Space-, object-, and feature-based attention interact to organize visual scenes

Dwight J. Kravitz and
Laboratory of Brain and Cognition, National Institute of Mental Health, Bethesda, MD 20892, USA

Marlene Behrmann
Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, USA

Dwight J. Kravitz: kravitzd@mail.nih.gov

Abstract

Biased-competition accounts of attentional processing propose that attention arises from distributed interactions within and among different types of perceptual representations (e.g., spatial, featural, and object-based). Although considerable research has examined the facilitation in processing afforded by attending selectively to spatial locations, or to features, or to objects, surprisingly little research has addressed a key prediction of the biased-competition account: that attending to any stimulus should give rise to simultaneous interactions across all the types of perceptual representations encompassed by that stimulus. Here we show that, when an object in a visual display is cued, space-, feature-, and object-based forms of attention interact to enhance processing of that object and to create a scene-wide pattern of attentional facilitation. These results provide evidence to support the biased-competition framework and suggest that attention might be thought of as a mechanism by which multiple, disparate bottom-up, and even top-down, visual perceptual representations are coordinated and preferentially enhanced.

Keywords

Neural attentional mechanisms; Object-based attention; Space-based attention

As we look around the world, we are confronted with visual scenes containing many disparate items, and even more numerous possible relationships between these items. Attention is generally thought of as a mechanism by which a subset of the overwhelming visual input is selected for further processing. While there has been much interest in characterizing the psychological and neural mechanisms underlying attentional selection, most of the studies have considered selection primarily on the basis of only one type of perceptual representation or dimension—for example, selection by spatial location (space-based attention [SBA]; Posner, 1980; Posner, Cohen, & Rafal, 1982), or by particular features of the input (e.g., color; feature-based attention [FBA]; see, e.g., Baylis & Driver, 1992; Driver & Baylis, 1989; Duncan & Nimmo-Smith, 1996; Harms & Bundesen, 1983; Kramer & Jacobson, 1991), or by object membership (object-based attention [OBA]; Duncan, 1984; Egly, Driver, & Rafal, 1994; Kramer, Weber, & Watson, 1997; Vecera & Farah, 1994). Additionally, in recent years, investigations have explored the neural correlates of attentional selection, but again, these studies have focused predominantly on selection from one domain of representation—be it, for example, spatial-based (Gandhi,

Correspondence to: Dwight J. Kravitz, kravitzd@mail.nih.gov.

Heeger, & Boynton, 1999; Tootell et al., 1998), feature-based (Martinez-Trujillo, Medendorp, Wang, & Crawford, 2004; Maunsell & Treue, 2006; Rossi & Paradiso, 1995; Saenz, Buracas, & Boynton, 2002; Schoenfeld et al., 2003), or object-based (Martinez et al., 2006; Martinez, Ramanathan, Foxe, Javitt, & Hillyard, 2007; Serences, Schwarzbach, Courtney, Golay, & Yantis, 2004).

Although substantial evidence now exists for the attentional enhancement and facilitation for processing those items selected for preferential processing, there has been little consideration of how these independent perceptual representations are coordinated in the service of coherent perception. One popular account of selection, "biased competition" (Deco & Zihl, 2004; Desimone, 1998; Desimone & Duncan, 1995; Duncan, 2006; Reynolds, Chelazzi, & Desimone, 1999), has considered this issue explicitly and has posited that selection arises from competitive and cooperative interactions *within* and *between* different perceptual representations or dimensions. Thus, the search for a red bar among red and green bars would be facilitated if attention were cued just to the color red (feature-based; green items suppressed by competition) or to the location of the target bar (space-based; nontarget locations suppressed by competition). Moreover, there would be an interaction between the featural and spacebased representations, given that these two dimensions are shared by the target stimulus. By virtue of this apparent interactivity between different perceptual representations shared by a target, selection based on a single dimension (e.g., color) should have direct consequences for the processing of other shared stimulus dimensions (e.g., spatial position). The result of this integration across many distinct competitive systems is a flexible, integrated, and behaviorally relevant description of the input (Duncan, 2006), and similar principles may be at work in the deployment of overt attention or eye movements (Zelinsky, 2008). Surprisingly, rather few investigations have explored the interaction of different forms of attention. Some studies have noted that selection can take place on the basis of spatial and object representations simultaneously (e.g., Egly et al., 1994; Shomstein & Behrmann, 2006) or on the basis of featural and spatial position simultaneously (Harms & Bundesen, 1983; Kim & Cave, 2001), but few have directly demonstrated the influence of one form of selection on another, as would be predicted by a biased-competition account.

One recent pertinent study that has explored this attention-based interactivity has reported that when a location within an object is cued, facilitation accrues not only to other locations within the same object (OBA), but also to information in the surrounding space (SBA) (Kravitz & Behrmann, 2008). In these experiments, one end of a single object (such as a rectangle or a barbell) was cued (see Fig. 1a and b). The ensuing target could appear in the cued, valid location, or at the opposite end of the object, in the invalid location. Additionally, targets could appear in one of two possible locations outside the bounds of the cued object, both of which were equidistant from the cue, but one of which was nearer to the center of mass of the cued object (near-object) and one of which was farther from the center of mass of the cued object (far-object). Unsurprisingly, detection or discrimination was better when the target appeared in the valid versus invalid locations, reflecting the object-based advantage. The more intriguing and novel finding was that performance for the outer target was facilitated in the near-object position relative to the far-object position, reflecting the influence of the object representation on the spatial topography of the surround of the object (Fig. 1a). These findings suggest that an interaction takes place between OBA and SBA, wherein the facilitated locations within the cued object give rise to spatial gradients such that facilitation in the surround is contingent on distance from their average location (the center of mass of the object; Fig. 1b).

In the present study, we perform a more extensive evaluation of the interactivity predicted by biased competition, and we show not only that attending to perceptual features, such as

color (Exp. 1) and shape (Exp. 2), gives rise to FBA, but that this FBA also interacts with SBA and OBA to form graded patterns of facilitation across the entire visual scene (Fig. 1c and d). Moreover, the interactivity is not restricted to similarity as defined by the perceptual properties of the input, because in Experiment 3 we demonstrate that similarity at the level of semantic category yields similar graded patterns of facilitation across the entire visual scene (see also Schmidt & Zelinsky, 2009; Yang & Zelinsky, 2009, for effects of semantic category on attention).

In all of these experiments, participants viewed displays consisting of two objects, which either were identical or differed in color, shape, or semantic category (Fig. 2). When one of the two objects was cued, reaction times (RTs) to targets appearing within that object were faster than RTs to equidistant targets in the uncued object, replicating the well-known OBA effect (Kramer & Jacobson, 1991; Martinez-Trujillo & Treue, 2004; Saenz et al., 2002). However, the strength of OBA was modulated by FBA, with targets in the uncued object, which shared features with the cued object, eliciting faster RTs than targets in an uncued object whose features differed from those of the cued object (see also Shomstein & Behrmann, 2008). Moreover, the additional facilitation afforded to the identical objects gave rise to faster processing of targets that appeared in the surround (or ground) of both the cued and uncued objects. In sum, as shown in Fig. 1c and d, a purely spatial cue initially evoked OBA. The magnitude of this OBA depended on the featural/categorical similarity between the cued and uncued objects (FBA), and furthermore, this joint OBA–FBA effect in turn determined the distribution of spatial attention to the surround (SBA). We argue, then, that visual selection and the ensuing attentional enhancement ultimately emerges from the interactions within and among the perceptual representations of spatial position, features, and objects in a visual scene, and that performance is ultimately determined by the results of these multiple attentional dynamics.

Experiment 1: Feature-based attention—Color

This experiment is designed, first, to show that FBA causes greater facilitation of processing in an uncued object when that object shares perceptual features with the cued object than when it does not. In this experiment, such greater facilitation is reflected in reduced median RTs to targets in an uncued object (invalid, between-object targets) that has the same color as the cued object, as compared to an uncued object that does not share color with the cued object. Furthermore, feature similarity across objects alters the attentional gradient in the surround of the objects. Given that an identical uncued object benefits from FBA, its processing is enhanced, and objects that appear in its surround also inherit some of the enhancement. Thus, the featural similarity between the cued and uncued objects modulates the amount of facilitation that accrues between objects and from the objects to locations in the surround.

Method

Participants—All 45 participants (20 female, 25 male) were undergraduate students who gave written informed consent and received either payment ($7) or course credit for their participation. Three of the participants were excluded due to high error rates (∼20%), and an additional participant whose median RT was more than 2.5 standard deviations from the group mean was also excluded.[1] All participants had normal or corrected-to-normal vision. This protocol was approved by the Institutional Review Board (IRB) of Carnegie Mellon University.

---

[1]Note that in both this experiment and the subsequent ones, we also used other rules for excluding participants, including boxplot and median absolute deviation rules. All of these methods produced the same exclusions, save for 1 participant in Experiment 3. The inclusion/exclusion of this participant had no qualitative impact on the reported results.

