

Object-Based Attentional Selection—Grouped Arrays or Spatially Invariant Representations?: Comment on Vecera and Farah (1994)

Arthur F. Kramer and Timothy A. Weber
University of Illinois at Urbana–Champaign

Stephen E. Watson
Brooks Air Force Base

S. P. Vecera and M. J. Farah (1994) have addressed the issue of whether visual attention selects objects or locations. They obtained data that they interpreted as evidence for attentional selection of objects from an internal spatially invariant representation. A. F. Kramer, T. A. Weber, and S. E. Watson question this interpretation on both theoretical and empirical grounds. First, the authors suggest that there are other interpretations of the Vecera and Farah data that are consistent with location-mediated selection of objects. Second, they provide data, using the displays employed by Vecera and Farah in conjunction with a postdisplay probe technique, that suggests that attention is directed to the locations of the target objects. The implications of the results for space and object-based attentional selection are discussed.

Over the last decade there has been increasing empirical support for the notion of multiple modes of visual selection. The modal view in the past had been that visual attention selects information in the visual field much in the same way that a spotlight might be focused on an actor on a stage. That is, attention is directed to a specific area of the visual field and information that falls within the attentional focus is actively processed, whereas information outside of the focus is ignored.

In fact, a number of metaphors such as spotlights (Broadbent, 1982; Posner, Snyder, & Davidson, 1980), zoom lenses (Eriksen & St. James, 1986; Eriksen & Yeh, 1985), and gradients (Downing & Pinker, 1985; LaBerge & Brown, 1989) have been invoked to describe the manner in which attention is directed in the visual field. Models based on these metaphors differ on a number of characteristics, the most important of which is the flexibility with which attention can be focused on information in visual space. For example, spotlight models often assume a fixed diameter beam, whereas zoom-lens models suggest that attention can either be focused tightly or expanded to encompass much of the visual field. Despite these differences, the spotlight, zoom-lens, and gradient models all assume that attention is directed to specific areas of the visual environment and that

stimuli falling within the focus of attention are selected for further processing.

In contrast to the space-based selection models, there is an increasing body of data that suggests that attention can be directed to objects or perceptual groups that have been created by a preattentive segmentation of the visual field. Within such a framework, processes of grouping, parsing, and figure-ground segregation create candidate objects that can be individually addressed or selected for further processing and action. Kahneman and colleagues (Kahneman & Henik, 1981; Kahneman & Treisman, 1984; Kahneman, Treisman, & Gibbs, 1992; Treisman, Kahneman, & Burkell, 1983) have further suggested that focusing attention on a particular object or perceptual group results in the mandatory processing of all properties of that object. Thus, different properties of an object are processed in parallel, whereas different objects are processed serially.

The theoretical assumptions underlying the object-based selection proposal have implications for both focused and divided-attention tasks. In the case of a focused-attention task, performance should improve to the degree that any conflicting information can be located on different objects from the relevant information. On the other hand, performance on divided-attention tasks should be best when all of the information can be located on a single object.

Duncan (1984) reported a series of important studies that are consistent with the object-based proposal. In these studies participants were instructed to report either two properties from a single object (i.e., the direction of tilt and texture of a line, the height or side of a box on which a gap appeared) or one property from each of two objects (e.g., the direction of tilt of the line, the height of the box, etc.). Although the two objects were superimposed, performance was better when both properties appeared on a single object than when participants reported one property from one object and the other property from the second object. Furthermore, Duncan found that as long as the two properties were located on a single object, performance was equivalent

Arthur F. Kramer and Timothy A. Weber, Beckman Institute, University of Illinois at Urbana–Champaign; Stephen E. Watson, Armstrong Laboratory, Brooks Air Force Base, San Antonio, Texas.

This research was supported by Office of Naval Research Grant N00014-92-J-1792 and National Institutes of Health Grant RO1 AG12203.

