = 0.12$, or target position, $F(2.01, 42.20) = 1.04$, $p = .36$, $\eta_p^2 = 0.05$, and no two-way or three-way interactions (all ps > .05). As a result, for all subsequent analyses, RTs were averaged across stimulus orientations (vertical and horizontal) and target locations.

Reaction-time data for all experiments are shown in Table 1, with differences between the valid and invalid conditions (RT cost, see Fig. 4) shown in parentheses. On average, RTs were fastest in the valid condition (288 ms), a little slower when the cue and target appeared on the same surface (294 ms), slowest when they appeared on different objects (327 ms), and intermediate when they appeared on different surfaces of the same object (304 ms). A score of zero indicated no cost in switching attention from the cued location to the target location, whereas positive scores reflected the additional time required to make the switch.

After averaging across target location and orientation, planned pairwise *t* tests were used to compare each of the three invalid cue types. The cueing effect was larger in the other-object condition than in either the other-surface, $t(22) = 6.03$, $p < .001$, Cohen's $d = 1.26$, or same-surface conditions, $t(22) = 3.96$, $p < .001$, Cohen's $d = 0.83$. There was no difference in cueing effect depending on whether the target appeared on the same or a different surface of the same object, $t(22) = 1.04$, $p = .307$, Cohen's $d = 0.22$.

Discussion

The results demonstrate that object-based cueing facilitation effects extend to adjacent surfaces that are perceived to belong the same object. RTs to targets presented on the same surface as a cued location were no different from those of targets presented on an adjacent surface. However, RTs to targets on surfaces that belonged to a different object were longer. Because the cue-to-target distances were the same across all conditions, the additional switching cost to uncued locations on other objects was likely due to having to shift attention across objects. Critically, when the target appeared on a different surface of the same object, there was little to no switching cost, despite the fact that the other surface was oriented differently in depth. suggesting that attention was initially allocated to the entire object as a whole and not just to one of its surfaces.

A recent study using rectangular stimuli found that facilitation effects were stronger for horizontally arranged bars (Chen & Cave, 2019). Although we did not find a statistically significant effect of orientation, the observed pattern of results was consistent with these findings with the RT cost in the
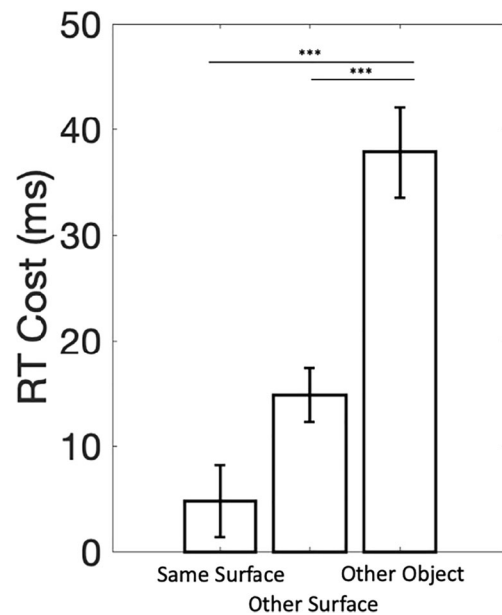

Fig. 4 Average RT cost: difference in reaction time between each invalid cueing condition and the valid cue condition from Experiment 1. Larger values mean RTs were slower in that condition compared with the valid cue condition. Negative values mean that the RT for those conditions was faster than in the valid cue condition. Error bars are standard errors. Each point represents the data for an individual participant. ***$p < .001$

other-surface condition being less for vertically aligned stimuli (10.6 s) compared with horizontally aligned stimuli (16.3 s). Note that when the cubes were vertically aligned, in the other-surface condition, the cue and target were horizontally arranged (one on each forward-facing surface of one cube); when the cubes were horizontally aligned, the cue and target were vertically arranged.

Facilitation effects have previously been found for overlapping objects where the cue and target locations are separated by another, occluding surface (Law & Abrams, 2002). In this case also, the cue and target are appearing on two different, bounded regions of the display. However, locations appearing at two ends of an amodally completed object are still appearing on a single surface, even if that surface is not continuous in the image. In contrast, in the other-surface condition in this experiment, the location of the target was on a distinct surface. It is possible that observers perceived the displays as flat, two-dimensional parallelograms and diamonds, but subjectively they do not appear as such. Furthermore, the objects

Table 1 Reaction time for all cue–target conditions from all experiments, in ms. Cueing effect (invalid − valid) shown in parentheses.

	Valid	Invalid (same surface)	Invalid (other/near surface)	Invalid (other/far object)
Exp. 1	289	294 (5)	304 (15)	327 (38)
Exp. 2a	296	307 (11)	313 (17)	323 (26)
Exp. 2b	376	381 (5)	398 (22)	398 (22)
Exp. 3	309	320 (11)	325 (16)	339 (30)

were present on the screen for the duration of a trial, which lasted several seconds, so it was unlikely that observers did not have the time to recover and perceive their structure.