Stimulus, apparatus, and procedure—Participants were seated in a dimly lit room approximately 48 cm from a 19-in. Dell M992 Flat CRT monitor. The experiment was run using the E-Prime software package, running on a Dell Dimension 8200 computer.

The sequence of stimuli and their timing within a trial followed a classic paradigm often used for studying OBA (Egly et al., 1994). Each trial began with a 1-s stimulus display, which contained a pair of objects (rectangles) flanking a fixation cross at the center of the screen. The background of this display was white (CIE: $x = .254$, $y = .325$) with a luminance of 57 cd/m^2. The objects were either blue (CIE: $x = 156$, $y = .106$) or red (CIE: $x = 628$, $y = .346$) with a luminance of 12 cd/m^2. Following the stimulus display, a black cue overlaying one end of the contour of one rectangle appeared for 100 ms. After a delay of 100 ms, the target (a black "T" or "L") appeared simultaneously with the stimulus display for a further 1 s. Participants had 1,500 ms from the target onset to respond (was a "T" or an "L" present?) by pressing one of two keys. The response triggered the onset of a 500-ms intertrial interval (ITI), after which the next trial started. In the absence of a response, the ITI began automatically after 1,500 ms. Participants were instructed to maintain fixation throughout each trial and to respond as quickly as possible to the target while minimizing errors in identifying the targets. Each participant completed five blocks of 256 randomized trials, each containing 128 valid (64 per object condition) and 128 invalid (32 per invalid location) trials.

Design—The stimulus display contained two vertical colored rectangles, one on either side of a central fixation cross. The color of these rectangles was manipulated to create two object conditions in which both rectangles either shared color (*identical*) or did not (*nonidentical*). The target ("T" or "L") could appear in the same location as the cue (*valid*) or in one of four invalid locations (*within-object*, *between-objects*, *near-object*, *far-object*), all equidistant from the cued location (5.8°) and all probed with equal frequency (see Fig. 2). The near- and far-object locations (exterior targets) were positioned 7.9° from fixation. The within- and between-object locations (interior targets) were 4.1° from fixation.

In both the identical and nonidentical object conditions, we expected to observe faster median RTs to targets located in the valid than in the invalid spatial location. Consistent with OBA, we also expected that targets appearing in the invalid position of the cued object (within-object target) would be more quickly discriminated than targets appearing in the uncued object (between-object target) in both the identical and nonidentical object conditions. If attentional facilitation in an uncued object is dependent on its featural similarity to the cued object (FBA), we should observe faster responses to targets in the uncued object in the identical than in the nonidentical color condition. This should be reflected in shorter median RTs to the between-object targets in the identical than in the nonidentical condition, which would indicate the interaction of FBA with OBA.

Under our hypothesis of multiple forms of attention operating interactively, we also expected the accrual of facilitation within the uncued object to have a knock-on or cascade effect on median RTs to the targets appearing exterior to the objects—that is, in the surround. Locations within the uncued object in the identical condition were predicted to receive more facilitation than locations in uncued objects in the nonidentical condition, with the result that, in the former case, there would be a greater shift in the scene-wide attentional distribution toward the identical uncued object. Figure 1c shows the distances between the exterior targets and a theoretical scene-wide center of mass of attentional distribution in both of the object conditions. In the identical condition, the facilitated locations within the uncued object would draw the center of mass toward fixation—that is, equidistant from the two exterior targets (Fig. 1c). Therefore, we would expect the difference in median RTs between the exterior targets to be smaller in the identical than in the nonidentical condition. As any

shift in the center of mass should be horizontal, the distance between it and the near-object target should change much more than the distance to the far-object target. Therefore, we expected slower median RTs to the near-object target to be paired with slightly faster median RTs to the far-object target in the identical condition relative to the nonidentical condition. This finding would be consistent with spatial facilitation afforded by OBA— which, in turn, would be afforded by FBA.

Results and discussion

Accuracy in identifying the target letter was uniformly high (>88%) and was unaffected by cue location, left-object color, object condition, or target location (all $p > .4$). Median RTs[2] for each participant for correct trials only were analyzed in a four-way repeated measures ANOVA, with cue location (upper left, upper right, lower left, lower right), left-object color (red, blue), object condition (identical, nonidentical), and target location (valid, within-object, between-objects, near-object, far-object) as within-subjects factors. No main effects or interactions were observed with either cue location ($p = .14$) or left-object color ($p = .22$). Median RTs differed as a function of target location [$F(4, 160) = 51.554$, $p < .001$, $\eta_p^2 = .56$], and an interaction between object condition and target location was observed [$F(4, 160) = 5.179$, $p = .005$, $\eta_p^2 = .12$] (Note that all p values reported in this experiment and the later ones are Greenhouse–Geisser corrected) (Fig. 3a).

At the outset, we confirmed SBA resulting from spatial proximity to the cue. A series of planned paired one-tailed t tests between the valid and the other target locations in both object conditions revealed that valid trials were responded to significantly more quickly than any other trials (all $ps < .001$). This finding verifies the enhancement at attended over unattended spatial locations (SBA: valid vs. invalid trials, both within and between objects).

To test interactions between OBA and FBA and then their joint effect on SBA, and to better visualize the results in light of any baseline median RT differences that might arise between subjects, difference scores (Fig. 3b) were computed by subtracting the median valid RT from the median RTs at the other target locations and were entered into a two-way repeated measures ANOVA with object condition (identical, nonidentical) and target location (within-object, between-objects, near-object, far-object) as within-subjects factors. This ANOVA yielded a main effect of target location [$F(3, 120) = 11.54$, $p < .001$, $\eta_p^2 = .22$] and an interaction between object condition and target location [$F(3, 120) = 4.78$, $p = .004$, $\eta_p^2 = .11$].

To examine the effects of object condition on OBA, we performed a series of planned paired one-tailed t tests that revealed that the pattern of difference scores for the interior targets followed the expected pattern (Fig. 3b). The presence of OBA was confirmed, with significantly larger difference scores for the between-object than for the within-object target locations in both object conditions [identical, $t(1, 40) = 2.27$, $p = .013$; nonidentical, $t(1, 40) = 3.99$, $p < .001$]. There was no significant difference between the difference scores for within-object targets [$t(1, 40) = 0.85$, $p = .2$]. As predicted by our hypothesis of an interaction between FBA and OBA, the difference score for between-object targets was significantly less in the identical than in the nonidentical condition [$t(1, 40) = 3.05$, $p = .002$], suggesting facilitation and reduced costs to shift attention between objects that share the feature of color. However, these effects were not strong enough to drive a significant object condition (identical, nonidentical) × target location (within-object, between-objects) interaction in an ANOVA limited to the interior targets [$F(3, 120) = 2.45$, $p = .13$, $\eta_p^2 = .06$].

[2]We also analyzed the data from this and the subsequent experiments using the participant means rather than medians, and we found no qualitative differences in any of the reported effects.

Finally, we examined the consequences of the conjoined effects of SBA, OBA, and FBA on attention to the spatial gradient in the surround of the rectangles. We entered the difference scores into a two-way ANOVA with object condition (identical, nonidentical) and target location (near-object, far-object) as within-subjects factors (Fig. 3b). This ANOVA revealed a significant main effect of target location [$F(1, 40) = 29.02$, $p < .001$, $\vert_p^2 = .43$] and a significant interaction between object condition and target location [$F (1, 40) = 4.66$, $p < .037$, $\vert_p^2 = .11$]. The source of this interaction was the predicted increase in median RTs to near-object targets and slight decrease in median RTs to far-object targets in the identical object condition, as revealed by a set of planned one-tailed t tests. First, we replicated our previous results (Kravitz & Behrmann, 2008) with significantly larger difference scores in the far-object than in the near-object target location in both object conditions [identical, $t(1, 40) = 2.27$, $p = .014$; nonidentical, $t(1, 40) = 5.44$, $p < .001$]. As predicted, the difference score for the near-object targets was significantly larger in the identical than in the nonidentical condition [$t(1, 40) = 2.61$, $p = .013$], suggesting an impact of identical uncued objects (FBA) on the spatial gradient in the surround. The difference scores for the far-object target location did not vary significantly between the identical and nonidentical conditions [$t(1, 40) = 0.51$, $p > .1$], as predicted by the small change in distance to center of mass between the two object conditions (Fig. 1c and d). The small difference that was observed went in the expected direction, with a slightly reduced cost in discriminating the far-object target in the identical condition (nonidentical – identical = 1.76 ms).

In sum, in this experiment, we replicated the standard object-based effects obtained using this well-established two-rectangle Egly et al. (1994) paradigm. We also report two findings that verify our prediction that identical uncued objects would accrue greater attentional facilitation than would nonidentical objects. First, median RTs to between-object targets were reduced when the two objects were identical, suggesting greater attentional facilitation between objects that share features (FBA interacting with OBA). Second, the enhanced facilitation for identical uncued objects resulting from the FBA/OBA interaction caused a shift in the attentional center of mass toward the fixation cross (Fig. 1c). This shift was reflected in slower median RTs to the near-object targets in the identical than in the nonidentical condition. The latter finding illustrates the reshaping of the spatial gradient in the surround by virtue of the interaction of FBA and OBA.

