

# Object-based selection from spatially-invariant representations: evidence from a feature-report task

Michi Matsukura · Shaun P. Vecera

Published online: 17 November 2010  
© Psychonomic Society, Inc. 2010

**Abstract** Attention selects objects as well as locations. When attention selects an object's features, observers identify two features from a single object more accurately than two features from two different objects (object-based effect of attention; e.g., Duncan, *Journal of Experimental Psychology: General*, 113, 501–517, 1984). Several studies have demonstrated that object-based attention can operate at a late visual processing stage that is independent of objects' spatial information (Awh, Dhaliwal, Christensen, & Matsukura, *Psychological Science*, 12, 329–334, 2001; Matsukura & Vecera, *Psychonomic Bulletin & Review*, 16, 529–536, 2009; Vecera, *Journal of Experimental Psychology: General*, 126, 14–18, 1997; Vecera & Farah, *Journal of Experimental Psychology: General*, 123, 146–160, 1994). In the present study, we asked two questions regarding this late object-based selection mechanism. In Part I, we investigated how observers' foreknowledge of to-be-reported features allows attention to select objects, as opposed to individual features. Using a feature-report task, a significant object-based effect was observed when to-be-reported features were known in advance but not when this advance knowledge was absent. In Part II, we examined what drives attention to select objects rather than individual features in the absence of observers' foreknowledge of to-be-reported features. Results suggested that, when there was no opportunity for observers to direct their attention to objects that possess to-be-reported features at the time of stimulus presentation, these stimuli must retain strong perceptual cues to establish themselves as separate objects.

**Keywords** Object-based attention · Spatially-invariant representations · Feature-report task

## Introduction

Because the visual system cannot process all the incoming inputs in the world, visual attention enables us to select and process inputs relevant to our goals while ignoring irrelevant inputs. Research has shown that visual objects control attention (e.g., Duncan, 1984; Egly, Driver, & Rafal, 1994; Matsukura & Vecera, 2009; Vecera, 1994; Vecera & Farah, 1994). More than a quarter century ago, Duncan (1984) reported that observers can identify two attributes of a single object more accurately than two attributes of two different objects. In Duncan's (1984) feature-report task, observers were presented with two superimposed objects, a box and a line (Fig. 1a). Each object possessed two properties that observers could report. At the beginning of each block, observers were told which two attributes of the objects to report (i.e., observers were given advance knowledge of to-be-reported features). Observers were more accurate at reporting two target attributes from one object (the *same-object* condition) rather than from two different objects (the *different-objects* condition). Duncan's (1984) finding suggests that attention selects an object and multiple features that belong to that single object with little or no cost.

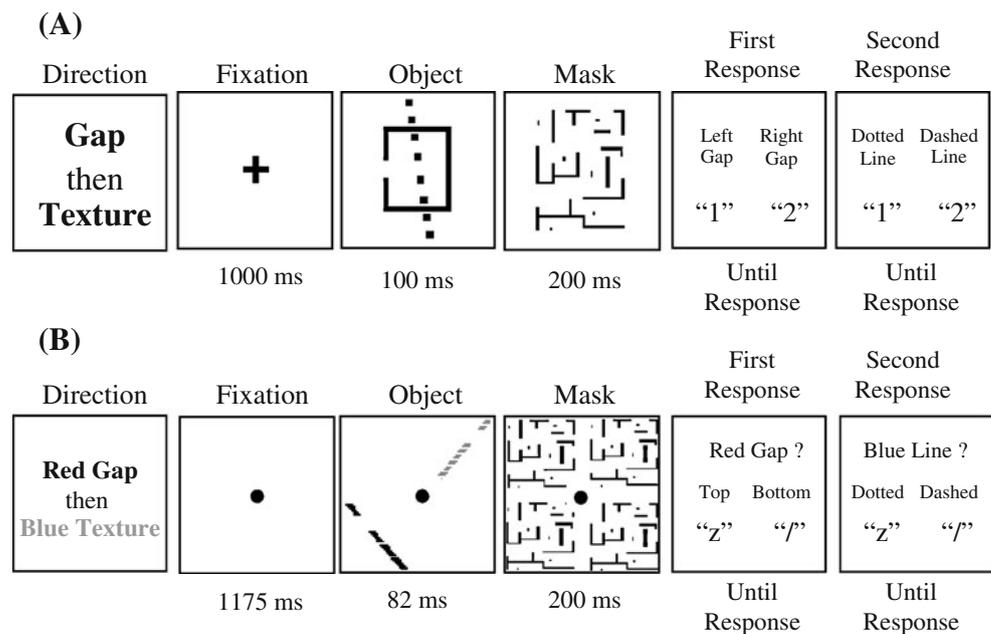
Following Duncan's (1984) results, Vecera and Farah (1994) suggested that there were two possible forms of object-based selection. Objects could be selected from a relatively early grouped-array representation, which allows locations to be selected by spatial attention (Vecera, 1994; Vecera & Farah, 1994). In a grouped array, objects are represented as a collection of features (e.g., edges) that perceptually group together based on gestalt principles such as proximity, connectedness, and good continuation. In this

---

M. Matsukura (✉) · S. P. Vecera  
Department of Psychology, University of Iowa,  
E11 Seashore Hall,  
Iowa City, IA 52242-1407, USA  
e-mail: michi-matsukura@uiowa.edu

S. P. Vecera  
e-mail: shaun-vecera@uiowa.edu

**Fig. 1 a** Event sequence for a typical feature-report trial (Duncan, 1984). **b** Event sequence for a variant of the feature-report trial used in Awh et al. (2001). Stimuli are not to scale. For this and all subsequent figures, different shades of gray are used to represent different colors



grouped-array selection, the object-based effect of attention arises because locations within an object are strongly grouped together, while locations between objects are weakly grouped (e.g., Egly et al., 1994; Vecera, 1994). Objects could also be selected from a relatively late selection stage, which represents the shape of an object and combines the features of that object, not the location occupied by that object. Selection from such an object-centered representation is independent of that object's spatial information, consistent with the idea of translation invariance from the object recognition literature (e.g., Biederman, 1987; Biederman & Gerhardstein, 1993, 1995; Marr, 1982). Thus, selection from this type of representation has been called *spatially-invariant* object-based attention (Vecera, 1997; Vecera & Farah, 1994; but see Kramer, Weber, & Watson, 1997).