However, there could have been other explanations for the facilitation and switching RT cost effects. One possibility is that eccentricity was not controlled across locations relative to the fixation point even though all cue-to-target distances were identical. In the 2-D version of these displays, distances between the fixation and all cued locations on the objects are the same. In the vertical arrangement in the current experiment, locations on the bottom part of the surfaces above fixation and locations on the top part of the surfaces below fixation were closer than the other locations on those surfaces. However, if fixation-target eccentricity was a factor, then RT to targets appearing on the same surface but a different location (which was farther from the fixation point) should have been slower than to targets appearing either on the near surface (closer to fixation) or on the other object (also closer to fixation). Since RTs were actually *faster* (lower RT cost) for the same-surface condition, the effects cannot be explained by fixation-target eccentricity. Another possibility is that facilitation in the other-surface condition may have arisen because of some other form of grouping between the surfaces other than the fact that they belong to the same object such as their symmetric arrangement, or the fact that they were closer to each other than the surfaces of the other object. In Experiments 2a and 2b, we explore some of these possibilities.

Experiment 2a

Experiment 2a was similar to Experiment 1, except the surfaces were moved slightly apart and the upper portion of the

cube was replaced with other edges that made the surfaces seem to be sides of two separate prisms (see Fig. 5, left panel). If facilitation occurs for surfaces belonging to the same object, then adding such extra cues that indicate that they belong to different objects should block the facilitation and increase RT cost. That is, RT to targets on nearby surfaces that belong to a different object should be as slow as those to targets on surfaces that belong to a distant object. In contrast, if the effects are due to a more general grouping of the surfaces (e.g., symmetry or proximity), then the facilitation effects should persist.

Method

Participants

Participants were 27 undergraduate students (10 male, 17 female, mean age = 24.7 years) at the University of Nevada, Reno. All had normal or corrected-to-normal vision and were naïve to the purposes of the study.

Stimulus and apparatus

The apparatus was the same as for Experiment 1. The objects in this experiment were pairs of prisms (see Fig. 5, left). The surfaces on which the cues and targets appeared were the same dimensions as in Experiment 1, but were shifted apart by 0.1°. In pilot work, we tested a version of the display without the gap in order to keep the stimuli as close as possible to those in the previous experiment, but participants reported that the prisms appeared to be one book-like object. The edges that formed the upper portion of each cube were shortened and another edge was added to make each object into two prisms.
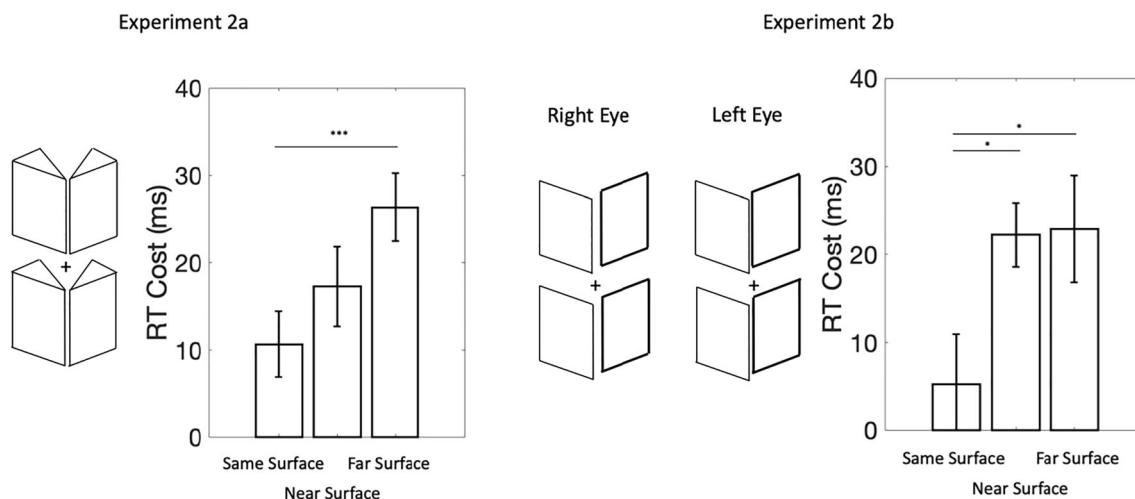

Fig. 5 Stimuli used in Experiments 2a (left) and 2b (right). The surfaces for Experiment 2b are stereo pairs that can be cross-fused. The surface with the thicker boundary appears behind the one with the thinner boundary. In the experiment, surfaces appeared in different colors and were viewed with anaglyph glasses (see text for details). Stimuli were also presented in horizontal arrangement (display rotated clockwise by 90°). Targets could appear in one of eight locations as in Experiment 1 (see Fig. 3). Next to the stimuli are RT costs, as in Fig. 4. Error bars are standard errors. Points represent individual participants. ***$p < .001$. **$p < .01$. *$p < .05$

As in Experiment 1, both vertical and horizontal arrangements of the objects were used. The properties of the cues and targets and the distance between them remained unchanged. The same types of valid and invalid trials were used. In general, apart from the underlying stimulus configuration, no other aspects of the experimental design were altered. The principle difference due to the new display configuration was that targets in the other surface invalid cue condition now appeared on a surface of a different object. We refer to surfaces on prisms adjacent to the one on which the cue appeared as the near surface and those that are part of the other pair of prisms as the far-surface (corresponding to the other surface and other-object conditions in Experiment 1, respectively).