Experiment 2: Feature-based attention—Shape

Experiment 1 showed that attentional facilitation within an uncued object and its concomitant effect on the scene-wide center of mass of attention were dependent on the sharing of the color feature between the cued and uncued objects. Here, we expand this result to another perceptual dimension, shape, to demonstrate the generality of the claim that SBA, OBA, and FBA interact to determine attentional facilitation throughout the scene. Specifically, we manipulated the shape similarity of the two objects in the display and showed that, as in Experiment 1, the OBA effects differed as a function of the perceptual similarity of the two objects, and furthermore that this had consequences for the distribution of attention in the surround of the display.

Method

Participants—All 43 participants (19 female, 24 male) were undergraduate students with normal or corrected-to-normal vision. All gave written informed consent, in compliance with the IRB approval from Carnegie Mellon University, and received either payment ($7) or course credit for their participation. The apparatus and procedure were identical to those of Experiment 1, unless noted otherwise in the Design section. Four of the participants were excluded from the final analyses due to high error rates (~23%; $n = 2$) or extremely long median RTs ($n = 2$).

Design—The stimulus display contained two objects, one on either side of a central fixation cross. The objects were either two rectangles (*identical*) or one rectangle and one barbell (*nonidentical*) (Fig. 4). The two-barbell display and the reversed nonidentical display were not included in this study because the particular type of display had had no observable impact on behavior in Experiment 1, so we restricted this experiment to the identical rectangles and one version of the nonidentical display. This design also allowed us to reduce the number of trials in a block from 256 to 128, making the task significantly less strenuous for the participants. Each participant completed 10 blocks, each consisting of 128 trials, 64 valid (32 per object condition) and 64 invalid (16 per target location) trials.

As in Experiment 1, the cue involved the darkening of one end of one of the objects. In the case of the barbell, the cue followed the curved contour of the object. The target could appear either at the cued location (valid) or in one of four invalid locations (within-object, between-objects, near-object, far-object), all in exactly the same positions as in Experiment 1; the invalid locations were all probed with equal probability.

Results and discussion

Accuracy in identifying the target letter was uniformly high (>92%) and was unaffected by cue location, object condition, or target location ($p > .4$). As in Experiment 1, a four-way ANOVA revealed no significant main effects or interactions involving cue location (all ps > .1), so the remaining analyses were collapsed over this factor. Participant median RTs (Fig. 5a) were analyzed in a two-way repeated measures ANOVA with object condition (identical, nonidentical) and target location (valid, within-object, between-objects, near-object, far-object) as within-subjects factors. A significant main effect of target location [$F(4, 152) = 60.79$, $p < .001$, $\eta_p^2 = .62$] and an interaction between object condition and target location [$F(4, 152) = 4.72$, $p = .002$, $\eta_p^2 = .14$] were observed.

To establish the presence of SBA, we conducted a series of planned paired one-tailed t tests between valid and the other target locations in both object conditions. Valid trials were responded to significantly more quickly than any other trials (all ps < .001), verifying the attentional facilitation resulting from spatial proximity to the cue.

As in Experiment 1, difference scores (valid – invalid locations; Fig. 5b) were derived and analyzed in a two-way repeated measures ANOVA with object condition (identical, nonidentical) and target location (within-object, between-objects, near-object, far-object) as within-subjects factors. A significant main effect of target location [$F(3, 114) = 15.99$, $p < .001$, $\eta_p^2 = .30$] and a significant interaction between object condition and target location [$F(3, 114) = 6.72$, $p < .001$, $\eta_p^2 = .15$] were obtained.

A series of planned paired one-tailed t tests revealed that the pattern of difference scores for the interior targets (Fig. 5b) matched that observed in Experiment 1. First, we verified the significantly reduced difference scores for the within-object relative to the between-object targets in both object conditions [identical, $t(1, 38) = 3.11$, $p = .002$; nonidentical, $t(1, 38) = 5.16$, $p < .001$], reflecting the standard object-based facilitation result. Then, we confirmed, as predicted, that the difference score for between-object targets was significantly less when the uncued object was identical rather than nonidentical to the cued object [$t(1, 38) = 3.29$, $p = .001$]. This result suggests an increased cost to shift attention to nonidentical objects and demonstrates the impact of shared features on attentional enhancement. There was no significant difference between the difference scores for the identical and nonidentical conditions for within-object targets [$t(1, 38) = 1.58$, $p = .06$]. As in Experiment 1, these effects were not strong enough to drive a significant object condition (identical, nonidentical) × target location (within-object, between-objects) interaction in an ANOVA limited to the interior targets [$F(3, 114) = 1.87$, $p = .18$, $\eta_p^2 = .05$]. However a combined

three-way ANOVA over the results of Experiments 1 and 2 with object condition (identical, nonidentical) and target location (within-object, between-objects) as within-subjects factors and experiment (1, 2) as a between-subjects factor did reveal a significant object condition × target location interaction [$F(3, 114) = 4.27$, $p = .04$, $_p^2 = .05$].

Performance across the exterior targets also matched the predicted pattern (Fig. 5b) and illustrates the alteration of the spatial gradient by virtue of the attentional facilitation resulting from the joint influences of FBA and OBA observed above. We entered the difference scores for the exterior objects into a two-way ANOVA with object condition (identical, nonidentical) and target location (near-object, far-object) as within-subjects factors (Fig. 5b). This ANOVA revealed a significant main effect of target location [$F(1, 38) = 12.82$, $p = .001$, $_p^2 = .24$] and a significant interaction between object condition and target location [$F(1, 38) = 5.58$, $p = .023$, $_p^2 = .13$]. We replicated our previous results, with significantly greater difference scores in the far-object than in the near-object target location in the nonidentical object condition [$t(1, 40) = 4.47$, $p < .001$]. The difference between the exterior targets did not reach significance in the identical object condition [$t(1, 40) = 0.80$, $p = .43$], perhaps suggesting a stronger effect of shape (Exp. 1) than of color (Exp. 2) on the exterior targets.[3] As predicted, the impact on the exterior targets was primarily evident in greater difference scores for the near-object targets in the identical as compared to the nonidentical condition [$t(1, 40) = 2.31$, $p = .013$]. Also, as in Experiment 1, the difference scores for the far-object target location did not vary significantly between the identical and nonidentical conditions [$t(1, 40) = 0.92$, $p = .36$]. The small difference that was observed went in the expected direction, with a slightly reduced cost in discriminating the far-object target in the identical condition (nonidentical – identical = 2.33 ms).

The only unexpected pattern in the data in either Experiment 1 or 2 was the overall lack of a median RT advantage for the interior over the exterior targets, which one might have expected given their relative proximity to the fixation cross. One possible explanation of this result lies in the demands of the discrimination task. In Experiment 1, the color of the objects reduced the contrast of the interior targets relative to the exterior targets, which could lead to longer median RTs. The proximity of the interior targets to the contour of the objects might also have placed them at a slight disadvantage in Experiment 2. The contours of the object were the same width and contained the same angles as the target letters, perhaps hindering the ability of participants to make the perceptual discrimination. This effect might be similar to crowding (Pelli & Tillman, 2008), wherein, for example, letter detection or reading in the periphery is hindered by the presence of flanking items with similar features.

Experiment 3: Category-based attention

Experiments 1 and 2 demonstrated that the speed of target discrimination in a two-object display was dependent on the extent to which perceptual features were shared across the cued and uncued objects. When the uncued object was perceptually identical to the cued object, a target in the uncued object was detected faster than it would be in a nonidentical uncued object, and target discrimination in the surround was also affected. Thus far, we have shown that perceptual similarity impacts the distribution of attention across a multiobject scene. In this final experiment, we examined whether the high-level conceptual or semantic similarity between objects might have a similar impact on the distribution of attention throughout the display. High-level categorical representations have been shown to impact

---

[3]While it is tempting to compare these two experiments directly, there was more variation in the stimulus displays in the Experiment 1 than in Experiment 2. Although the data do suggest a stronger effect of shape than of color similarity, a direct comparison would require a more closely matched design.

visual search; for example, Yang and Zelinsky (2009) showed that a category-defined target (for example, a teddy bear) was fixated far sooner than would be expected by chance, and that this effect derived from observers using a categorical model of common features from the target class (see also Schmidt & Zelinsky, 2009). Here we tested whether top-down categorical representations interact with OBA and SBA in ways akin to those of the perceptual dimensions of color and shape. Specifically, we manipulated whether the two objects in the scene were of the same ("H"/"h") or different ("H"/"4") categories or labels. We found that shared semantic category (independent of perceptual similarity) had effects on attentional facilitation similar to those found for the perceptual dimensions. This experiment demonstrated the generality of the effects found in the first two experiments, and based on these findings, we propose that similarity across objects, defined either by bottom-up or top-down properties, has similar consequences for the deployment of visual attention in a multiobject scene.