The current experiments investigate the nature of this late, spatially-invariant object-based selection. We ask if observers' top-down expectancy (i.e., advance knowledge of to-be-reported features) is required for object-based selection or if the bottom-up perceptual inputs alone are sufficient to establish object representations. There are relevant findings in the literature. For example, Awh, Dhaliwal, Christensen, and Matsukura (2001) demonstrated that spatially-invariant selection could operate in the absence of advance knowledge of to-be-reported features. As depicted in Fig. 1b, Awh et al. (2001) presented a red line and a blue line separately on both sides of the fixation. The observers' task was a variant of Duncan's (1984) feature-report task: observers were asked to discriminate two features of these lines, gap and texture. A gap was located either on the top or the bottom of the line. Texture indicated whether the line was composed of either dots or dashes. In Experiment 1, observers were told two to-be-reported features of colored lines at the beginning of each

block (e.g., “Report gap of the red line first, and then texture of the blue line.”). In Experiment 2, while keeping the stimuli and task of Experiment 1 constant, the cuing information of to-be-reported features was withheld until after two lines were presented and masked. Both experiments replicated the key results of Duncan (1984): Observers reported two features that belonged to a single line more accurately than each feature of each line.

When observers are asked to identify to-be-reported features after objects are presented and masked as in Awh et al. (2001), attentional selection must occur from a relatively late representation, such as visual short-term memory (VSTM; see Matsukura & Vecera, 2009). For this reason, Awh et al.'s finding offers some of the strongest evidence to support the existence of spatially-invariant selection of object-based attention to date.

At a first glance, Awh et al. (2001) study appears to suggest that observers were able to quickly select objects without advance knowledge of to-be-reported features, suggesting that bottom-up perceptual inputs were sufficient to form object representations that attention selects. However, one overlooked aspect of Awh et al.'s (2001) procedure is that the to-be-reported object type (the same-object condition vs the different-objects condition) in Experiment 2 was *blocked*, while the observers were not explicitly told about this blocked design. In other words, observers reported two attributes from a particular object type (i.e., either from the same object or two different objects) for 32 consecutive trials, despite specific to-be-reported features and stimuli for every trial. When trials were implicitly blocked such as this, another type of advance knowledge might influence object-based attention. Specifically, attentional selection on the previous

trial might affect selection on the current trial. For example, the scale of attention on the current trial can be configured based on the preceding trial (Theeuwes, Kramer, & Belopolsky, 2004). In the same-object blocks, attention might be set to a narrower focus than in the different-objects blocks. This difference in the attention scale might have produced higher accuracy in the same-object condition than in the different-objects condition, even though specific to-be-reported features vary from trial to trial.

The goal of the current experiments was to examine the nature of spatially-invariant object-based selection under the condition of maximal uncertainty; that is, the condition that neither object type (the same-object condition vs the different-objects condition) nor the order of reporting two features is known before the target objects are presented. If the object-based effect of attention is observed under such maximal uncertainty, then it suggests that attention selects objects irrespective of an observer's top-down goal setting. In contrast, if the object-based effect is eliminated under the maximal uncertainty condition, then it suggests that object-based selection may be closely linked to an observer's attentional set.

Using Duncan's (1984) feature-report task, observers identified two features either from a single object or from two different objects, while to-be-reported features were withheld until after objects were presented and masked. Critically, the same-object and different-objects trials were randomly intermixed in order to prevent observers from predicting a pair of to-be-reported features of a upcoming trial based on preceding trials and from setting a particular attention scale based on preceding trials within the same block.

To preview our results, we found that the object-based effect of attention was abolished in the maximally uncertain condition (Part I). Next, in the second sets of experiments (Part II), we searched for a critical factor that drove attention to select objects in the absence of observers' foreknowledge of to-be-reported features, as demonstrated in Awh et al. (2001). Our results suggest that object-based selection can operate in the face of uncertainty when these objects are highly discriminable and perceptually distinct from each other.

## Part I: Effects of advance knowledge

### Experiment 1

To examine whether attention selects objects purely from a spatially-invariant representation, we simply randomized pairs of to-be-reported features while withholding the cuing information of two to-be-reported features until after the objects were masked. Complete randomization of trial types prevented observers from predicting a pair of to-be-reported

features for an upcoming trial and from setting a specific attentional scale within a particular object-type block. In *structured trials* (Experiment 1), the two to-be-reported features were blocked (unknown to the observers) and observers did not know what features to report until after the objects were presented and masked. This structured-trials design allowed us to replicate Awh et al.'s findings with our stimuli (Duncan, 1984) and procedure. In *mixed trials* (Experiment 2), the two to-be-reported features were randomly selected and observers did not know what features to report until after the objects were presented and masked.

For both structured and mixed trials, observers were asked to report features from a single object (the *same-object* condition) on half of the trials and features from two different objects on the other half of the trials (the *different-objects* condition). The object-based effect of attention was defined as higher accuracy of the same-object trials than that of the different-objects trials.