Results and discussion

The same analysis procedure was applied to the RT data from this experiment as in Experiment 1. Across all participants, an average of 5.32% of trials were excluded for having either too short or too long RTs. Average accuracy on the catch trials was 86.57%. Data are shown in Fig. 5 and in Table 1. As in Experiment 1, RTs were fastest in the valid cue condition (295 ms), slightly slower in the same-surface condition (307 ms), slowest in the far-surface condition (323 ms), and intermediate for the near-surface condition (313 ms). As in Experiment 1, RTs from the valid cue condition were subtracted from the invalid cue conditions to form RT cost scores. A 2 (orientation) × 3 (invalid cue condition) × 4 (target position) repeated-measures ANOVA (Greenhouse–Geisser corrected) revealed an overall effect of cue condition, $F(1.66, 39.89) = 4.44$, $p = .020$, $\eta_\mathrm{p}^2 = 0.16$. There was no effect of orientation, $F(1, 19) = 0.15$, $p = .701$, $\eta_\mathrm{p}^2 = 0.006$, or target position, $F(1.69, 40.60) = 0.64$, $p = .509$, $\eta_\mathrm{p}^2 = 0.03$, and no two-way or three-way interactions (all ps > .05). Data were therefore collapsed across orientation and target position. Paired comparisons between the three invalid cue conditions showed a difference in RT cost between the same-surface and far-surface conditions, $t(26) = 3.14$, $p = .004$, Cohen's $d = 0.60$, but not between same-surface and near-surface, $t(26) = 1.14$, $p = .263$, Cohen's $d = 0.22$, nor between near-surface and far-surface conditions, $t(26) = 1.81$, $p = .082$, Cohen's $d = 0.35$.

By changing the scene information in the display to make each pair of surfaces appear independent, switching costs were increased for targets on those surfaces compared with Experiment 1. However, we predicted that RTs for the near-surface condition would be similar to those of the far-surface condition: that any two prisms would be treated as two different objects. Instead, RTs were intermediate in the near-surface condition, suggesting that two prisms in a pair were not as distinct from each other as they were from the other pair of prisms. It may be that surfaces in each pair were grouped in some way, for example, due to symmetry

or proximity of the surfaces themselves. For example, in the near-surface condition, the cued and target surfaces are closer together than the cued and target surfaces in the far-surface condition. Perhaps attentional facilitation extends to groups of surfaces or objects and basic Gestalt grouping principles such as similarity or proximity determine grouping strength.

In the design and conceptualization of this experiment, we tacitly assumed that the two prisms would be perceived as two separate objects because no parts of them were touching. Although each display was repeated many times over the course of the experiment and were visible for several seconds on each trial, it is possible that some observers did not notice the small gap between them. If the displays were sometimes seen as containing two objects and sometimes four, then this could potentially account for the intermediate results in the near-surface condition. In order to test this, we conducted a control experiment in which the exact same stimuli were shown (without the cues and targets), and observers were asked to simply count how many objects they saw on the screen. No additional information was given about what counted as an object, and no examples were shown. We report the details of the design in the Supplementary Materials. Of the eight participants, all but one consistently reported the vertical configuration as consisting of two objects (see Supplementary Fig. S1). All reported the horizontal configuration as consisting of two objects. It is not clear why one participant thought there was only one object in the vertical configuration. However, a majority of the time, the prisms were seen as two objects, and so the intermediate results in the near-surface condition were unlikely to be due to ambiguity in the total number of objects in the displays. For all other experiments, the reported number of objects matched what was expected.

Experiment 2b

Additional cues in Experiment 2a that made the pairs of surfaces appear as separate objects only partially slowed reaction times relative to the same-surface condition. In Experiment 2b, additional cues were added to further break the grouping of the surfaces. The small gap was retained between the surfaces, but the additional edges that made them appear as parts of prisms were removed. The stimuli were presented in stereo so that each surface in a pair appeared on a different depth plane. Surfaces were also shifted apart vertically, and one was made to have thicker edges to further distinguish between the two and to aid fusion. With these additional cues, if grouping was sufficiently disrupted, then the switching cost in the near-surface condition should be more similar to that of the far-surface than to the same-surface condition.

Method

Participants

Participants were 20 undergraduate students (one male, 19 female, mean age = 21.75 years) at the University of Nevada, Reno. All had normal or corrected-to-normal vision and were naïve to the purposes of the study.

Stimulus and procedure

The apparatus was the same as in the other experiments. The stimuli were created by starting with the shapes used in Experiment 2a and removing all edges except those that defined each pair of surfaces. On each trial, one of each pair of surfaces was randomly shifted upward or downward relative to the other by 0.025°. The same surface from each pair was shifted (e.g., in the vertical configuration, the right surface from each pair would be shifted). The edges of the right surface in the vertical configuration and those of the bottom surface in the horizontal configuration were made thicker than the edges of the surface the same pair (0.04° vs. 0.02°). Both the displacement and the manipulation of the thickness were done to facilitate stereo fusion. Each pair of surfaces was drawn twice on the screen, once in red (RGB values [0.4, 0, 0] on 0-1 scale) and once in blue (RGB values [0, 0.2, 0.7]). On each trial, one of the surfaces in each pair was shifted to the right or left by 0.05° in one eye and by the same amount in the opposite direction in the other eye. This offset made one of the surfaces in each pair appear closer or farther than the other when the two displays were fused. Which surface was shifted and in which direction was randomized across trials. However, across pairs of surfaces, the same surface was shifted in the same direction. For example, in the vertical configuration, if the left surface of the upper pair appeared shifted away from the observer in depth, then the left surface of the bottom pair was also shifted in the same direction. We therefore refer to the case where a target appears on an adjacent surface that is shifted in depth as the near/different-Depth condition, and the case where a target appears on a distant surface that is on the same depth plane as the far/same-depth condition. In all other respects, the procedure was the same as the other experiments. The only feature that was different was the objects on which the cues and targets appeared. Participants wore red–blue anaglyph glasses for the duration of the experiment.