Method

Participants—All 57 participants (24 female, 33 male) were undergraduate students with normal or corrected-to-normal vision. All gave written informed consent, in compliance with the IRB approval from Carnegie Mellon University, and received either payment ($7) or course credit for their participation. The apparatus and procedure were identical to those in Experiments 1 and 2, unless noted otherwise in the Design section. A total of 21 participants were excluded from the final analyses due to high false alarm rates ($\sim$35%; $n = 16$) or extremely long median RTs ($n = 5$) under the same criteria used in our previous study (Kravitz & Behrmann, 2008). The increased rate of participant exclusions relative to Experiments 1 and 2 resulted primarily from the increased fatigue caused by the target detection, as compared to target discrimination, and the need to use false alarms as opposed to accuracy as the exclusion criteria. Including these excluded participants did not qualitatively alter the reported pattern of results below, but it did significantly increase the noise.

Design—The stimulus display contained two objects, one on either side of a central fixation cross. One object was always an uppercase "H," while the other object was either a lowercase "h" (identical category) or a "4" (nonidentical category) (Fig. 6). Note that the perceptual similarity of the displays was constant, as the "4" and " h" were exactly the same object, presented either upright or inverted. Thus, any difference between the identical and nonidentical category displays could not be attributed to low-level image characteristics and must be ascribed to the semantic or categorical relationship between the objects. As in Experiment 1, the cue involved the darkening of one end of one of the objects closest to fixation. The interior targets were in exactly the same positions as those in Experiments 1 and 2. The exterior targets were slightly farther away from fixation (10.4°) and from the cue (8.04°) to accommodate the additional width of the objects. The exterior targets remained equidistant from both the cue and fixation locations.

Unlike in the previous experiments, the target was a black square that participants had to respond to with a keypress as quickly as possible. The switch to target detection from discrimination was made in order to eliminate any effects of the letter stimuli on the judgment of letter identities (we were concerned about possible interference effects from doing target discrimination of letters embedded in global letters/numbers). Catch trials, in which no target appeared, were introduced to ensure that participants waited until the target appeared to press the key. As in the first two experiments, the target could appear either at the cued location (valid) or in one of four invalid locations (within-object, between-objects, near-object, far-object), and these invalid locations were all probed with equal probability.

Trials were organized into 15 blocks of 80 trials, each consisting of 40 valid (20 per object condition), 32 invalid (8 per target location), and 8 catch (2 per cue location) trials.

Results and discussion

As in Experiment 1, a four-way ANOVA revealed no significant main effects or interactions involving cue location ($p > .1$), so the remaining analyses were collapsed over this factor. Participant median RTs (Fig. 7a) were analyzed in a two-way repeated measures ANOVA, with object condition (identical category, nonidentical category) and target location (valid, within-object, between-objects, near-object, far-object) as within-subjects factors. Significant main effects of target location [$F(4, 140) = 116.56$, $p < .001$, $\eta_p^2 = .77$] and object condition [$F(1, 35) = 6.24$, $p = .017$, $\eta_p^2 = .15$] were observed, as well as an interaction between object condition and target location [$F(4, 140) = 3.87$, $p = .006$, $\eta_p^2 = .1$]. This interaction implies that categorical similarity affected detection times across target locations.

We broke down this interaction by first establishing the presence of SBA, via a series of planned paired one-tailed t tests between the valid and the other target locations in both object conditions. Valid targets were detected significantly more quickly than any other target type (all $ps < .001$), verifying the attentional facilitation resulting from spatial proximity to the cue.

We then removed the effect of validity on the pattern of results by deriving difference scores, as in Experiment 1 (valid – invalid locations; Fig. 7b) and entered them into in a two-way repeated measures ANOVA with object condition (identical, nonidentical) and target location (within-object, between-objects, near-object, far-object) as within-subjects factors. A significant main effect of target location [$F(3, 105) = 75.72$, $p < .001$, $\eta_p^2 = .68$] and a significant interaction between object condition and target location [$F(3, 105) = 3.21$, $p = .026$, $\eta_p^2 = .10$] were obtained, again indicating the effect of our categorical manipulation on the pattern of median RTs across locations.

Performance across the exterior targets matched the predicted pattern (Fig. 7b) as well as the previously observed alteration of the spatial gradient observed in Experiments 1 and 2. We entered the difference scores into a two-way ANOVA with object condition (identical, nonidentical) and target location (near-object, far-object) as within-subjects factors (Fig. 7b). This ANOVA revealed a significant main effect of target location [$F(1, 35) = 5.91$, $p = .02$, $\eta_p^2 = .14$] and a significant interaction between object condition and target location [$F(1, 35) = 4.49$, $p = .04$, $\eta_p^2 = .11$]. We replicated the results of Experiments 1 and 2, with significantly greater difference scores in the far-object than in the near-object target location in the nonidentical object condition [$t(1, 35) = 2.97$, $p = .003$]. The difference between the exterior targets did not reach statistical significance in the identical object condition [$t(1, 35) = 0.12$, $p = .90$]. As in Experiments 1 and 2, only the near-object targets evidenced a significant effect of object condition [$t(1, 35) = 1.73$, $p = .045$]. These results verify that the categorical similarity between objects has an impact similar to that of perceptual similarity on the distribution of attention in the surround.

Regarding the interior targets, a series of planned paired one-tailed t tests revealed the presence of OBA in both object conditions [identical category, $t(1, 35) = 2.35$, $p = .013$; nonidentical category, $t(1, 35) = 3.06$, $p = .002$]. Unlike in the previous experiments, we did not find a significant difference between median RTs to between-object targets in the two object conditions [$t(1, 35) = 0.87$, $p = .39$]. The absence of this result might be attributable to the reduced power and increased participant fatigue resulting from the target detection paradigm and to the shape difference between the cued and uncued objects present in both object conditions. The objects were also far more complex in shape and larger than those used in Experiments 1 and 2, making the precise distribution of attention across them more

difficult to predict, and perhaps also more variable across both trials and participants. Note also that median RTs were faster in this experiment than in the previous experiments, potentially reflecting a ceiling effect that might also obscure some condition differences.

To quantify whether the data from this third experiment followed the same pattern as those from the two previous experiments, we conducted an omnibus ANOVA with object condition (identical, nonidentical) and target location (valid, within-object, near-object, far-object) as within-subjects factors, and experiment (1, 2, 3) as a between-subjects factor. This analysis revealed a highly significant main effect of target location [$F(3, 339) = 59.16$, $p < .001$, $_p^2 = .34$] and an interaction between object condition and target location [$F(3, 339) = 10.25$, $p < .001$, $_p^2 = .08$]. The only observed interaction or main effect of experiment was an interaction between experiment and target location [$F(3, 339) = 29.41$, $p < .001$, $_p^2 = .34$], resulting from the longer RTs to exterior as compared to interior targets in Experiment 3 (Fig. 7). Since Experiment 3 utilized target detection rather than discrimination, this finding is in keeping with our interpretation that the objects interfered with target discrimination in Experiments 1 and 2 (see above). The lack of a significant three-way interaction between the factors suggests that the patterns of results across object conditions and target locations were very similar across all experiments. As such, we aggregated the data from each of the three experiments and produced boxplots to give an impression of the distribution of the effect of object condition on median RTs to the interior (OBA) and exterior (SBA) targets (Fig. 8). These analyses suggest that even using displays in which the perceptual similarity of the two objects differed, a purely semantic similarity between the cued and uncued objects was sufficient to modulate attention in a way similar to that observed with perceptual similarity in Experiments 1 and 2. In all three cases, there was a strong effect of object perceptual/categorical similarity on SBA in the surround (far – near) and a weaker effect on the strength of OBA (between – within).

General discussion

Much recent research has focused on demonstrating that attention can be directed not only to locations in space (SBA), but also to particular features (FBA) and objects (OBA) that appear in visual displays. We have now shown that all three of these forms of attention can function simultaneously, and that the consequences of attending in one domain extend to other domains as well. The major finding of the present investigation is that not only does OBA facilitate processing within the bounds of an attended object, and not only does it interact with SBA to determine facilitation throughout a scene (Kravitz & Behrmann, 2008), but FBA (both bottom-up and top-down) is also engaged and modulates that interaction.

Specifically, we have shown that uncued objects that are either perceptually or semantically identical to a cued object accrue greater facilitation than do uncued objects that differ from the cued object along these dimensions. The increased facilitation is reflected in a reduction in the between-object cost, with faster median RTs to targets in identical than in nonidentical uncued objects. Additionally, the facilitation extends to targets that are exterior to the objects in the scene, in that enhanced facilitation of performance also occurs in the surround of the uncued object when this uncued object is identical to the cued object, but not when it is nonidentical. Thus, SBA, FBA, and OBA all interact to create a pattern of attentional facilitation across the entire scene based on the similarity of the uncued object to the cued stimulus across multiple dimensions (spatial proximity, object membership, featural similarity, and categorical similarity). While many other studies have shown the scene-wide (e.g., Busse, Roberts, Crist, Weissman, & Woldorff, 2005; Serences & Yantis, 2007) and early (e.g., Schoenfeld et al., 2003; Zhang & Luck, 2009) effects of FBA, and others have shown the pairwise effects, say, of SBA and OBA or of OBA and FBA (O'Craven, Downing, & Kanwisher, 1999) operating simultaneously, our results are the first to establish

the joint interaction of FBA, SBA, and OBA. Furthermore, we have established that the impact of categorical similarity on the pattern of attentional facilitation throughout a scene is similar to the impact of perceptual similarity. That categorical similarity, which must be dependent on top-down input, has an impact on attention similar to that of bottom-up perceptual similarity suggests that in both cases the attentional dynamics operate in comparable manners, regardless of the source of the input. Collectively, our results suggest that attention facilitation is determined, simultaneously, by similarity across both perceptual and categorical dimensions (see also Zelinsky, 2008, for similar principles applied to overt eye movements during search).