We would like to thank Kyle Cave, Bill Johnston, and Shaun Vecera for their helpful comments on a draft of this article.

Correspondence concerning this article should be addressed to Arthur F. Kramer, Beckman Institute, University of Illinois, 405 North Mathews Avenue, Urbana, Illinois 61801. Electronic mail may be sent via Internet to akramer@s.psych.uiuc.edu.

to a control condition in which only a single property was judged (see also Baylis & Driver, 1993; Duncan, 1993; Kramer & Watson, 1995). Other researchers have found larger performance costs when an irrelevant and response-incompatible distractor was embedded in the target object or perceptual group than when the distractor was located in a different object (Baylis & Driver, 1992; Kramer & Jacobson, 1991).

An important issue concerning object-based selection, and the focus of this article, is the representation from which selection takes place. One suggestion, offered by Vecera and Farah (1994; see also Vecera, 1994), is that visual attention selects from an internal object-based representation. An important characteristic of such a proposal is that spatial location plays no role in the selection process because objects are selected from an internal representation in which they are coded in a spatially invariant fashion. Such a proposal is consistent with object-recognition theories that suggest that objects are represented in memory in an object-centered coordinate system in which object parts are specified relative to the main axis of the object (Marr, 1982; Marr & Nishihara, 1978) or that simple object parts are defined relative to other object parts (Lowe, 1987). Hereinafter, we refer to this suggestion as the *spatially invariant proposal*.

In contrast to the spatially invariant proposal, it is conceivable that attention activates those locations in a retinotopic or spatiotopic representation of the visual field that correspond to an object's shape. Thus, in this case, attention might precisely silhouette a specific object. This proposal, referred to as the *grouped-array model* by Vecera and Farah (1994), is similar to space-based attentional models in the sense that attention serves to activate particular regions of a representation of the visual field rather than spatially invariant object representations. However, the grouped-array proposal differs from space-based models in the sense that attention is, for all practical purposes, infinitely flexible with respect to its shape.

Vecera and Farah (1994) performed a series of studies in an effort to contrast the spatially invariant and grouped-array models of object-based attentional selection. They employed the same stimuli and similar tasks used by Duncan (1984). Participants judged two properties that could be located on either a single object or distributed between two different objects. Two different conditions were contrasted. In the *superimposed condition*, participants were presented with the overlapping box and line at fixation. In the *separated condition*, the box was presented on one side of fixation and the line was presented on the other side of fixation.

Vecera and Farah (1994) reasoned that if participants had selected the objects from a grouped array in Duncan's (1984) study then they would expect to find an increase in the size of the object effect (i.e., better performance when two properties were located on a single object than when the properties were distributed between the two objects) for the separate than for the superimposed condition. This expectation was based on the assumption that attention would take longer to move between the two objects in the sepa-

rated than in the superimposed condition, thereby resulting in poorer performance in the *different-object condition* (i.e., when the two properties were located on different objects) when the two objects were separated as compared to when they were overlapping. On the other hand, if participants selected the objects from an internal spatially invariant array, then the magnitude of the object effect should be equivalent in the superimposed and separated conditions. Their results were consistent with the spatially invariant proposal. The magnitude of the object effect was the same in the superimposed and separated conditions.

Vecera and Farah (1994) interpreted the statistically equivalent object effects in the superimposed and separated conditions as evidence in support of object-based selection from an internal spatially invariant representation. This interpretation is critically dependent upon their assumption that attention moves in an analog fashion through visual space. Indeed, a number of researchers have argued that attention movement occurs in this manner, with the time required to reorient attention increasing with the separation between two locations (Jolicoeur, Ullman, & MacKay, 1991; Remington & Pierce, 1984; Shulman, Remington, & McLean, 1979; Tsal, 1983). However, this assumption and the studies that purportedly support it has been criticized on both logical and methodological grounds (Eriksen & Murphy, 1987; Yantis, 1988). Furthermore, there now appears to be a substantial body of literature that suggests that, under many conditions, spatial attention appears to be reoriented from one location to another in a discrete manner, such that it does not take more time to move longer distances (Eriksen & Webb, 1989; Kwak, Dagenbach, & Egeth, 1991; Sagi & Julesz, 1985; Sperling & Weichselgartner, 1995). Of course, if attention moved in a discrete manner between the two objects in the Vecera and Farah study, then the lack of a performance difference between the separated and superimposed conditions would not be diagnostic with respect to whether attention was focused on the location of the objects or instead on a spatially invariant internal representation of the objects.¹