Results and discussion

The same analysis procedure was used as in the previous experiments. Across all participants, an average of 0.65% of trials were excluded for having either too short or too long RTs. Average accuracy on the catch trials was 97.31%. RTs are shown in Table 1 and in Fig. 5 (right). It is not clear why overall accuracy was higher in this experiment than in the previous ones. RTs for all conditions were slower by ~100 ms compared with those in Experiments 1 and 2a. Perhaps it took some time for the stereo to "kick in"; however, the display was on the screen for some time before the cue appeared, and it was not necessary to see the displays in stereo to do the task. It may also be the case that this group of participants was simply more careful in their responses.

As in Experiment 1, RTs were fastest in the valid cue condition (376 ms), slower in the same-surface condition (381 ms), and slowest in the near-surface (398 ms) and far-surface conditions (398 ms). Valid RTs were subtracted from the invalid RTs. A 2 (orientation) × 3 (invalid cue condition) × 4 (target position) repeated-measures ANOVA revealed a main effect of cue condition, $F(1.87, 35.57) = 4.33$, $p = .022$, $\eta_p^2 = 0.185$. There was no effect of orientation, $F(1, 19) = 1.18$, $p = .292$, $\eta_p^2 = 0.058$, or target position, $F(2.72, 51.67) = 1.04$, $p = .379$, $\eta_p^2 = 0.052$. There were no significant two-way interactions (all ps > .05). There was a significant three-way interaction between orientation, cue condition, and target position, $F(5.02, 95.37) = 2.57$, $p = .031$, $\eta_p^2 = 0.12$. However, because there were no a priori hypotheses about which contrasts to test and because all possible pairwise comparisons of conditions would have resulted in hundreds of tests, data were collapsed across orientation and target position as before. Pairwise t tests were used to compare each of the three invalid cue types. The cueing effect was weaker in the same-surface condition than in either the near-surface, $t(19) = 2.52$, $p = .021$, Cohen's $d = 0.56$, or far-surface conditions, $t(19) = 2.33$, $p = .031$, Cohen's $d = 0.52$. There was no difference in cueing effect depending on whether the target appeared on the near or far surface, $t(19) = 0.068$, $p = .947$, Cohen's $d = 0.02$. The additional cues of depth, edge thickness, and misalignment made the RTs for invalid near-surface and other-surface conditions more similar than those in Experiment 2a.

Experiment 3

In Experiments 2a and 2b, cue facilitation effects were reduced by separating the surfaces, moving them to different depth planes and adding other cues that suggested that they belonged to separate surfaces. Here, we sought to regroup the two surfaces into a single object while still keeping them separated in 2-D. We further sought to test facilitation effects when the surfaces were nonadjacent on the 3-D object. The prism stimuli from Experiment 2a were modified so that each prism in a pair was connected (see Fig. 5). The two surfaces appeared to belong to the same object, but shared no mutual edges as in Experiment 1. If facilitation effects are object based, then they should apply even to nonadjacent surfaces.

Method

Participants

Participants were 23 undergraduate students (four male, 19 female, mean age = 21.65 years) at the University of Nevada, Reno. All had normal or corrected-to-normal vision and were naïve to the purposes of the study.

Stimulus and procedure

The apparatus was the same as in the other experiments. Stimuli were created by beginning with the prisms in Experiment 2a and connecting the tops of each prism with a lines to make the forward-facing surfaces of the prisms appear to belong to a single object (see Fig. 6, left). A small line was added between the two prisms to be the "bottom" of this new object. All other aspects of the experiment were the same.

Results and discussion

Average accuracy on catch trials was 85.06% and an average of 7.54% of trials were excluded. This large proportion of excluded trials prompted us to check individual participant performance, and we found that two participants responded correctly to fewer than 50% of the catch trials, suggesting that they were simply hitting the response keys without waiting for the target to appear and were not doing the task properly. We therefore excluded these two participants in all subsequent analyses. When included, the data are qualitatively similar. No participants in any of the other experiments performed as
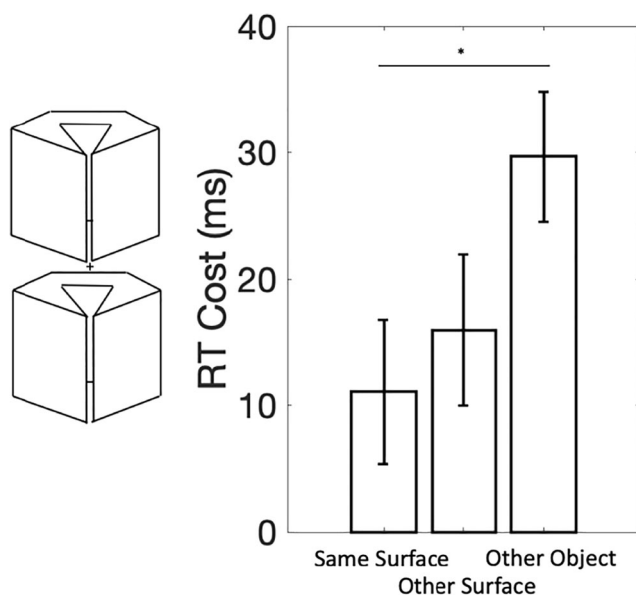

Fig. 6 Left: Stimulus in Experiment 3. Additional edges were added to the tops of the prisms from Experiment 2a to make the surfaces appear to belong to the same object

poorly on the catch trials. For the remaining 21 participants, average accuracy on catch trials was 90.85% and errors on catch trials was 3.43%.