Taken together, these results argue against the notion that attentional selection occurs on the basis of cooperation or competition within a single stimulus dimension, such as spatial position, in which one position is enhanced relative to other positions in the field. Rather, the findings suggest that the observed pattern of facilitation is the product of multiple interactions within and between several distinct representations. Even a purely spatial cue affected both OBA and FBA, and in turn, the modulation of those representations influenced the final distribution of SBA. These results are consistent with a distributed attention mechanism (e.g., Desimone & Duncan, 1995; Duncan, 2006) based on cooperation and competition within and between different levels of representation or dimensions (e.g., color, space, or semantics). Within any particular level, these interactions are defined by similarity along that dimension, such that uncued items that share colors, locations, shapes, or labels (H/h) similar to those of cued items will come to be facilitated as well. In turn, these dimensions will interact with each other, such that the overall pattern of facilitation is determined by similarity on all dimensions simultaneously. First, our results, along with our previous findings (Kravitz & Behrmann, 2008), provide a critical verification of the effects of similarity and the interactions between levels of representation predicted by distributed accounts of attention. Second, the present results demonstrate that these interactions and effects are observed even when one of the levels of representation is categorical rather than perceptual. The key implication of these findings is that attention results from dynamic interactions across a distributed system in which common principles and/or mechanisms operate in parallel over different forms of perceptual and categorical representations.

Two recent studies, one focused on SBA (Lee & Maunsell, 2009) and the other on an extension of FBA (Reynolds & Heeger, 2009), have argued that one such principle, normalization, might account for many attentional effects. Our findings are clearly compatible with any approach proposing that attention arises from distributed interactions within the perceptual system, but they may also require some extension of these models. We have demonstrated that a spatial cue can simultaneously lead to an advantage for the cued location, for OBA, for FBA across locations, and for the formation of a scene-wide spatial gradient. As such, our results necessitate an expansion of these models to account for FBA beyond a binary spatial/featural strategy (Reynolds & Heeger, 2009).

We have demonstrated that attentional selection, as reflected by pattern of median RTs to targets throughout a scene, is intimately tied to the perceptual representations mediating the entire scene. The amount of facilitation that accrues to any item (location, feature, object, label, or category) in the scene is dependent on its similarity to the attended item. Previous studies have also reported that attentional facilitation of an item intensifies its perception (Carrasco, Ling, & Read, 2004), and a growing literature is also demonstrating that attention alters the response of perceptual cortex (e.g., Müller & Kleinschmidt, 2003; Shomstein & Behrmann, 2006; Treue & Maunsell, 1999; Vandenberghe et al., 1997). Attention may then better be thought of as a process not just of selection, but also of organization, wherein the entire contents of visual perception come to be configured relative to the attended item. Attention then reduces the complexity of a scene not in terms of the number of items [O(n)],

but in terms of the relationships between those items [pairwise: $O(n^2)$]. Thus, attention may serve as a mechanism by which scenes are organized, rather than selected from once the organization is derived. Such a mechanism would be integrated with and operate in tandem with perceptual processing, rather than serving as a distinct process in which preattentive mechanisms establish the initial structure of the input, followed by selection from that structure.

In sum, our findings suggest that multiple factors play roles in determining the organization of visual input and that attentional dynamics participate in this process. Many open questions remain, however. For example, in all of our experiments we only used displays of 2 objects, but recent evidence has suggested that as many as 12 objects can be potentially selected (de-Wit, Cole, Kentridge, & Milner, 2011). Whether the full set of attentional dynamics we have illustrated play out equivalently in highly complex scenes remains to be determined.

Of course, we are plagued, as is the field as a whole, with the perennial question of what constitutes an object. One possibility is that the findings we have observed emerges from the fact that, in the identical condition, the two rectangles present in the displays form one larger perceptual entity/gestalt, while in the different condition, the two rectangles are segregated into two disjunct units. In the former case, cuing one of the identical objects would be similar to cuing one half of the greater entity, which would explain the spread of attention from one side to the other in the identical condition but not in the nonidentical condition. One recent study has indicated that only "objecthood" affords OBA facilitation; using displays containing identical features that were perceived either as a bound object or as a set of unbound features, Naber, Carlson, Verstraten, and Einhäuser (2011) demonstrated OBA in the former but not in the latter case. Our two rectangles may have been integrated into a higher-order "meta-object," in which case the effects we observed might be attributable to this objecthood. The existing literature suggests, however, that attention can be deployed at multiple hierarchical scales, from parts of an object (Vecera, Behrmann, & Filapek, 2001), to objects in their own right, to global entities (Yeari & Goldsmith, 2011). While we cannot pinpoint decisively where in the hierarchy our displays fit (indeed, this is the kernel of the problem of defining what an object is), our findings support the notion that objects may serve as units for attentional selection.

Finally, we have shown that objects drawn from the same semantic category facilitate the distribution of attention to the uncued object. This type of selection is consistent with studies showing that "late" representations also influence selection. Much of the work in the OBA literature, however, has been concerned with a particular type of "late" representation—that is, a grouped array or one that is spatially invariant (for recent examples, see Hollingworth et al., in press; Matsukura & Vecera, 2011))—whereas we have focused on higher-order representations that are semantic or categorical in nature. Clearly, much research remains to be conducted to elucidate the nature of these higher-level representations and whether items that are drawn from the same semantic category but that do not share the same label (as in our case) also constrain the deployment of attention in a scene.

Taken together, our experiments provide clear support for the biased-competition account of attention, and we show that multiple representations serve as potential candidates for attentional selection. Moreover, when these representations are selected, they have considerable influence on the remaining representations to the extent that the representations are similar along some perceptual or conceptual dimension. Eventually the dynamics resolve, and a coherent scene-wide description is derived.

Acknowledgments

References

Baylis GC, Driver J. Visual parsing and response competition: The effect of grouping factors. Perception & Psychophysics. 1992; 51:145–162.10.3758/BF03212239 [PubMed: 1549433]

Busse L, Roberts KC, Crist RE, Weissman DH, Woldorff MG. The spread of attention across modalities and space in a multisensory object. Proceedings of the National Academy of Sciences. 2005; 102:18751–18756.10.1073/pnas.0507704102

Carrasco M, Ling S, Read S. Attention alters appearance. Nature Neuroscience. 2004; 7:308–313.10.1038/nn1194

Deco G, Zihl J. A biased competition based neurodynamical model of visual neglect. Medical Engineering & Physics. 2004; 26:733–743. [PubMed: 15564110]

Desimone R. Visual attention mediated by biased competition in extrastriate visual cortex. Philosophical Transactions of the Royal Society B. 1998; 353:1245–1255.10.1098/rstb.1998.0280

Desimone R, Duncan J. Neural mechanisms of selective visual attention. Annual Review of Neuroscience. 1995; 18:193–222.10.1146/annurev.ne.18.030195.001205

de-Wit LH, Cole GG, Kentridge RW, Milner AD. The parallel representation of the objects selected by attention. Journal of Vision. 2011; 11(4):1–10. 13. 10.1167/11.4.13

Driver J, Baylis GC. Movement of visual attention: The spotlight metaphor breaks down. Journal of Experimental Psychology Human Perception and Performance. 1989; 15:448–456.10.1037/0096-1523.15.3.448 [PubMed: 2527954]

Duncan J. Selective attention and the organization of visual information. Journal of Experimental Psychology General. 1984; 113:501–517.10.1037/0096-3445.113.4.501 [PubMed: 6240521]

Duncan J. EPS Mid-Career Award 2004: Brain mechanisms of attention. Quarterly Journal of Experimental Psychology. 2006; 59:2–27.10.1080/17470210500260674

Duncan J, Nimmo-Smith I. Objects and attributes in divided attention: Surface and boundary systems. Perception & Psychophysics. 1996; 58:1076–1084.10.3758/BF03206834 [PubMed: 8920843]

Egly R, Driver J, Rafal RD. Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. Journal of Experimental Psychology General. 1994; 123:161–177.10.1037/0096-3445.123.2.161 [PubMed: 8014611]

Gandhi SP, Heeger DJ, Boynton GM. Spatial attention affects brain activity in human primary visual cortex. Proceedings of the National Academy of Sciences. 1999; 96:3314–3319.10.1073/pnas.96.6.3314