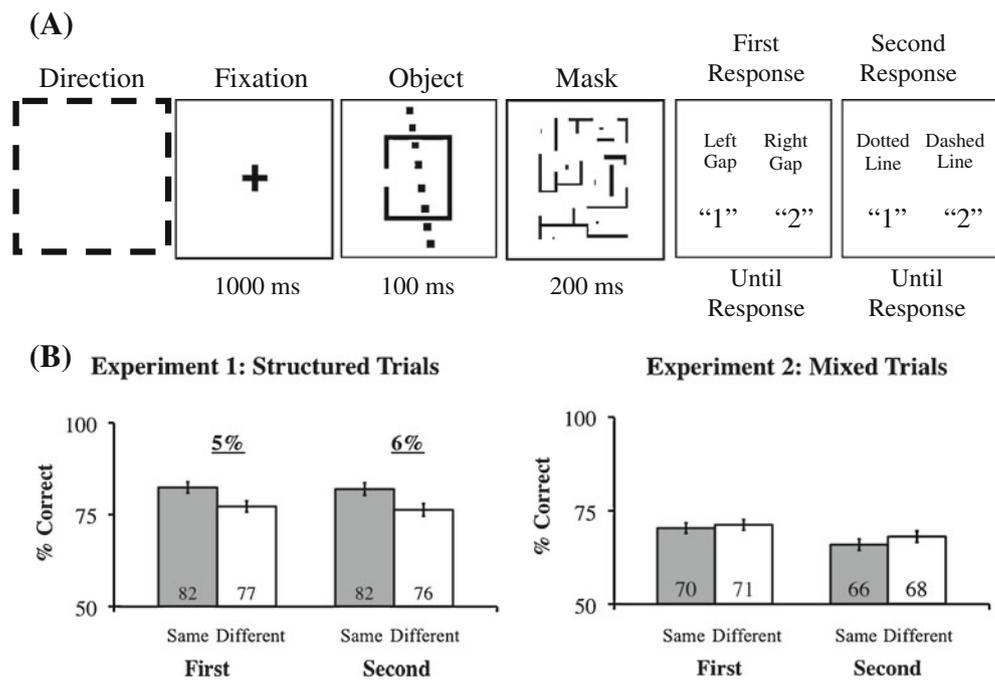
### Method

**Participants** Observers were 16 University of Iowa undergraduate students who received partial course credit for their involvement (age 18–30); all reported having normal or corrected-to-normal vision. None of these observers participated in any of the other experiments.

**Stimuli** Stimuli were similar to those used by Duncan (1984), shown in Fig. 2a. Stimuli consisted of two objects, a box and a line. These stimuli were viewed from a distance of 60 cm and presented on a 17-inch color monitor with a black background. Each object had two dimensions. The box was either short or tall and had a gap either on the left or right. The line was either dotted or dashed and tilted either to the left or right. These two objects were superimposed. The width of each box was  $0.67^\circ$  of visual angle. The height was  $1.14^\circ$  for the tall box and  $0.86^\circ$  for the short box. The box's gap was centered on either the left or the right of the box. Six pixels ( $0.20^\circ$ ) were removed from the side of the box to create this gap. The length of each line was  $1.53^\circ$  and was tilted  $8^\circ$  clockwise (rightward) or counterclockwise (leftward) from the vertical position. The line was also either dotted or dashed; both a dot and a dash contained 4 pixels but their configuration differed. A dot was 2 pixels in width and 2 pixels in height whereas a dash was 1 pixel in width and 4 pixels in height. The pattern mask was  $2.1^\circ$  in width and  $1.62^\circ$  in height.

**Procedure** Each trial began with an observer beginning an articulatory suppression task, in which the observer was required to repeat either "A, B, C, D" or "1, 2, 3, 4" aloud through the duration of the trial. This task effectively

**Fig. 2 a** Event sequence for a feature-report trial in Experiments 1 and 2. **b** Feature-report accuracy as a function of object type (same object vs different objects) and response order (first response vs second response) in Experiments 1 and 2. For this and all subsequent figures, *error bars* represent 95% within-subjects confidence intervals (Loftus & Masson, 1994). *Underlined* percentage represents the size of the object-based effect



discourages verbal recoding of visual information (Baddeley, 1986; Besner, Davies, & Daniels, 1981; Murray, 1968). Observers were instructed to speak at a rate of 3–4 digits or letters per second, and the experimenter continuously monitored the observers to ensure adequate performance.<sup>1</sup>

The observers' task was to discriminate two features of two objects. Each trial display began with a fixation cross that was presented for 1,000 ms. A box and a line appeared for 100 ms and were replaced by the patterned mask. The mask remained on the screen for 200 ms. The mask was replaced by the first feature-report question, for example, "Left Gap or Right Gap?" The observers made an unspeeded manual response to indicate whether the gap was located on the left or right side of the box using the keyboard. After reporting the first feature, the second feature-report question appeared. After reporting the second feature in the same manner with the first feature, the fixation cross appeared for the next trial.

Each observer participated in a single experimental session. Although the observers were not explicitly told

about the trial structure, the trials were blocked by pairs of to-be-reported features, and each observer reported all possible pairings in all possible orders. Thus, four features from two objects created 4 same-object pairs of to-be-reported features (gap/height, height/gap, texture/tilt, tilt/texture) and 8 different-objects pairs of to-be-reported features (gap/texture, texture/gap, gap/tilt, tilt/gap, height/texture, texture/height, height/tilt, tilt/height). Because the same-object condition had only 4 pairs, the observers performed the same-object pairs twice to equate the number of the same-object and different-objects trials. As a result, there were 8 blocks of the same-object trials and 8 blocks of the different-objects trials. Same-object and different-objects blocks were alternated, and the starting block was counterbalanced across the observers. The observers performed 256 trials in four 64-trial blocks. The observers were allowed to take a break between these blocks. Independent of this 4-block structure, these 256 trials were blocked by two to-be-reported features; i.e., the observers reported the identical pair of to-be-reported features for 16 consecutive trials.

### Results and discussion

The left graph in Fig. 2b shows accuracy (percent correct) of the feature-report task as a function of object type (same object vs different objects) and response order (first vs second). Replicating Awh et al. (2001) results, the observers reported two features from a single object more

<sup>1</sup> We conducted a number of feature-report tasks, including replication of Awh et al. (2001), *with and without* articulator suppression. Both for box-line (Duncan, 1984) and two-line stimuli (Awh et al., 2001), observers' feature-report performance was not affected by the use of articulatory suppression in terms of accuracy and pattern of the data. Regardless of whether articulatory suppression was performed or not, (1) the robust object-based effects were observed as long as to-be-reported features were known before objects were presented, and (2) no object-based effect was observed when the observers did not know two to-be-reported features in advance. Performance accuracy remained within the same range across *with and without* articulatory suppression experiments.

accurately than from two different objects, and this typical object-based effect was present for both the first-response and second-response trials. These observations are supported by an analysis of variance (ANOVA) with within-subjects factors of object type and response order. Greater accuracy for the same-object trials than for the different-objects trials led to a significant main effect of object type,  $F(1, 15) = 15.4$ ,  $p < .001$ . There was no main effect of response order,  $F(1, 15) = 0.9$ ,  $p = .4$ , indicating that accuracies for the first-response and second-response trials were approximately the same. Finally, the two-way interaction of object type and response order was not significant,  $F(1, 15) = 0.2$ ,  $p = .7$ , which indicates that the magnitude of the object-based effect for the first-response trials (5%) and the second-response trials (6%) did not significantly differ from one another. Planned pair-wise comparisons indicated that the observed object-based effect was significant for both the first-response trials,  $t(15) = 3.7$ ,  $p < .002$ , and the second-response trials,  $t(15) = 3.5$ ,  $p < .0003$ .