¹ There is at least one additional plausible account of the results obtained by Vecera and Farah (1994) that does not depend on the objects being selected from a spatially invariant internal representation. Given that the separated and superimposed conditions were randomly presented within the same block of trials in the Vecera and Farah studies, it is conceivable that participants initially distributed attention in a diffuse fashion across the three possible object locations (i.e., the objects could occur at fixation or to the right and left of fixation). When the display was presented and participants were to report two properties from a single object, participants may have constricted their focus of attention to the target object. Given that the location of the target object varied randomly over trials, a shift from a diffuse to a focused mode of attention (Eriksen & St. James, 1986; Jonides, 1983), which is a focus on the shape of the object, was expected to take the same amount of time regardless of whether the target object appeared at the locations employed on the separated or superimposed trials. In the different-object blocks participants were uncertain as to the position of the objects and were also uncertain as to whether the target properties were to be extracted from the close (i.e., on the

Experiment 1

It is prudent to employ additional methods to examine the issue of whether object-based attentional selection occurs from a grouped-array or internal spatially invariant representation. This is especially true given the present uncertainty as to whether attention shifts in a discrete or analog fashion coupled with the importance of this assumption for the interpretation of the Vecera and Farah (1994) results. One reasonable way to address this question is through the use of the postdisplay probe technique. Two different variants of this technique have been reported in the literature. Both of these variants are based upon the same logic. That is, if selection is based on location, then the benefits of attention should be obtained for any stimulus that occurs in the selected locations.

Cave (1994; see also Kim & Cave, 1995) examined this assertion by requiring participants to perform two different tasks. In the primary task, participants attended to a digit that appeared in a particular color and reported its value. Digits that appeared in other colors in the same display were to be ignored. Participants responded, by naming the appropriately colored digit, when prompted 2s after the disappearance of the digits. The secondary task, which occurred on 25% of the trials, required participants to make a speeded response to the occurrence of a small black rectangle that appeared at a location that was previously occupied by either a target or distractor digit. Consistent with the logic outlined above, Cave found that probes that occurred at the location previously occupied by a target digit were responded to more quickly than probes that appeared at a location previously occupied by a distractor digit.

In another variant of this postdisplay probe technique, Luck, Fan, and Hillyard (1993) recorded event-related brain potentials (ERPs) elicited by the probe in a visual search task. The probes, which in this case did not require an overt response, were presented at locations that had either previously contained a conjunction target or a distractor. Luck et al. found larger amplitude P100 and N100s for probes that had occurred at a previously occupied target location than for probes that were presented at distractor locations. Given that P100 and N100s are well documented to reflect the spatial distribution of attention in a number of paradigms (Hillyard et al., 1995; Mangun & Hillyard, 1990), these results, like those obtained by Cave (1994), suggest that the postdisplay cue procedure can be used to map the distribution of attention in the visual field.

superimposed trials) or distant (i.e., on the separated trials) locations. Therefore, one reasonable strategy given this uncertainty (see Eriksen & St. James, 1986, for empirical support for this speculation) is to leave attention in the diffuse mode during the presentation of the objects in the different-object blocks. Such a strategy leads to equivalent performance on both the separated and superimposed trials. Thus, this account of the Vecera and Farah results is consistent with previous proposals regarding the strategic distribution of attention in the visual field (Eriksen & St. James, 1986; Jonides, 1983), albeit grouped-array selection in the focused mode, rather than selection of the objects from an internal spatially invariant representation.