Harms L, Bundesen C. Color segregation and selective attention in a nonsearch task. Perception & Psychophysics. 1983; 33:11–19.10.3758/BF03205861 [PubMed: 6844088]

Hollingworth A, Maxcey-Richard AM, Vecera SP. The spatial distribution of attention within and across objects. Journal of Experimental Psychology: Human Perception and Performance. in press. 10.1037/a0024463

Kim MS, Cave KR. Perceptual grouping via spatial selection in a focused-attention task. Vision Research. 2001; 41:611–624.10.1016/S0042-6989(00)00285-6 [PubMed: 11226506]

Kramer AF, Jacobson A. Perceptual organization and focused attention: The role of objects and proximity in visual processing. Perception & Psychophysics. 1991; 50:267–284.10.3758/BF03206750 [PubMed: 1754368]

Kramer AF, Weber TA, Watson SE. Object-based attentional selection—Grouped arrays or spatially invariant representations? Comment on Vecera and Farah (1994). Journal of Experimental Psychology General. 1997; 126:3–13.10.1037/0096-3445.126.1.3 [PubMed: 9090141]

Kravitz DJ, Behrmann M. The space of an object: Object attention alters the spatial gradient in the surround. Journal of Experimantal Psychology: Human Perception and Performance. 2008; 34:298–309.10.1037/0096-1523.34.2.298

Lee J, Maunsell JH. A normalization model of attentional modulation of single unit responses. PloS One. 2009; 4:34561.

Martínez A, Teder-Sälejärvi W, Vazquez M, Molholm S, Foxe JJ, Javitt DC, et al. Objects are highlighted by spatial attention. Journal of Cognitive Neuroscience. 2006; 18:298–310.10.1162/089892906775783642 [PubMed: 16494688]

Martínez A, Ramanathan DS, Foxe JJ, Javitt DC, Hillyard SA. The role of spatial attention in the selection of real and illusory objects. Journal of Neuroscience. 2007; 27:7963–7973. [PubMed: 17652587]

Martinez-Trujillo JC, Treue S. Feature-based attention increases the selectivity of population responses in primate visual cortex. Current Biology. 2004; 14:744–751.10.1016/j.cub.2004.04.028 [PubMed: 15120065]

Martinez-Trujillo JC, Medendorp WP, Wang H, Crawford JD. Frames of reference for eye–head gaze commands in primate supplementary eye fields. Neuron. 2004; 44:1057–1066.10.1016/j.neuron. 2004.12.004 [PubMed: 15603747]

Matsukura M, Vecera SP. Object-based selection from spatially-invariant representations: Evidence from a feature-report task. Attention, Perception, & Psychophysics. 2011; 73:447–457.10.3758/ s13414-010-0039-9

Maunsell JHR, Treue S. Feature-based attention in visual cortex. Trends in Neurosciences. 2006; 29:317–322.10.1016/j.tins.2006.04.001 [PubMed: 16697058]

Müller NG, Kleinschmidt A. Dynamic interaction of object- and space- based attention in retinotopic visual areas. Journal of Neuroscience. 2003; 23:9812–9816. [PubMed: 14586009]

Naber M, Carlson TA, Verstraten FA, Einhäuser W. Perceptual benefits of objecthood. Journal of Vision. 2011; 11(4):1–9. 8. 10.1167/11.4.8

O'Craven KM, Downing PE, Kanwisher N. fMRI evidence for objects as the units of attentional selection. Nature. 1999; 401:584–587.10.1038/44134 [PubMed: 10524624]

Pelli DG, Tillman KA. The uncrowded window of object recognition. Nature. 2008; 11:1129–1135.10.1038/nn1208-1463b

Posner MI. Orienting of attention. Quarterly Journal of Experimental Psychology. 1980; 32:3–25.10.1080/00335558008248231 [PubMed: 7367577]

Posner MI, Cohen Y, Rafal RD. Neural systems control of spatial orienting. Philosophical Transactions of the Royal Society B. 1982; 298:187–198.10.1098/rstb.1982.0081

Reynolds JH, Heeger DJ. The normalization model of attention. Neuron. 2009; 61:168–185. [PubMed: 19186161]

Reynolds JH, Chelazzi L, Desimone R. Competitive mechanisms subserve attention in macaque areas V2 and V4. Journal of Neuroscience. 1999; 19:1736–1753. [PubMed: 10024360]

Rossi AF, Paradiso MA. Feature-specific effects of selective visual attention. Vision Research. 1995; 35:621–634.10.1016/0042-6989(94)00156-G [PubMed: 7900301]

Saenz M, Buracas GT, Boynton GM. Global effects of feature-based attention in human visual cortex. Nature Neuroscience. 2002; 5:631–632.10.1038/nn876

Schmidt J, Zelinsky GJ. Search guidance is proportional to the categorical specificity of a target cue. Quarterly Journal of Experimental Psychology. 2009; 62:1904–1914.10.1080/17470210902853530

Schoenfeld MA, Tempelmann C, Martinez A, Hopf JM, Sattler C, Heinze HJ, et al. Dynamics of feature binding during object-selective attention. Proceedings of the National Academy of Sciences. 2003; 100:11806–11811.10.1073/pnas.1932820100

Serences JT, Yantis S. Spatially selective representations of voluntary and stimulus-driven attentional priority in human occipital, parietal and frontal cortex. Cerebral Cortex. 2007; 17:284–293. [PubMed: 16514108]

Serences JT, Schwarzbach J, Courtney SM, Golay X, Yantis S. Control of object-based attention in human cortex. Cerebral Cortex. 2004; 14:1346–1357. [PubMed: 15166105]

Shomstein S, Behrmann M. Cortical systems mediating visual attention to both objects and spatial locations. Proceedings of the National Academy of Sciences. 2006; 103:11387–11392.10.1073/ pnas.0601813103

Shomstein S, Behrmann M. Object-based attention: Strength of object representation and attentional guidance. Perception & Psychophysics. 2008; 70:132–144.10.3758/PP.70.1.132 [PubMed: 18306967]

Tootell RBH, Hadjikhani N, Hall EK, Marrett S, Vanduffel W, Vaughan JT, et al. The retinotopy of visual spatial attention. Neuron. 1998; 21:1409–1422.10.1016/S0896-6273(00)80659-5 [PubMed: 9883733]

Treue S, Maunsell JHRM. Effects of attention on the processing of motion in macaque visual cortical areas MT and MST. Journal of Neuroscience. 1999; 19:7603–7616. [PubMed: 10460266]

Vandenberghe R, Duncan J, Dupont P, Ward R, Poline JB, Bormans G, et al. Attention to one or two features in left or right visual field: A positron emission tomography study. Journal of Neuroscience. 1997; 17:3739–3750. [PubMed: 9133394]

Vecera SP, Farah MJ. Does visual attention select objects or locations? Journal of Experimental Psychology General. 1994; 123:146–160.10.1037/0096-3445.123.2.146 [PubMed: 8014610]

Vecera SP, Behrmann M, Filapek V. Attending to the parts of a single object: Part-based selection mechanisms. Perception & Psychophysics. 2001; 63:308–321. [PubMed: 11281105]

Yang H, Zelinsky GJ. Visual search is guided to categorically-defined targets. Vision Research. 2009; 49:2095–2103.10.1016/j.visres.2009.05.017 [PubMed: 19500615]

Yeari M, Goldsmith M. Organizational and spatial dynamics of attentional focusing in hierarchically structured objects. Journal of Experimental Psychology Human Perception and Performance. 2011; 37:758–780.10.1037/a0020746 [PubMed: 21443386]

Zelinsky GJ. A theory of eye movements during target acquisition. Psychological Review. 2008; 115:787–835.10.1037/a0013118 [PubMed: 18954205]