The results of Experiment 1 replicated Awh et al. (2001) results of the structured-trials experiment: the observers exhibited significant object-based effects for both the first-response and second-response trials.

One minor difference between our findings and those from Awh et al. (2001) is that we found an object-based effect for both the first-response and second-response trials; Awh et al. (2001) found the effect only for the second-response trials (as in Duncan, 1984). This is likely to derive from the scale of attention effect mentioned earlier (Theeuwes et al., 2004). During a different-objects run, there is greater uncertainty about the second to-be-reported feature, and this uncertainty could affect the first response on the subsequent trial. Regardless of this finding, Experiment 1 produced an overall object-based effect. Having replicated the object-based effect with structured trials, we now examine whether object-based effects will be observed when both to-be-reported features are known before the objects are presented.

## Experiment 2

In Experiment 2, the stimuli and task of Experiment 1 were kept constant, but object type and response order were randomly presented. If attention selects objects without observers' advance knowledge of to-be-reported features, then the object-based effect should be observed. By contrast, if attention does not select objects in the absence of the advance knowledge, then the object-based effect should disappear.

## Method

The method of Experiment 2 was the same as that of Experiment 1 except that the trials were not blocked by

pairs of to-be-reported features. Thus, object type and response order were completely randomized throughout the experiment (i.e., the mixed-trials design).

## Results and discussion

The right graph in Fig. 2b shows accuracy of the feature-report task as a function of object type and response order. Surprisingly, all object-based effects were abolished. Observers did not show higher accuracy in the same-object condition than in the different-objects condition for either the first-response or second-response trials. Overall accuracy was lower than that observed in Experiment 1, and accuracy for the first-response trials was higher than for the second-response trials. These observations were supported by an ANOVA with within-subjects factors of object type and response type. Neither the main effect of object type nor the interaction of object type and response order was significant,  $F(1, 15) = 3.3$ ,  $p = .1$ ,  $F(1, 15) = 0.4$ ,  $p = .6$ , respectively. The main effect of response type was significant,  $F(1, 15) = 12.2$ ,  $p < .0003$ , with higher accuracy for the first-response trials than for the second-response trials.

Admittedly, the results of Experiment 2 were a surprise. When the to-be-reported features were post-cued, object-based effects disappeared, which indicates that attention did not select objects. Thus, object-based selection from spatially-invariant representation may require observers' advance knowledge of to-be-reported features before the objects are presented.

Because object-based effects disappeared in the mixed-trials design, we were motivated to understand how object-based effects of attention were produced when the objects were not perceptually available in Awh et al. (2001). In a series of follow-up experiments that used the same procedure as Experiment 2, we first tested whether the absence of an object-based effect was due to a *data-limited* condition (Norman & Bobrow, 1975) in the mixed-trials design. Specifically, randomization of object type (same-object vs different-objects) coupled with a brief exposure duration (100 ms) might have made too little information available for observers to deploy attention to objects. Although longer exposure duration (200 ms) increased observers' overall accuracy 10% higher than that of Experiment 2, the object-based effect remained absent,  $F(1, 15) = 0.5$ ,  $p = .5$ .

Another possible cause for the absence of the object-based effect in Experiment 2 is that there might not have been enough time for visual items to be consolidated into VSTM when observers could not direct their attention to objects that possess to-be-reported features. To examine this possibility, while keeping the mixed-trials design constant, we increased

the duration for memory consolidation. Similar to a typical change-detection task (Luck & Vogel, 1997; Vogel, Woodman, & Luck, 2001), the exposure duration of the box-line stimuli was 100 ms, followed by a delay of 900 ms without any mask before the onset of the first feature question display. This sequence creates 1,000-ms consolidation time between the onset of the box-line stimuli and the onset of the first feature-report question. Based on the estimates from Vogel, Woodman and Luck (2006), this 1,000-ms duration should be sufficient to consolidate perceptual object representations into VSTM. However, although overall accuracy was raised to approximately 86%, the object-based effect remained absent,  $F(1, 15) = 1.0, p = .3$

Having ruled out exposure duration and VSTM consolidation time as the source of the absent object-based effect, it is possible that object-based selection ceases to operate when to-be-reported features vary from trial to trial, regardless of whether the trials are presented in the structured or mixed design. Observers may be unable to select objects when to-be-reported features and the objects to which these features belong change on every trial. To test this possibility, in the next two experiments, we pre-cued the to-be-reported features on each trial while keeping the mixed-trials design.

### Experiment 3

In Experiment 3, we kept the timing and mixed-trials design used in Experiment 2 constant; however, we cued both to-be-reported features at the beginning of each trial. In essence, we examined whether spatially-invariant object-based selection would fail whenever to-be-reported features vary from trial to trial. The design of Experiment 3 also allowed us to test whether the attention scale setting based on the structured-trials runs (Theeuwes et al., 2004) contributes to the object-based effect. By pre-cuing the to-be-reported features in the mixed-trials design, we could measure the effect of observers' expectancy on the object-based effect independently from the effect of potential attention scale setting.

### Method

The method of Experiment 3 was the same as that of Experiment 2 except that the trials were not blocked by pairs of to-be-reported features but the two to-be-reported features were cued at the beginning of each trial. At the beginning of each trial, the information of the to-be-reported features was displayed (e.g., "Report GAP then TEXTURE"). Observers were instructed to press any button to start a trial after reading this direction. As soon as the observers press any button, a fixation cross was presented for 1,000 ms, and the same event sequence with Experiment 2 followed. Thus, object type and response

order were completely randomized throughout the experiment (i.e., the mixed-trials design).