We opted to employ the Cave (1994; Kim & Cave, 1995) variant of the postdisplay cue technique to examine whether the objects in the Vecera and Farah (1994; Duncan, 1984) displays were selected from a grouped-array or spatially invariant internal representation. To that end, we presented participants with the box and the line in the superimposed and separated arrangements and asked them to indicate the level of each of two properties on each trial. On different trials the two properties could occur on the same or on different objects. On 25% of the trials a small red dot appeared at a location that was previously occupied by either the box or the line. Participants were instructed to respond to this probe as quickly as possible.

The critical question addressed in this study is whether, on the separated trials, the probe is responded to more quickly when it is presented in a location that was previously occupied by the object that possessed both of the target properties (i.e., a *same-object trial*) than when it is presented in the location that was previously occupied by the object that possessed neither of the target properties. Such a finding would be consistent with the proposal that participants were attending to the locations occupied by the box and the line (i.e., the grouped-array proposal). On the other hand, a failure to find a difference in probe reaction time (RT) in these conditions would provide further support for the proposal that selection took place from an internal spatially invariant representation.

Given that we presented the superimposed displays at the same physical locations as the separated displays (i.e., to the left or right of fixation), we can also examine whether the probe is responded to more quickly when it occurs at the position previously occupied by both the box and the line as compared to the position on the other side of fixation that was occupied by a filler (nontarget) object. Such a pattern of results would be consistent with selection from a grouped-array representation.

Method

Participants. Twenty-four undergraduate students between the ages of 18 and 25 participated in the study. Thirteen of the students were women. All of the students had normal or corrected-to-normal vision, which was confirmed with a Snellen vision test. The students were paid \$4 per hour for their participation.

Stimuli and apparatus. The stimuli were presented on a VGA monitor using the Micro Experimental Laboratory (Schneider, 1988) software. The participants used their left hands to depress the *D* and *F* keys and their right hands to depress the *K* and *L* keys on an IBM keyboard to make their responses. Response time was accurate to 1 ms.

The experimental stimuli were similar to those used by Vecera and Farah (1994). The task relevant stimuli consisted of two stimuli: a box and a line. The box and the line each possessed two properties with two levels per property. The box subtended a visual angle of 0.67° horizontally. The box could be either short (0.86° in height) or tall (1.14° in height). The gap was centered on either the left or the right side of the box and was 0.20° in extent. The line subtended 1.53° of visual angle and was tilted either 8.00° to the left or right of the vertical. The dotted line was constructed using