Reaction times are shown in Table 1 and RT costs in Fig. 6 (right). As in Experiment 1, RTs were fastest in the valid cue condition (309 ms), slower in the same-surface condition (320 ms), slowest in the other-object condition (339 ms), and intermediate for the other-surface condition (325 ms). There was no main effect of orientation, $F(1, 19) = 1.02$, $p = .326$, $\eta_p^2 = 0.05$, nor of target position, $F(2.59, 49.15) = 1.62$, $p = .201$, $\eta_p^2 = 0.08$) The effect of cue condition was not significant, $F(1.91, 36.36) = 3.19$, $p = .055$, $\eta_p^2 = 0.14$, likely due to the increased variability in the same-surface condition (see Fig. 6). No two-way interactions were significant (all $ps > .05$). There was again a three-way interaction between orientation, cue condition, and target position, $F(3.51, 66.76) = 2.73$, $p = .043$, $\eta_p^2 = 0.13$. After averaging across orientation and position, pairwise comparisons were performed between the three cue conditions. Similar to Experiment 1, there was no difference between RT costs when the targets appeared on the same-surface or on the other-surface, $t(20) = 0.438$, $p = .666$, Cohen's $d = 0.10$, or between the other-surface and other-object conditions, $t(20) = 1.99$, $p = .060$, Cohen's $d = 0.43$. RT costs were greater for the other-object compared to the same-surface condition, $t(20) = 2.82$, $p = .011$, Cohen's $d = 0.62$.

By manipulating contextual factors (i.e., addition of connecting edges between the surfaces to make them appear to belong to the same object), the pattern of facilitation effects reverted to that of Experiment 1 in which the surfaces were two sides of a cube. An important difference is that a gap has now been introduced between the two surfaces so that they are no longer adjacent. Facilitation effects therefore appear to extend to nonadjacent surfaces as long as they are part of the same object. Qualitatively, however, the pattern of results matches more closely those of Experiment 2a, in which the magnitude of the facilitation effect for the other-surface condition is in between that of the same-surface and the other-object conditions. This may indicate that, despite the gap between the two surfaces, the prisms in Experiment 2a were still treated as one object. Alternatively, it may be the case that facilitation only partially extends to distant parts or surfaces of multipart objects. We consider these possibilities in the next section.

General discussion

Behavioral paradigms that reveal effects of object-based attention commonly employ "objects" that are defined by a single 2-D surface lying on a single depth plane (Reppa, Schmidt, & Leek, 2012). As such, it remains unknown as to whether or not object-based attention is in fact object based, or whether it is

actually surface based (i.e., restricted to the cued surface). The results of the experiments described above conclusively demonstrate that object-based attention can indeed be object based, facilitating the detection of targets located on noncued surfaces of an object. This was revealed in the results of Experiment 1, in which cue-facilitation effects were observed for targets presented on surfaces adjacent to the cued surface even though that surface was oriented differently in depth. In this case, attention can be said to extend around the corners of an object to adjacent surfaces. This result was replicated and extended in Experiment 3, in which cue-facilitation effects were observed for targets on surfaces other than the one on which a cue appeared, even though the two surfaces did not share an edge in common.

In addition to these primary observations, we conducted Experiments 2a and 2b to determine if breaking up the volumetric objects used in Experiment 1 into distinct and separate objects would make the facilitation effect go away. In both of these experiments, the cross-surface facilitation observed in Experiment 1 was reduced. This again suggests that object-based attention is indeed object based. However, the pattern of results observed in Experiments 2a and 2b are also suggestive of an alternative and unexpected hypothesis that object-based attention may in some instances extend to the surfaces of uncued objects. In Experiment 2a, this hypothesis is supported by the fact that targets presented on the uncued proximal surface are on average detected more rapidly that on the distal surface. In Experiment 2b, a similar pattern is observed where targets presented on the uncued surface lying in the same depth plane are on average detected faster than those lying on a surface in different depth plane. In our viewing of these stimuli, we note that the proximal surface in Experiment 2a tends to perceptually group with the cued object, and in Experiment 2b the surfaces group according to their depth planes. It may therefore be the case that object-based attention may not necessarily be object based after all, but instead may extend to multiple objects that are perceptually grouped together. Or it may indicate that what counts as an object for the attentional system is itself a continuous rather than a discrete notion. We consider the implications of these graded effects below. We note however, that the results of Experiments 2a and 2b are merely suggestive of this hypothesis. Moreover, we did not explicitly measure observers' experiences of perceptual grouping. That said, we believe the hypothesis suggested by these results is noteworthy enough to highlight and warrants further investigation.

We note that for both the near-surface and far-surface conditions in Experiment 2, the RT cost is smaller than those observed in Experiment 1 and Experiment 3. One speculative hypothesis for this is that in both of these conditions we are observing some grouping-based benefit. For example, on a given trial in Experiment 2b, an observer may group two other the surfaces according to depth or according to 2-D proximity.

To test hypothesis, we ran a second control experiment (see Supplementary Materials) where we asked observers to report their subjective grouping response. Consistent with this hypothesis, we found a bimodal pattern of responses for the stimuli of Experiment 2b—half of the observers grouped according to proximity and half grouped according to depth. These conclusions are tempered by the fact that this experiment was done in a separate group of participants.