### Results and discussion

As shown in the left graph of Fig. 3b, there was a significant main effect of object type,  $F(1, 15) = 31.0, p < .0001$ , with higher accuracy for the same-object trials than for the different-objects trials. Higher accuracy for the first-response trials than for the second-response trials also led to the main effect of response order,  $F(1, 15) = 34.4, p < .0001$ . Finally, the interaction between object type and response order was also significant,  $F(1, 15) = 20.9, p < .0001$ , because the object-based effect was present only for the second-response trials but not for the first-response trials. Planned pair-wise comparisons indicated that the observed object-based effect was significant for the second-response trials,  $t(15) = 5.6, p < .0001$ , but not for the first-response trials,  $t(15) = 1.4, p = .2$ . These results suggest that object-based selection from a spatially-invariant representation can operate when features are cued in advance of stimuli presentation, even though the to-be-reported features vary from trial to trial.

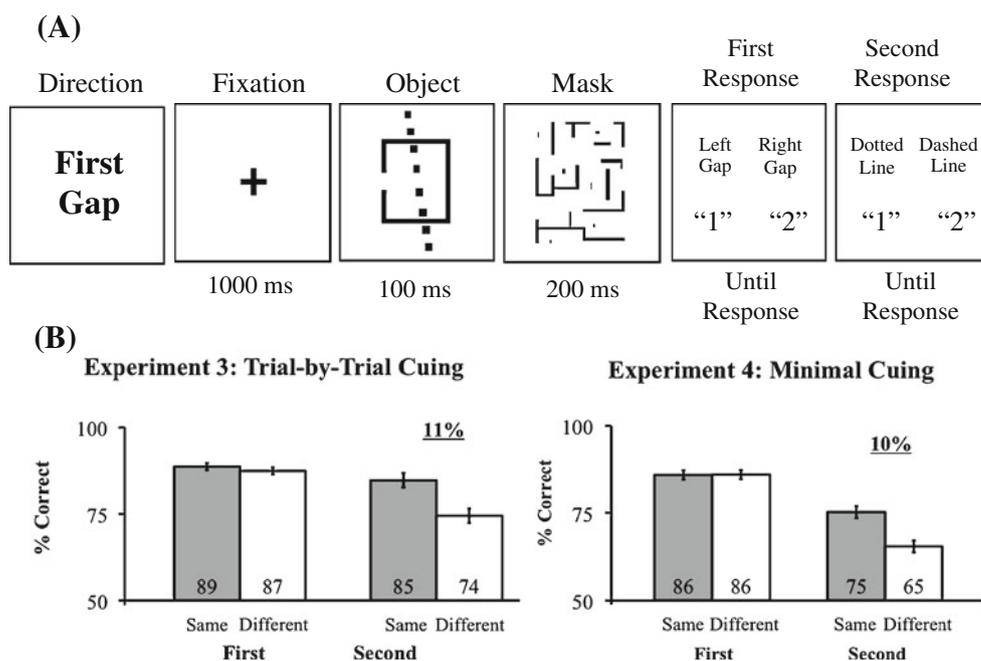
### Experiment 4

Although it appears that object-based attention can operate on a trial-by-trial basis when both to-be-reported features are pre-cued, it remains unclear whether advance knowledge of both to-be-reported features are necessary to reduce uncertainty regarding which objects attention should be directed to. It is possible that cuing of a single to-be-reported feature is sufficient to select an object; when one feature of the target object is selected, the other feature of that object should be also selected. To test this hypothesis of the object-based selection mechanism, in Experiment 4, the observers were presented with the pre-cuing information of the first to-be-reported feature only (Fig. 3a). The feature for the second response was chosen randomly from the remaining three features. If the observers are able to select a visual item with two features as a single object, then the object-based effect should remain strong when the minimal cuing information about the object is provided at the beginning of each trial.

### Method

The method of Experiment 4 was the same as that of Experiment 3. While the trials were not blocked by pairs of to-be-reported features, only one to-be-reported feature was cued at the beginning of each trial (Fig. 3a). Thus, object type and response order were completely randomized throughout the experiment (i.e., the mixed-trials design).

**Fig. 3 a** Event sequence of a feature-report trial in Experiment 4 (with a minimal pre-cue in the mixed-trials design). **b** Feature-report accuracy as a function of object type and response order in Experiments 3 and 4



*Results and discussion*

As shown in the right graph of Fig. 3b, the results demonstrated that the observers reported two features from a single object more accurately than from two different objects. An ANOVA with within-subjects factors of object type and response order revealed that higher accuracy for the same-object trials than for the different-objects trials led to a significant main effect of object type,  $F(1, 15) = 23.4, p < .0002$ . Higher accuracy for the first-response trials than for the second-response trials also led to a significant effect of response order,  $F(1, 15) = 143.1, p < .0001$ . Finally, the two-way interaction of object type and response order was also significant,  $F(1, 15) = 26.0, p < .0001$ , because the object-based effect was present only for the second-response trials but not for the first-response trials. Pair-wise comparisons confirmed that the object-based effect was significant for the second-response trials,  $t(15) = 6.3, p < .0001$ , but not for the first-response trials,  $t(15) = 0.1, p = .9$ . Thus, object-based selection operates even when the minimal information of a target object was provided, as long as the observers had a chance to restrict visual processing to a single target object.

**Interim discussion**

The results of Experiments 1–4 demonstrate that attention does not appear to select objects unless observers have advance knowledge of the objects or features to report. Such results suggest the possibility that observers’ expectancy for upcoming to-be-reported features contribute to the operation of object-based attention. In other words, at the time of the

object presentation, some knowledge of the relevant features is required to observe object-based selection.