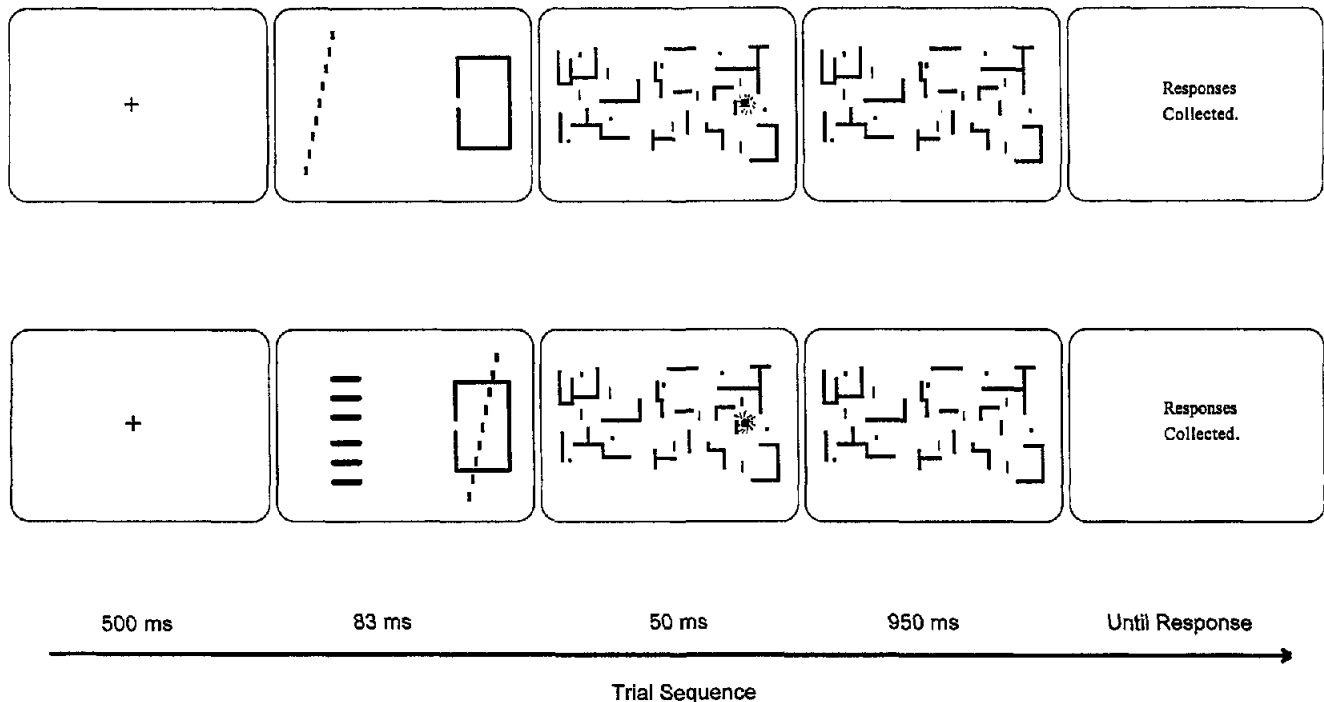


Figure 1. A graphic illustration of the trial sequence for a separated (top) and superimposed (bottom) trial. Although a postdisplay probe is illustrated for each of the conditions, these probes occurred on only 25% of the trials.

a 2×2 array of 4 pixels, whereas the dashed lines consisted of a 4×1 array of 4 pixels.

Two types of displays, which are illustrated in Figure 1, were used within each block of trials: separated and superimposed displays. For the superimposed displays, the box and the line were presented together. The superimposed objects were presented, on an equiprobable basis, to the left or the right of fixation.² For separated displays, the box was presented on one side of fixation and the line was presented on the other side of fixation. The box and the line were separated by 3.90° of visual angle in this condition. The participants viewed the stimuli binocularly at a distance of approximately 60 cm. The entire display subtended a visual angle of approximately 5.00° horizontally.

Stimuli were presented in the left and right positions on each experimental trial. In the separated condition, these two objects were the box and the line. In the superimposed conditions, a filler object appeared in the position that was unoccupied by the task-relevant stimuli (i.e., box and line). This filler object was a vertical array of -s whose luminance was equal to the average luminance of the task-relevant stimuli. The filler object was included to ensure that forward masking of the infrequent postdisplay probe would be equivalent at each of the two display locations in the superimposed conditions.

On 25% of the trials a postdisplay probe was presented immediately after the presentation of the stimulus. The remaining 75% of the trials were not followed by a probe. For all trial types there are only two possible locations for the probe to appear: the center of the display positions for the box and the line on the left and right of fixation. The probe was a red dot measuring 3×3 pixels. For the separated *same-object trials* (i.e., trials on which both of the reported properties were on a single object), 50% of the postdisplay probes were presented at the target location, and the other half of the probes appeared at the location that contained the object

with neither of the target properties. For the separated *different-object trials*, 50% of the probes appeared at the location containing the line; the other half of the probes appeared at the location that contained the box stimulus. On the superimposed condition trials, the post-display probe appeared half of the time at the location of the box-line and half the time at the position of the filler object.