Alternatively, the fact that RTs in the far-surface condition had a smaller facilitation effect (22 ms) compared with the other-object condition in Experiments 1 (39 ms) and far-surface condition in Experiments 2a (27 ms) might indicate that the two surface that were on the same depth plane were grouped more strongly than (i.e., treated more as one unit) than two surfaces that were closest in the image plane (near surface). This is consistent with previous work showing cueing facilitation along a depth plane, but not across depth planes (He & Nakayama, 1992, 1995; Reppa et al., 2010).

What distinguishes one part of the scene as belonging to an object versus another part not belonging to the object or belonging to a different object? Unlike the selection of spatial location, features like color or orientation, or even depth planes (none of which require object individuation to be detected), objects require additional processing. One possibility is that objects for the visual system at this level are any enclosed or bounded region. L shapes arranged into the corners of a square pop out amongst other arrangements of Ls (Donnelly, Humphreys, & Riddoch, 1991; Pomerantz & Pristach, 1989), closed contours are more easily detected in noise than are open curves (Kovacs & Julesz, 1993), and closure strengthens object-based cueing effects in the two-rectangle paradigm (Marino & Scholl, 2005). However, closure cues do not always indicate that an object is present. Consider, for example, an outline drawing of a cube, as in Fig. 1. Some of the edges are boundaries between the cube and background, whereas others are interior edges that form the boundaries between two surfaces, both of which are part of the same object. Closure may not even be explicitly necessary as rectangles made of dots also produce the same facilitation effects (Marrara & Moore, 2003). Nor are closure cues sufficient for individuating objects. Object-based facilitation effects occur in the two-rectangle paradigm even when the rectangles are defined by illusory contours that are not closed in the image (Moore et al., 1998) and also for occluded rectangles when the two ends in which the cue and target appear are separated by an intervening surface in the image (Behrmann, Zemel, & Mozer, 1998; Chen & Cave, 2008; Haimson & Behrmann, 2001; Law & Abrams, 2002; Leek et al., 2003; Moore et al., 1998; Reppa & Leek, 2003, 2006; for a review, see Reppa et al., 2012). Contextual cues may also have an effect on the perceived grouping of the rectangles and subsequently on whether they are or are not seen as two separate objects or parts of one object. For example, if the rectangles

are shown as parts of a single surface seen through two rectangular holes in a nearer surface, then the facilitation effects occur for all targets (Albrecht et al., 2008).

Taken together, these studies and the current findings suggest that the targets of attention are not only bounded objects as determined by closure, but may be *perceptual groups*. Sometimes, this grouping is a result of closure, in which case bounded surfaces may correspond to individual objects. In other cases, other grouping cues like proximity or surface similarity (e.g., color) may determine what constitutes an "object" (Chen, 2012; Marino & Scholl, 2005; Scholl, 2001). For example, in the two-rectangle paradigm, if the ends of each rectangle are colored differently, the cue facilitation effects are reduced or eliminated (Hecht & Vecera, 2007; see also Chen, 1998). Similarly, if the objects are formed from several differently colored segments, the effects can also be reduced (Watson & Kramer, 1999; although see Matsukura & Vecera, 2006). Indeed, attention can be selectively allocated to object parts instead of to the whole object (Hollingworth, Maxcey-Richard, & Vecera, 2012; Vecera et al., 2001; Vecera et al., 2000). Likewise, in the current experiments, contextual cues may make the two surfaces seem to more or less strongly belong together, and this in turn determines attentional facilitation strength.

Such cues may also aid in decomposing an object into parts. For example, in the two-rectangle paradigm, if the two ends of each rectangle are made to be different colors, the facilitation effect is reduced (Hecht & Vecera, 2007). Adding notches at the color change boundary to facilitate segmentation causes the effect to reappear. In general, results have been mixed in part-based allocation of attention, with some studies finding that spatial gaps or changes in surface features between a cue and an object disrupt facilitation (Hollingworth et al., 2012; Vecera et al., 2001; Vecera et al., 2000; Watson & Kramer, 1999), while other studies have found facilitation effects for adjacent parts or across differently textured areas of a single surface (Albrecht et al., 2008; Behrmann et al., 1998; Matsukura & Vecera, 2006; Moore et al., 1998). However, it may be that just as certain cues strengthen or reduce the perceived grouping of parts into a single object, similarly contextual scene cues may influence the grouping of surfaces into volumetric wholes.

Our results are consistent with the notion that the objects of attention are perceptual groups (Chen, 2012; Scholl, 2001). The strength of object-based effects, whether they are faciliatory or inhibitory, may depend the grouping strength of contextual cues with respect to the different surfaces in a display. Proximity of surfaces seems to be an important feature, although in Experiment 3 we found that facilitation effects can extend to nonadjacent surfaces. The nature of these grouping cues in the context of object-based attention remains underexplored. It would be interesting, for example, if grouping strength as determined in other paradigms such as visual search predicted facilitation or IOR effect magnitude in these experiments. This may suggest a common mechanism linking multiple perceptual organization processes.

Open practices statement The data and all materials for all experiments is available via the Open Science Framework.

Compliance with ethical standards

Conflict of interest None.