Another potential source of uncertainty may come from the stimuli themselves. Duncan’s (1984) overlapping box and line stimuli contain relatively few perceptual cues to segregate themselves as separate objects. In contrast, line stimuli used in Awh et al. (2001) had clearer segregation, which allowed these two lines to be easily perceived as two independent entities. Specifically, the lines occupied two separate locations instead of overlapping, and they possessed distinctively different colors from one another. In addition, the lines ( $2.1^\circ$ ) were larger than the tall box ( $1.14^\circ$ ) and longer than the tilted line ( $1.53^\circ$ ), which may also have enhanced the overall stimulus perceptibility. Accumulation of these stimuli differences suggest that observers’ top-down expectancy might be required for relatively ambiguous stimuli, such as those used in Experiments 1–4, to drive object-based selection.<sup>2</sup> At the same time, observers’ top-down expectancy might not be required to generate object-based effects when the objects are readily distinguishable and easy-to-segregate from one another. We tested these possibilities in Experiments 5 and 6.<sup>3</sup> To

<sup>2</sup> A single perceptual cue (larger size, different colors, separate locations) alone did not make the object-based effect return,  $ps > .05$ .

<sup>3</sup> In Experiments 5 and 6, we presented different-colored objects on separate locations to closely follow the properties of the two line stimuli in Awh et al. (2001). Alternatively, as one reviewer suggested, we could have used transparent motions with different directions in place of having two objects on separate locations (e.g., Valdes-Sosa, Cobo, & Pinilla, 1998). If transparent motion cues are as efficient as the separate location cue to segregate the box and the line, object-based effects should be observed.

anticipate the results, object-based effects were observed without any advance knowledge of to-be-reported features when the objects were perceptually distinct from each other. These results allow us to integrate our findings from Experiments 1–4 with Awh et al. (2001) results.

## Part II: Effects of perceptual segregation

### Experiment 5

To determine if observers' top-down expectancy is required to observe object-based effects with perceptually distinct objects, we added several perceptual cues to the box-line stimuli to increase the probability that the stimuli would be perceived as two objects separate from one another. While keeping the mixed-trials design constant, the box and line were twice as large as those used above to make features more perceptible; two objects had different colors from one another and located on opposite sides of the fixation (Fig. 4a). In addition to combining these three perceptual cues (larger size, different colors, separate locations), the objects were presented for 200 ms to enhance stimuli perceptibility in general. Critically, these changes aid to segregate two objects, and they are task-irrelevant features to perform the feature-report task. If the increased segregability provides sufficient bottom-up information to facilitate object-based selection, then an object-based effect should return, as reported in Awh et al. (2001).

### Method

The method was the same as that of Experiment 2, except now the stimuli were twice as large as those used in the previous experiments. The width of each box was  $1.81^\circ$  of visual angle. The height was  $3.05^\circ$  for the tall box and  $2.77^\circ$  for the short box. Twelve pixels ( $0.48^\circ$ ) were removed from the side of the box to create this gap. The length of each line was  $1.53^\circ$ . The pattern mask was adjusted to  $3.15^\circ$  in width and  $4.76^\circ$  in height. Also, unlike previous experiments, the box and line had different colors: one appeared in blue while the other appeared in yellow. Two objects never appeared in the same color at the same time. The RGB values for the colors were as follows: blue = 51, 255, 255; yellow = 255, 255, 0. Finally, each object and mask appeared at  $2.62^\circ$  to the left and right of the fixation, and the exposure duration was 200 ms.

### Results and discussion

As shown in Fig. 4b, there was a significant main effect of object type,  $F(1, 15) = 6.7$ ,  $p < .02$ , with higher

accuracy for the same-object trials than for the different-objects trials. Higher accuracy for the first-response trials than for the second-response trials also led to the main effect of response order,  $F(1, 15) = 36.4$ ,  $p < .0001$ . Finally, the larger object-based effect for the second-response trials than for the first-response trials produced a significant interaction of object type and response order,  $F(1, 15) = 8.3$ ,  $p < .01$ . Planned pair-wise comparisons indicated that the observed object-based effect was significant for the second-response trials,  $t(15) = 3.7$ ,  $p < .002$ , but not for the first-response trials,  $t(15) = 0.04$ ,  $p = 1.0$ .

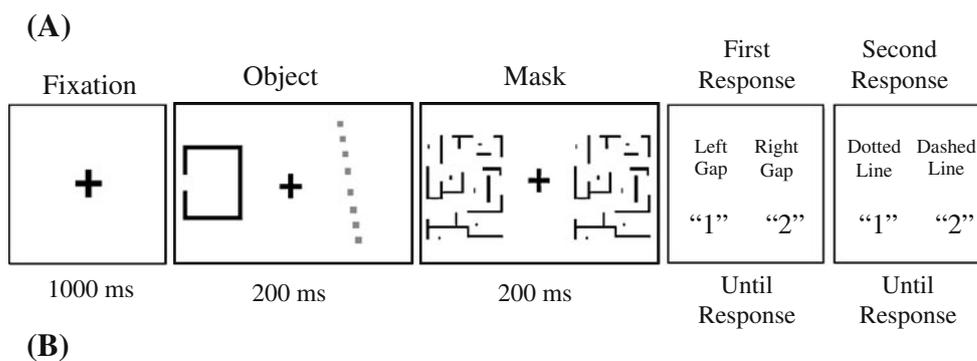
The results of Experiment 5 suggest that attention can select objects without advance knowledge. The critical factor, however, appears to be the segregation of the objects. When observers do not have any foreknowledge of to-be-reported features, it is necessary for objects to have strong perceptual cues that enable observers to perceive two objects as separate entities that are independent from one another.

One potential concern with Experiment 5 is that the separate locations of the objects, which are necessary for the two objects to easily segregate, might have allowed the observers to select objects based on their locations, not their object feature properties. It is critical to confirm that the object-based effect observed in Experiment 5 actually arose from the same format of object-based selection operated in Experiments 1–4. To test that attention indeed selected objects based on their feature properties but not on the locations that these objects occupied, in Experiment 6, we manipulated the distance between the objects (e.g., see Vecera, 1994; Vecera & Farah, 1994, for the use of this method).