Procedure. Participants performed two sessions of the experiment. Each session occurred on a separate day. There were six pairs of judgments that could be made about the stimuli (i.e., box height-box gap, line texture-line tilt, box height-line tilt, box height-line texture, box gap-line tilt, and box gap-line texture) and each participant made each of the six judgments. The same-object judgment pairs were each repeated twice to equate the number of same and different-object judgments. Participants received two blocks of same-object and two blocks of different-object decisions in each session. The order of the blocks during a session was counterbalanced across participants, and no judgment pairing was ever repeated during a session.

Participants received instructions on the computer monitor introducing the task. They were instructed at the beginning of each block as to which of the two dimensions they would be asked to report on during that block. The participants also received a visual

² The presentation of the box and the line in the superimposed condition differs from that of Vecera and Farah (1994). Vecera and Farah presented the superimposed box and line at fixation while presenting the separated box and line to either side of fixation. In an effort to equate the retinal eccentricity of the box and the line in the two display conditions, we decided to present the superimposed objects at the same retinal positions (i.e., 1.90° to the left or right of fixation) that the stimuli were presented in the separated condition.

example of the objects, and their respective properties, before each block of trials to ensure that they knew which properties they were to judge. Participants were told to be as accurate as possible and to wait to respond to the properties until a message ("please type your responses") was presented on the computer screen. Participants were also told that a postdisplay probe would appear following the box and the line on a small percentage of trials. The participants were instructed to press the space bar as quickly as possible whenever they saw the probe.

Participants were then given one block of 48 practice trials prior to each of the eight experimental blocks (i.e., four same-object and four different-object blocks). The practice block was used to familiarize the participants with the response assignments and the properties to be judged on that particular block. The participants started each trial by pressing the space bar on the keyboard. After the participants pressed the space bar, a 500-ms fixation cross (a "+" sign) appeared in the center of the screen. Following erasure of the fixation cross, the box and the line appeared. In the superimposed condition the box and the line appeared either to the left or right of fixation for 83 ms and were then erased. In the separated condition the line appeared on one side of fixation and the box appeared on the other side of fixation. As in the superimposed condition, the stimuli were presented for 83 ms and then erased. Immediately following the stimulus presentation, a pattern mask similar to the one used by Vecera and Farah (1994) was presented until the judgments about the box, the line, or both were made.

On 25% of the trials, a postdisplay probe appeared along with the pattern mask for 50 ms. The participants were given 1 s to respond to the probe. An error was recorded if the participant failed to respond to the probe within the 1-s period (i.e., miss) or responded when a probe had not been presented (i.e., false alarm). Errors were signaled to the participant by the presentation of a 75-ms, 400-Hz tone. Following the 1 s period allotted for the probe response (on both the probe and nonprobe trials), the participant was queried about the properties of the box, the line, or both. The participant then typed in his or her two responses (i.e., one response for each of the target properties) by depressing a *D* or and *F* and a *K* or an *L* key on the computer keyboard. If either of the responses were incorrect, a 75-ms, 800-Hz tone was presented. The next trial then proceeded in the same fashion.

Each of the eight experimental blocks included 128 trials (32 postdisplay probe trials) for a total of 1,024 trials per participant per session. Within each trial block, half of the trials were separated and half were superimposed. The two levels of each of the target properties were also represented with equal probability in each block. Finally, the box and the line appeared with equal probability on the left and the right of the display.

The participants received a break at the beginning of each block and after the first 64 trials were presented. At the completion of each block of trials the participants received the percentage of correct information about each decision type. The order of trials within each block was random.