References

Albrecht, A. R., List, A., & Robertson, L. C. (2008). Attentional selection and the representation of holes and objects. *Journal of Vision*, *8*(13), 1–8. doi:https://doi.org/10.1167/8.13.8

Atchley, P., & Kramer, A. F. (2001). Object and space-based attentional selection in three-dimensional space. *Visual Cognition*, *8*(1), 1–32. doi:https://doi.org/10.1080/13506280042000009

Avrahami, J. (1999). Object of attention, objects of perception. *Perception & Psychophysics*, *61*, 1604–1612. doi:https://doi.org/10.3758/bf03213121

Behrmann, M., Zemel, R. S., & Mozer, M. C. (1998). Object-based attention and occlusion: Evidence from normal participants and a computational model. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(4), 1011–1036. doi:https://doi.org/10.1037//0096-1523.24.4.1011

Ben-Shahar, O., Scholl, B. J., & Zucker, S. W. (2007). Attention, segregation, and textons: Bridging the gap between object-based attention and texton-based segregation. *Vision Research*, *47*(6), 845–860. doi:https://doi.org/10.1016/j.visres.2006.10.019

Bourke, P. A., Partridge, H., & Pollux, P. M. (2006). Additive effects of inhibiting attention to objects and locations in three-dimensional displays. *Visual Cognition*, *13*(5), 643–654. doi:https://doi.org/10.1080/13506280544000309

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433–436.

Brown, J. M., Weisstein, N., & May, J. G. (1992). Visual search for simple volumetric shapes. *Perception & Psychophysics*, *51*(1), 40–48.

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*(13), 1484–1525. doi:https://doi.org/10.1016/j.visres.2011.04.012

Cave, K. R., & Bichot, N. P. (1999). Visuospatial attention: Beyond a spotlight model. *Psychonomic Bulletin & Review*, *6*(2), 204–223. doi:https://doi.org/10.3758/bf03212327

Chen, Z. (1998). Switching attention within and between objects: The role of subjective organization. *Canadian Journal of Experimental Psychology, 52*, 7–16. doi:https://doi.org/10.1037/h0087274

Chen, Z. (2012). Object-based attention: A tutorial review. *Attention, Perception, & Psychophysics*, *74*(5), 784–802. doi:https://doi.org/10.3758/s13414-012-0322-z

Chen, Z., & Cave, K. R. (2008). Object-based attention with endogenous cueing and positional certainty. *Perception & Psychophysics, 70*(8), 1435–1443. doi:https://doi.org/10.3758/PP.70.8.1435

Chen, Z., & Cave, K. R. (2019). When is object-based attention not based on objects? *Journal of Experimental Psychology: Human Perception and Performance*, *45*(8), 1062–1082. doi:https://doi.org/10.1037/xhp0000657

Donnelly, N., Humphreys, G. W., & Riddoch, M. J. (1991). Parallel computation of primitive shape descriptions. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(2), 561–570. doi:https://doi.org/10.1037//0096-1523.17.2.561

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, *113*(4), 501–517. doi:https://doi.org/10.1037//0096-3445.113.4.501

Egly, 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. doi:https://doi.org/10.1037//0096-3445.123.2.161

Enns, J. T., & Rensink, R. A. (1990). Sensitivity to three-dimensional orientation in visual search. *Psychological Science*, *1*(5), 323–326.

Enns, J. T., & Rensink, R. A. (1991). Preattentive recovery of three-dimensional orientation from line drawings. *Psychological Review*, *98*(3), 335–351. doi:https://doi.org/10.1037/0033-295x.98.3.335

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & psychophysics*, *16*(1), 143–149. doi:https://doi.org/10.3758/BF03203267

Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(5), 583–597. doi:https://doi.org/10.1037//0096-1523.11.5.583

Gibson, B. S., & Egeth, H. (1994). Inhibition of return to object-based and environment-based locations. *Perception & Psychophysics*, *55*(3), 323–339. doi:https://doi.org/10.3758/bf03207603

Haimson, C., & Behrmann, M. (2001). Cued visual attention does not distinguish between occluded and occluding objects. *Psychonomic Bulletin & Review, 8*(3), 496–503. doi:https://doi.org/10.3758/bf03196184

He, Z. J., & Nakayama, K. (1992). Surfaces versus features in visual search. *Nature*, *359*(6392), 231–233. doi:https://doi.org/10.1038/359231a0

He, Z. J., & Nakayama, K. (1995). Visual attention to surfaces in three-dimensional space. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(24), 11155–11159. doi:https://doi.org/10.1073/pnas.92.24.11155

Hecht, L. N., & Vecera, S. P. (2007). Attentional selection of complex objects: Joint effects of surface uniformity and part structure. *Psychonomic Bulletin & Review, 14*(6), 1205–1211. doi:https://doi.org/10.3758/bf03193114

Hollingworth, A., Maxcey-Richard, A. M., & Vecera, S. P. (2012). The spatial distribution of attention within and across objects. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(1), 135–151. doi:https://doi.org/10.1037/a0024463

Kahneman, D., & Henick, A. (1981). Perceptual organization and attention. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization.* Hillsdale, NJ: Erlbaum.