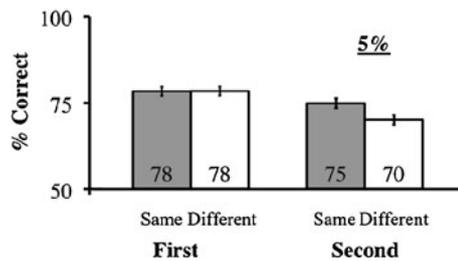
### Experiment 6

The purpose of Experiment 6 was to replicate the results observed in Experiment 5 even when the distance between two objects differed (Fig. 5). For half of the trials, a box and a line were presented in the exactly same manner as in Experiment 5 (*far trials*). For the other half of the trials, a box and a line were presented closer to the fixation (*near trials*). If the object-based effect observed in Experiment 5 was produced because two locations occupied by separate box and line were selected, then one would expect to see a decrease in the size of the object-based effect in the near condition, compared to the far condition. However, if the object-based effect observed in Experiment 5 arose from selecting an object's features from a spatially-invariant representation, then one would predict no difference in the size of the object-based effect between the near and far conditions (a spatially-invariant selection account); that is, when an object's

**Fig. 4 a** Event sequence of a feature-report trial in Experiment 5 (with a group of perceptual cues in the mixed-trials design). **b** Feature-report accuracy as a function of object type and response order in Experiment 5



**Experiment 5: Perceptual Cues**



features are selected, the spatial information of this object is irrelevant.

*Method*

The method of Experiment 6 was identical to that of Experiment 5, with the following exceptions. (1) Thirty-two new participants served as observers.<sup>4</sup> (2) For the half of the trials, a box and a line were presented closer to the fixation (near trials). In the near condition, each object and mask appeared at 1.24° to the left and right of the fixation. The other half of the trials was the replication of Experiment 5 (far trials) (Fig. 5a). (3) While keeping the mixed-trials design constant, distance type (near vs far) was randomly presented.

*Results and discussion*

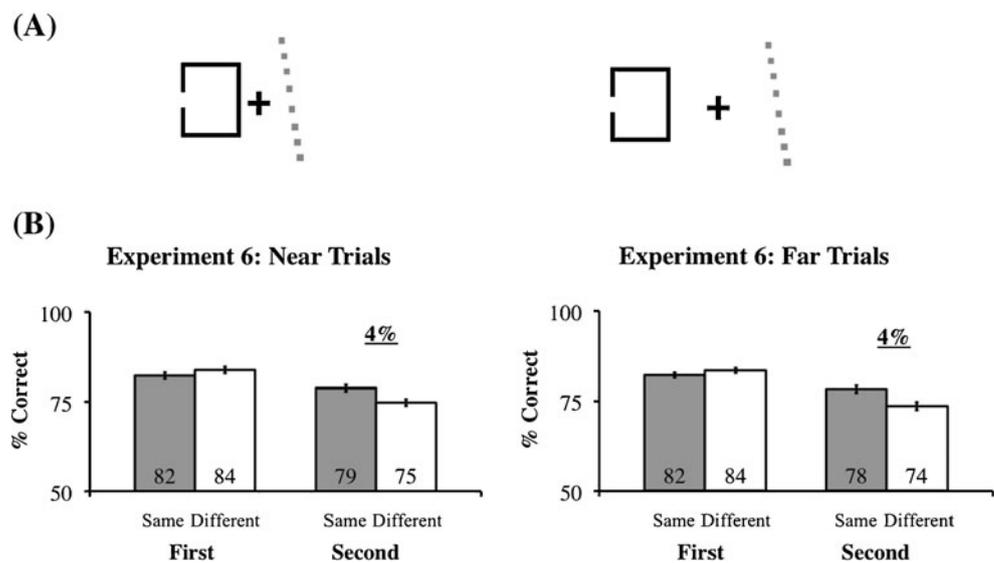
Figure 5b shows accuracy of the feature-report task as a function of object type and response order. The left graph depicts accuracy in the near condition whereas the right graph illustrates accuracy in the far condition. Significant object-based effects were observed in both the near and far conditions. Furthermore, as the spatially-invariant representation account predicted, the size of the object-based effect

did not differ between the near and far conditions. These observations were supported by a three-way ANOVA with within-subjects factors of object type, response order, and distance type (near vs far). Higher accuracy for the same-object trials than for the different-objects trials led to a significant main effect of object type,  $F(1, 31) = 10.6, p < .003$ . Higher accuracy for the first-response trials than for the second-response trials also led to a significant main effect of response order,  $F(1, 31) = 100.1, p < .0001$ . However, accuracy for the near trials did not significantly differ from that for the far trials,  $F(1, 31) = 1.0, p = .3$ . Replicating the results of Experiment 5, a significant object-based effect was observed for the second-response trials but not for the first-response trials. This pattern of the data produced a significant two-way interaction of object type and response order,  $F(1, 31) = 26.1, p < .0001$ . Finally, the size of the object-based effect observed in the near condition did not differ from that in the far condition, which led to a failure to produce a significant three-way interaction of object type, response order, and distance type,  $F(1, 31) = 0.14, p = .7$ . Pair-wise comparisons confirmed that the object-based effect observed for the second-response trials was significant in both the near condition (4%),  $t(31) = 4.5, p < .0001$ , and the far condition (4%),  $t(31) = 4.3, p < .0001$ .

One possible concern is whether visual acuity difference in the two distance conditions (due to different retinal eccentricities) would contribute to lower accuracy for the different-objects trials in the far condition, which would also lead to the larger object-based effect in the far condition than in the near condition. As reported earlier,

<sup>4</sup> The main results in Experiment 6 were significant with 16 or fewer subjects. We collected the data from 32 observers to counterbalance response orders and equate the number of the trials categorized by object type and distance type.

**Fig. 5** **a** Example stimuli of near and far trials in Experiment 6. **b** Feature-report accuracy as a function of object type and response order in Experiment 6



while no main effect of distance type was found, the magnitude of the object-based effect observed for the near and far trials is almost identical, suggesting that eccentricity has little effect across the two distance types.

The results of Experiment 6 indicate that the presence of the near trials did not affect the format of object-based selection for the far trials. To provide statistical support for this observation, we conducted a three-way ANOVA with a between-subjects factor of experiment type (Experiments 5 and 6) and within-subjects factors of object type and response order on Experiment 5 trials and far trials of Experiment 6. Although the two-way interaction of object type and response order was significant,  $F(1, 46) = 23.8, p < .0001$ , as observed in the previous experiments (except for Experiment 2), the three-way interaction of object type, response type, and experiment type was not significant,  $F(1, 46) = .3, p = .6$ . Lack of the three-way interaction was driven by an equivalent size of the object-based effect for the second-response trials observed in both Experiments 5 and 6. Thus, we can safely conclude that the spatially-invariant format of object-based attention for the far trials was not affected by the presence of the near trials in Experiment 6.