Zhang W, Luck SJ. Feature-based attention modulates feedforward visual processing. Nature Neuroscience. 2009; 12:24–25.10.1038/nn.2223
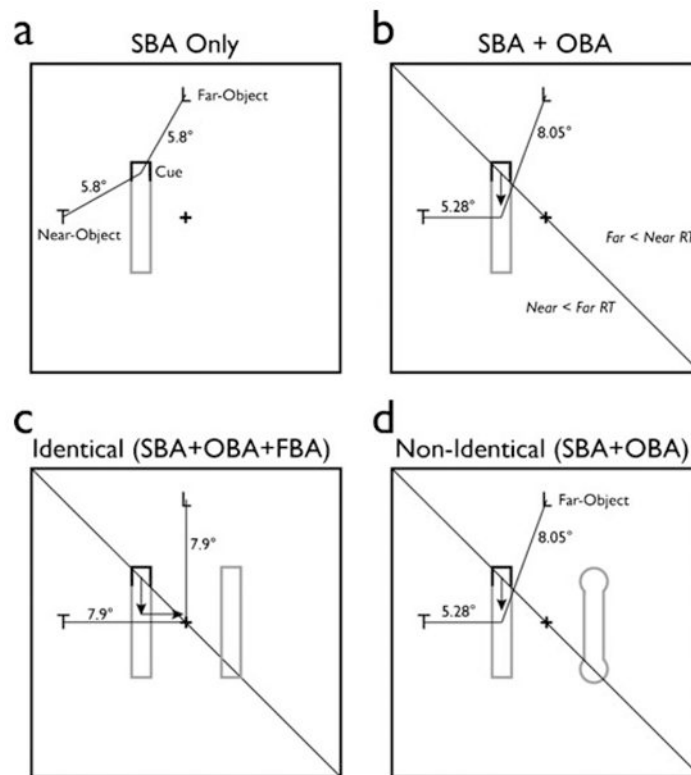
Fig. 1.
Model. (**a**) Illustration of the effect of space-based attention (SBA) following a cue. If the object has no effect on the attentional gradient, its peak will directly overlap the location of the cue, which is equidistant from the near- and far-object target locations. Thus, one would anticipate no difference in reaction times (RT) to these targets. (**b**) Illustration of the interaction between SBA and object-based attention (OBA) (Kravitz & Behrmann, 2008). If the cued object does have an impact, it will cause the peak of the gradient to shift downward (black arrow) toward the cued object's center of mass. Any shift below the black diagonal line moves the peak closer to the nearthan to the far-object target, leading to faster RTs to near-object targets. Note that the distance between the far-object targets and the peak changes more radically than does the near-object distance due to the vertical nature of the shift. (**c**) Illustration of the anticipated interaction between SBA, OBA, and feature-based attention (FBA). If an additional uncued object is added to the display, its impact on the scene-wide attentional gradient should be modulated by its similarity to the cued object. In the case of identical cued objects, the attentional modulation should shift the gradient peak toward the center of the display, and therefore closer to the diagonal black line. Any shift toward this line will reduce the difference in distance between the near- and far-object targets and the peak. (**d**) In the case of nonidentical objects, this shift should be greatly reduced due to the reduced impact of FBA, preserving the difference in RTs between the near- and far-object targets. Note that the shift caused by FBA is horizontal, in contrast to the vertical shift caused by OBA, and as a result, the distance between the near-object target and the peak of the attentional gradient changes more quickly
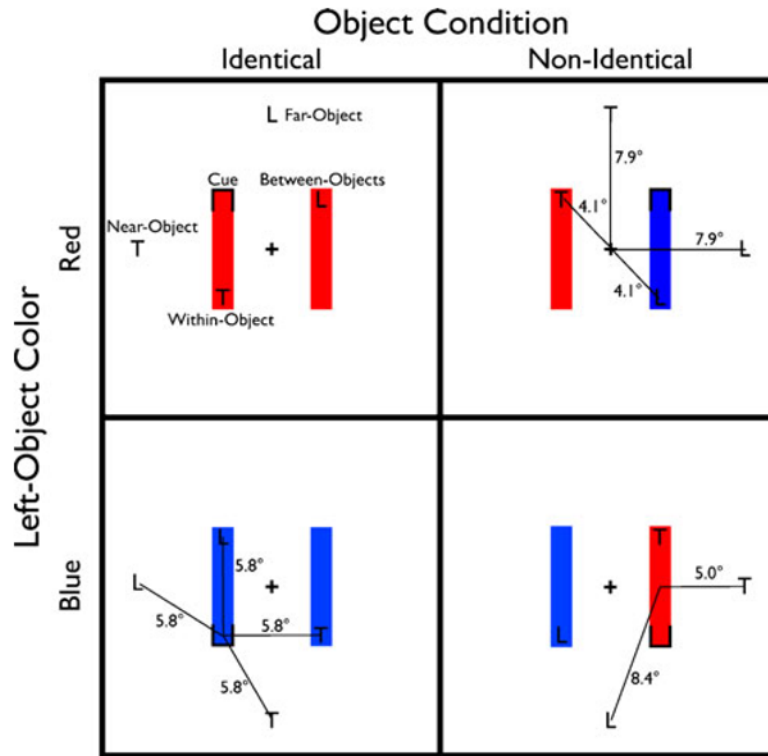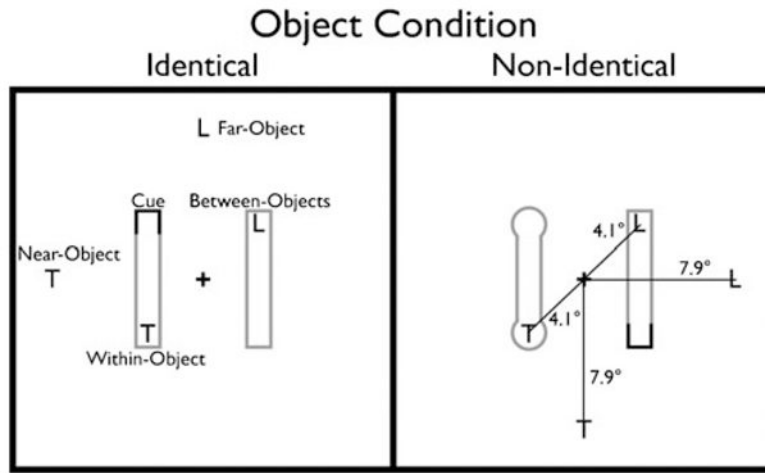
Fig. 2.
Experiment 1: Color task paradigm, illustration of the stimuli and task conditions. Participants were cued to attend to one of four possible locations (one of the ends of the two rectangles). Following the appearance of the cue, a target letter appeared at either the cued location (valid 50%; not shown) or at one of four other invalid locations (upper left panel). All of these invalid locations were equidistant from the cued location (lower left panel). The invalid locations were either interior (within an object) or exterior (outside of any object) conditions, and these varied in their distances from fixation (upper right panel). Finally, note that, although the two exterior targets were equidistant from the cued location, they differed in their distances from the center of mass of the cued object (lower right panel)
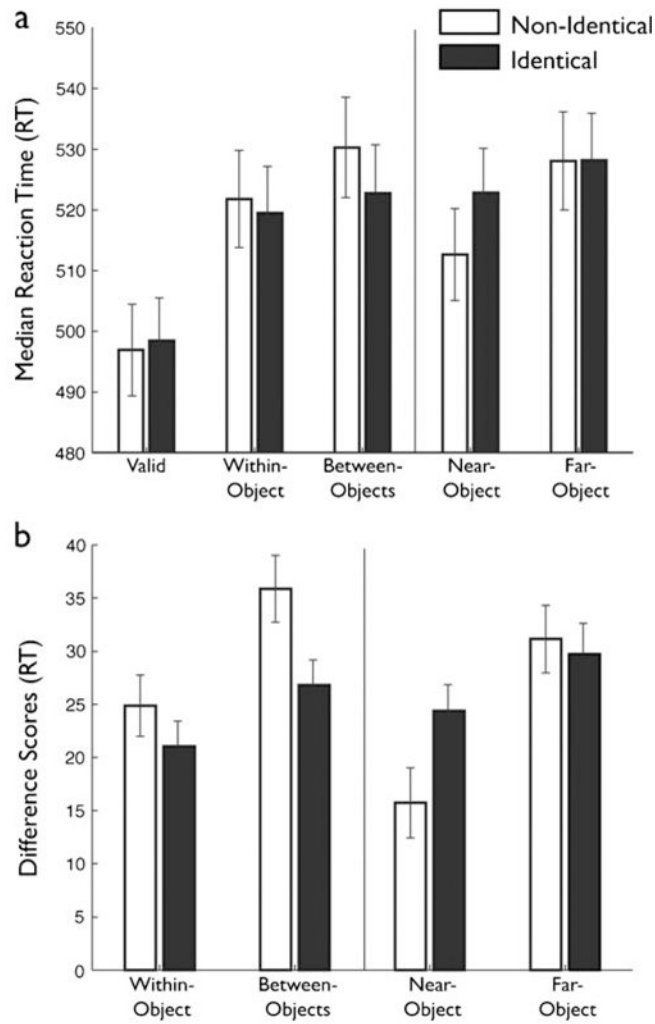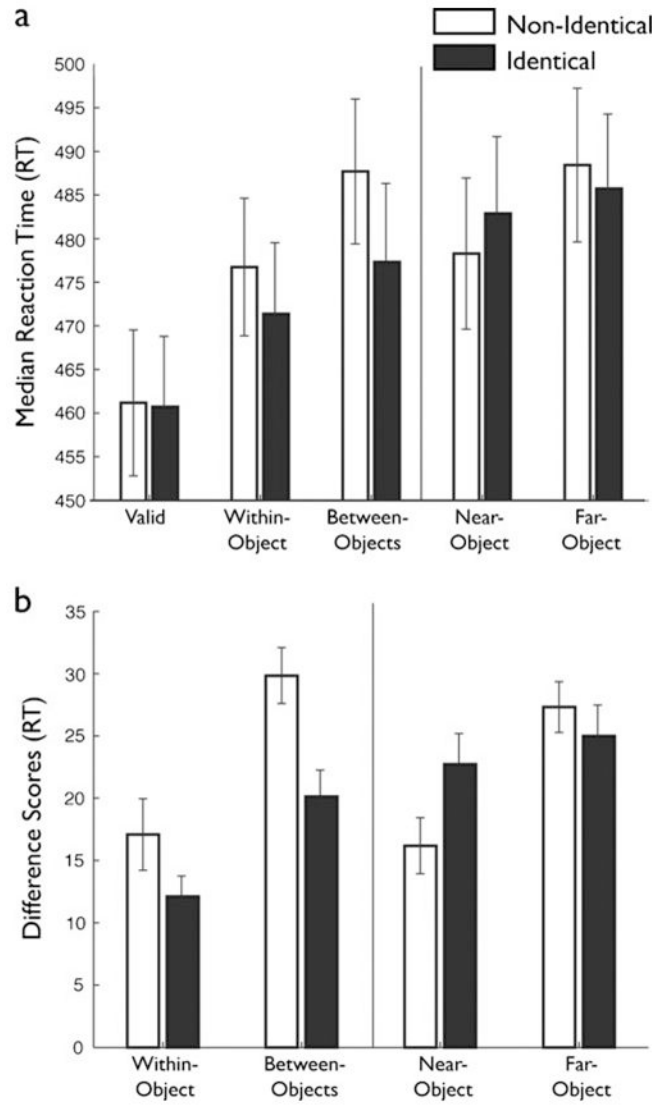
Fig. 3.
Color task results. (**a**) Raw reaction times (RTs) across the five conditions in the color experiment. Identical-trial RTs are depicted with filled bars, and nonidentical-trial RTs with open bars. The vertical line indicates the divide between interior and exterior target locations, which should not be directly compared because of clear stimulus differences. Error bars indicate the between-subjects standard errors. (**b**) Difference scores across the four invalid target locations in the color experiment. Difference scores were derived by subtracting the valid RTs from the invalid RTs. All other plotting conventions are as in panel (a)