Klein, R. (1988). Inhibitory tagging system facilitates visual search. *Nature, 334*, 430–431. doi:https://doi.org/10.1038/334430a0

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences, 4*, 138–147. doi:https://doi.org/10.1016/S1364-6613(00)01452-2

Kovacs, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one: Effect of closure in figure-ground segmentation. *Proceedings of the National Academy of Sciences of the United States of America, 90*(16), 7495–7497. doi:https://doi.org/10.1073/pnas.90.16.7495

Law, M. B., & Abrams, R. A. (2002). Object-based selection within and beyond the focus of spatial attention. *Perception & Psychophysics*, *64*(7), 1017–1027. doi:https://doi.org/10.3758/bf03194753

Leek, E. C., Reppa, I., & Tipper, S. P. (2003). Inhibition of return for objects and locations in static displays. *Perception & Psychophysics, 65*(3), 388–395. doi:https://doi.org/10.3758/bf03194570

Marino, A. C., & Scholl, B. J. (2005). The role of closure in defining the "objects" of object-based attention. *Perception & Psychophysics*, *67*(7), 1140–1149. doi:https://doi.org/10.3758/bf03193547

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*(7), 1103–1109. doi:https://doi.org/10.3758/bf03194837

Matsukura, M., & Vecera, S. P. (2006). The return of object-based attention: Selection of multiple-region objects. *Perception & Psychophysics, 68*(7), 1163–1175. doi:https://doi.org/10.3758/bf03193718

Maunsell, J. H., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, *29*(6), 317–322. doi:https://doi.org/10.1016/j.tins.2006.04.001

Maylor, E. A., & Hockey, R. (1985). Inhibitory component of externally controlled covert orienting in visual space. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 777–787. doi:https://doi.org/10.1037//0096-1523.11.6.777

Moore, C. M., Yantis, S., & Vaughan, B. (1998). Object-based visual selection: Evidence from perceptual completion. *Psychological science*, *9*(2), 104–110. doi:https://doi.org/10.1111/1467-9280.00019

Müller, H. J., & O'Grady, R. B. (2009). Object-based selection operating on a spatial representation made salient by dimensional segmentation mechanisms: A re-investigation of Egly and Homa (1984). *Psychological Research*, *73*(2), 271–286. doi:https://doi.org/10.1007/s00426-008-0213-z

Pelli, D. G. (1997). The video toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Pomerantz, J. R., & Pristach, E. A. (1989). Emergent features, attention, and perceptual glue in visual form perception. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(4), 635. doi:https://doi.org/10.1037/0096-1523.15.4.635

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. doi:https://doi.org/10.1080/17470218.2014.937446

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 531–556). Hillsdale, NJ: Erlbaum.

Posner, M. I., Rafal, R.D., Choate, L. S., &Vaughan, J. (1985). Inhibition of return: Neural basis and function. *Cognitive Neuropsychology, 2*, 211–228. doi:https://doi.org/10.1080/02643298508252866

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*(2), 160–174. doi:https://doi.org/10.1037/0096-3445.109.2.160

Possin, K. L., Filoteo, J. V., Song, D. D., & Salmon, D. P. (2009). Space-based but not object-based inhibition of return is impaired in Parkinson's disease. *Neuropsychologia, 47*(7), 1694–1700. doi:https://doi.org/10.1016/j.neuropsychologia.2009.02.006

Reppa, I., Fougnie, D., & Schmidt, W. C. (2010). How does attention spread across objects oriented in depth?. *Attention, Perception, & Psychophysics*, *72*(4), 912–925. doi:https://doi.org/10.3758/APP.72.4.912

Reppa, I., & Leek, E. C. (2003). The modulation of inhibition of return by object internal structure: Implications for theories of object-based attentional selection. *Psychonomic Bulletin & Review*, *10*(2), 493–502. doi:https://doi.org/10.3758/BF03196512

Reppa, I., & Leek, E. C. (2006). Structure-based modulation of inhibition of return is triggered by object-internal but not occluding shape features. *Quarterly Journal of Experimental Psychology, 59*(11), 1857-1866. doi:https://doi.org/10.1080/17470210600872113

Reppa, 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. doi:https://doi.org/10.3758/s13414-011-0211-x

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, *80*(1/2), 1–46. doi:https://doi.org/10.1016/S0010-0277(00)00152-9

Theeuwes, J. (2013). Feature-based attention: It is all bottom-up priming. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *368*(1628), 20130055. doi:https://doi.org/10.1098/rstb.2013.0055

Theeuwes, J., & Pratt, J. (2003). Inhibition of return spreads across 3-D space. *Psychonomic Bulletin & Review, 10*(3), 616–620. doi:https://doi.org/10.3758/BF03196523

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance*, *8*(2), 194–214. doi:https://doi.org/10.1037//0096-1523.8.2.194

Treisman, A., Kahneman, D., & Burkell, J. (1983). Perceptual objects and the cost of filtering. *Perception & Psychophysics*, *33*(6), 527–532. doi:https://doi.org/10.3758/BF03202934

Umiltà, C., Castiello, U., Fontana, M., & Vestri, A. (1995). Object-centred orienting of attention. *Visual Cognition*, *2*(2/3), 165–181. doi:https://doi.org/10.1080/13506289508401729

Vecera, S. P., Behrmann, M., & Filapek, J. C. (2001). Attending to the parts of a single object: Part-based selection limitations. *Perception & Psychophysics*, *63*(2), 308–321. doi:https://doi.org/10.3758/bf03194471

Vecera, S. P., Behrmann, M., & McGoldrick, J. (2000). Selective attention to the parts of an object. *Psychonomic Bulletin & Review*, *7*(2), 301–308. doi:https://doi.org/10.3758/BF03212985

Watson, S. E., & Kramer, A. F. (1999). Object-based visual selective attention and perceptual organization. *Perception & Psychophysics, 61*(1), 31–49. doi:https://doi.org/10.3758/bf03211947

Zhang, X., Huang, J., Yigit-Elliott, S., & Rosenholtz, R. (2015). Cube search, revisited. *Journal of Vision*, *15*(3), 9–9. doi:https://doi.org/10.1167/15.3.9