## General discussion

In the six experiments, we investigated how attention selects objects stored in a late, spatially-invariant representation. Experiments 1–4 (Part I) demonstrated that, when superimposed box and line can be seen as an ambiguous single object, advance knowledge of to-be-reported features is necessary for object-based selection. Foreknowledge of to-be-reported features enables observers to preferentially direct attention to the target objects that the to-be-reported

features belong to at the time of stimulus presentation. In Experiments 5 and 6 (Part II), we searched for a critical factor that allowed attention to select objects from a spatially-invariant representation in the absence of observers' foreknowledge of to-be-reported features. We demonstrated that strong perceptual cues to establish each object as an independent entity (i.e., enhanced perceptibility of visual items) enables object-based selection to return without pre-cuing the to-be-reported features. Therefore, attention can select objects purely from a spatially-invariant representation without any top-down assistance, provided that objects are easily segregated from one another.

The current results also argue for a wide range of control that spatially-invariant object-based attention can exert. That is, a spatially-invariant form of object-based attention can extend its operation to the conditions of maximal uncertainty. Attention selects objects from a spatially-invariant representation without being preferentially directed to objects that to-be-reported features are part of (Experiments 5 and 6) or without setting a particular attention scale around a single object across successive trials (Experiments 3). Moreover, in order for objects to be selected, it is not necessary for both of the to-be-reported features to be cued before the target objects are presented. The results of Experiment 4 suggest that, when one feature of the target object is selected, the other feature of that object is also selected. Given the only difference between Experiments 3 and 4 is whether the second to-be-reported feature was pre-cued or not, higher accuracy of the second-response trials in Experiment 3 than that of Experiment 4 must have been caused by the number of pre-cued features. Specifically, even though pre-cuing a single feature is sufficient to facilitate object-based selection, pre-cuing the two features leads to higher accuracy than cuing a single to-be-reported feature.

We would also like to note that, despite a recent report that spatially-invariant object-based selection operates within observers' VSTM representations (Matsukura & Vecera, 2009), the goal-driven and stimulus-driven controls of visual attention examined in the current experiments are different from the feature-binding role of visual attention that is extensively studied in the context of maintenance of VSTM (e.g., Johnson, Hollingworth & Luck, 2008; but see Wheeler & Treisman, 2002). On the one hand, the bi-directional attention mechanisms examined in the current experiments were a goal-driven function of preferentially directing or restricting observers' visual processing to task-relevant features that are part of an object, and a stimulus-driven function of selecting objects purely from a spatially-invariant representation, without any advance knowledge of to-be-reported features. On the other hand, visual attention extensively studied in the context of VSTM maintenance refers to a function of binding features into cohesive object representations in memory. Accordingly, based on the current results, we cannot draw any conclusions about whether visual items are stored as a unit of objects or features in VSTM.

In summary, object-based attention from a late spatially-invariant representation operates in both goal-driven and stimulus-driven manners. For attention to select objects in the absence of advance knowledge of to-be-reported features, strong perceptual cues to segregate two objects are necessary. Clear separation of two objects is the key for attention to select one object over the other without foreknowledge of to-be-reported features.

**Author Notes** This research was made possible by a Sigma Xi Grants-in-Aid of Research Award to Michi Matsukura and National Science Foundation award 03-39171 to Shaun P. Vecera. We thank Steve Luck, Joshua Cosman, Lauren Hecht, Hyunkyu Lee, Jeremy Wolfe, Jun Saiki, and two anonymous reviewers for their helpful comments on earlier drafts, as well as Emily Muff, Shaunna Kulp, Kristine Nurse, Jean Marie Hartin, Jessica M. Johnson, and Julia Meyer for assistance with data collection.

## References

- Awh, E., Dhaliwal, H., Christensen, S., & Matsukura, M. (2001). Evidence for two components of object-based selection. *Psychological Science*, *12*, 329–334.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon.
- Besner, D., Davies, J., & Daniels, S. (1981). Reading for meaning: The effects of concurrent articulation. *The Quarterly Journal of Experimental Psychology*, *33A*, 415–437.
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–147.
- Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 1162–1182.
- Biederman, I., & Gerhardstein, P. C. (1995). Viewpoint-dependent mechanisms in visual object recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1506–1514.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, *113*, 501–517.
- 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*, 161–177.
- Johnson, J. S., Hollingworth, A., & Luck, S. J. (2008). The role of attention in the maintenance of feature bindings in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 41–55.
- Kramer, A. F., Weber, T. A., & Watson, S. E. (1997). Object-based attentional selection—grouped array or spatially invariant representations? Comments on Vecera and Farah (1994). *Journal of Experimental Psychology: General*, *126*, 3–13.
- Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*, 476–490.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281.
- Matsukura, M., & Vecera, S. P. (2009). Interference between object-based attention and object-based memory. *Psychonomic Bulletin & Review*, *16*, 529–536.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: Freeman.
- Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology*, *78*, 679–684.
- Norman, D. A., & Bobrow, D. B. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, *7*, 44–64.
- Theeuwes, J., Kramer, A. F., & Belopolsky, A. V. (2004). Attentional set interacts with perceptual load in visual search. *Psychonomic Bulletin & Review*, *11*, 697–702.
- Valdes-Sosa, M., Cobo, A., & Pinilla, T. (1998). Transparent motion and object-based attention. *Cognition*, *66*, B13–B23.
- Vecera, S. P. (1994). Grouped locations and object-based attention: Comment on Egly, Driver, and Rafal (1994). *Journal of Experimental Psychology: General*, *123*, 316–320.
- Vecera, S. P. (1997). Grouped arrays versus object-based representations: Reply to Kramer et al. (1997). *Journal of Experimental Psychology: General*, *126*, 14–18.
- Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General*, *123*, 146–160.
- Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 92–114.
- Vogel, E. K., Woodman, G. F., & Luck, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 1436–1451.
- Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131*, 48–64.