Fig. 4.
Experiment 2: Shape task paradigm, illustration of the stimuli and task conditions. The task, stimulus sizes, and target locations in this experiment were identical to those in Experiment 1 (Fig. 2). The only difference was in the displays used, in which objects were defined not by their color but by their shape (either identical or not)

Fig. 5.
Shape results. (**a**) Raw reaction times (RTs) across the five conditions in the shape experiment. Identical-trial RTs are depicted with filled bars, and nonidentical-trial RTs with open bars. The vertical line indicates the divide between interior and exterior target locations, which should not be directly compared because of clear stimulus differences. Error bars indicate the between-subjects standard errors. (**b**) Difference scores across the four invalid target locations in the shape experiment. Difference scores were derived by subtracting the valid RTs from the invalid RTs. All other plotting conventions are as in panel (a)
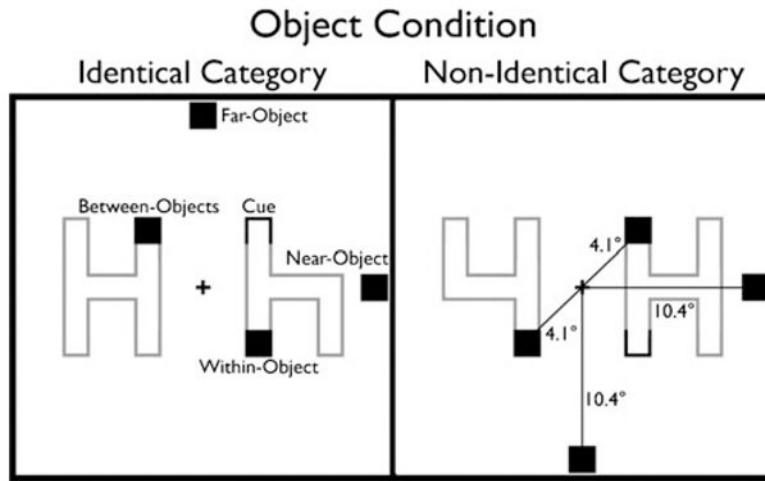
Fig. 6.
Experiment 3: Category task paradigm, illustration of the stimuli and task conditions. The task in this experiment was identical to that in Experiment 1 (Fig. 2). The only difference was in the displays used, in which the objects were letters and numbers; the identical displays contained two symbols from the same category ("H"/"h"), and the nonindentical displays contained two symbols from different categories ("H"/"4"). The targets were black squares rather than letters, to avoid any interference from the letter objects. Furthermore, the exterior targets were moved farther away from fixation to accommodate the wider objects
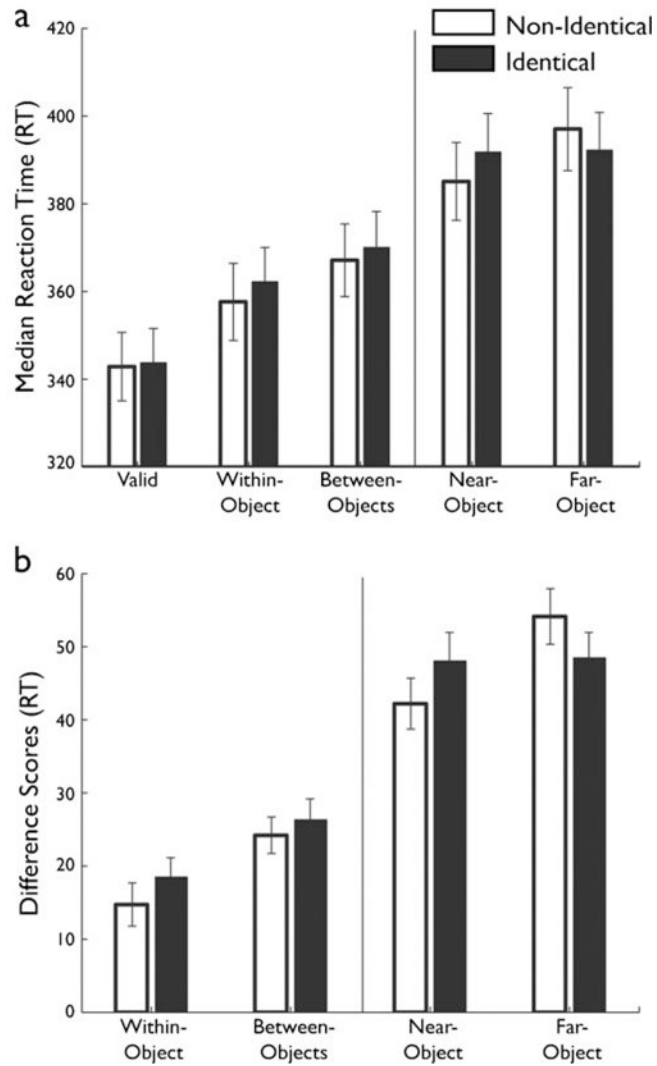
Fig. 7.
Category results. (**a**) Raw reaction times (RTs) across the five conditions in the category experiment. Identical-trial RTs are depicted with filled bars, and nonidentical-trial RTs with open bars. The vertical line indicates the divide between interior and exterior target location, which should not be directly compared because of clear stimulus differences. Error bars indicate the between-subjects standard errors. (**b**) Difference scores across the four invalid target locations in the category experiment. Difference scores were derived by subtracting the valid RTs from the invalid RTs. All other plotting conventions are as in panel (a)
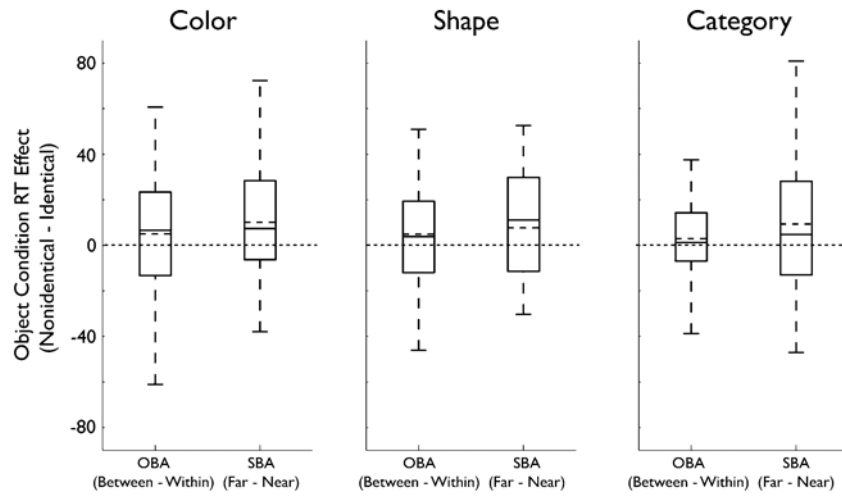
Fig. 8.
Boxplots of aggregated results. The boxplots show the distributions of the effects of object condition on both OBA and SBA in the surround, across participants. To obtain these measures in each participant, the difference of difference scores was calculated [e.g., (nonidentical between – within) – (identical between – within)]. These scores directly reflect the interaction between object condition and target location for the interior (OBA) and exterior (SBA) targets. The long dashed lines indicate 0, which would indicate no effect of object condition. The dashed line within each boxplot represents the median, and the solid line represents the mean across participants