Design. The experiment was a within-subjects $2 \times 2 \times 2$ design. The factors included display condition (separated and superimposed), object type (properties located on the same or different objects), and probe presence (present or absent). Within the probe factor, the probe could either appear ipsilateral or contralateral to the target object (i.e., the object that possessed both of the target properties) on the separated same-object trials. The probe could also appear either ipsilateral or contralateral to the location of the box and the line on the superimposed trials (i.e., in the contralateral case the probe would appear in the position that had been occupied by the filler object). Finally, on the separated

different-object trials the distinction between ipsilateral and contralateral positions is irrelevant because one target property is located to the right and the other target attributed is located to the left of fixation.

Results

The accuracy data for the property judgments are presented in Figure 2. Preliminary analyses failed to find differences in accuracy as a function of the visual field of presentation of the objects or the specific pairs of properties that were being judged. Therefore, the data are collapsed over these factors.

Several interesting trends are visible in the data. First, participants responded more accurately when judging two properties located on a single object than when judging properties located on two different objects. Thus, we have replicated the same-object benefit that was reported by Vecera and Farah (1994) and Duncan (1984). Second, the same pattern of effects can be seen whether the postdisplay probe was present or absent. Therefore, it appears that the presence of the probe did not change the manner in which the properties were judged. Finally, it appears that a larger same-object effect was obtained for the separated than for the superimposed display conditions. This effect is quite different from the results obtained by Vecera and Farah who found equivalent magnitude same-object benefits for separated and superimposed display conditions. In fact, the finding of equivalent magnitude same-object benefits in the superimposed and separated conditions served as the basis for Vecera and Farah's proposal for the selection of objects from spatially invariant representations.

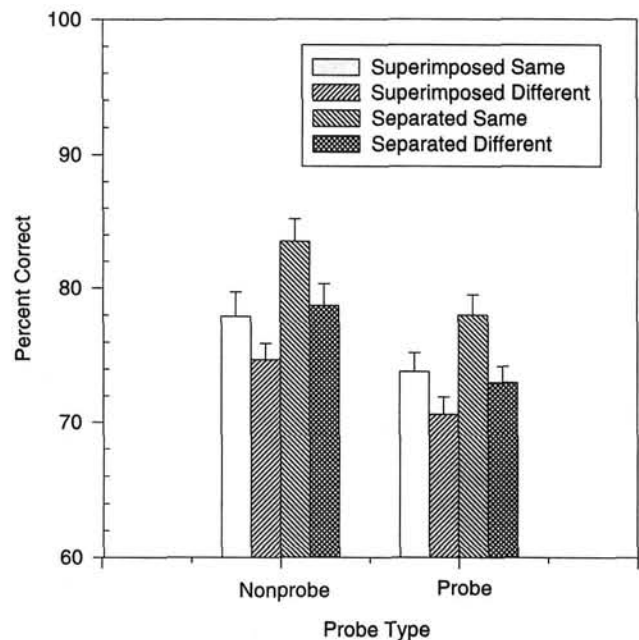


Figure 2. The percentage of correct responses for the property judgment task. The small bars represent standard errors.

The accuracy data were analyzed in a three-way, repeated-measures analysis of variance (ANOVA). The factors included display condition (separated and superimposed), object type (properties located on the same or different objects), and probe presence (present or absent). Main effects were obtained for each of the three factors. Participants were more accurate when judging two properties located on a single object than when judging one property on the box and one property on the line, $F(1, 23) = 15.10, p < .01$. Participants were also more accurate when the box and the line were separated than when they were superimposed, $F(1, 23) = 35.00, p < .01$. Finally, participants performed the judgment task more accurately when the postdisplay probe was absent than on the 25% of the trials on which the probe was present, $F(1, 23) = 63.50, p < .01$.

A two-way interaction was obtained between the object type and the display condition factors, $F(1, 23) = 5.70, p < .05$. As can be seen in Figure 2 the difference in accuracy between same- and different-object conditions was larger in the separated than in the superimposed displays.

Figure 3 presents the RT to the postdisplay probes in each of the experimental conditions. As can be seen in the figure, RTs were faster when the probes were presented at the location of the object or objects (i.e., box and line) in which

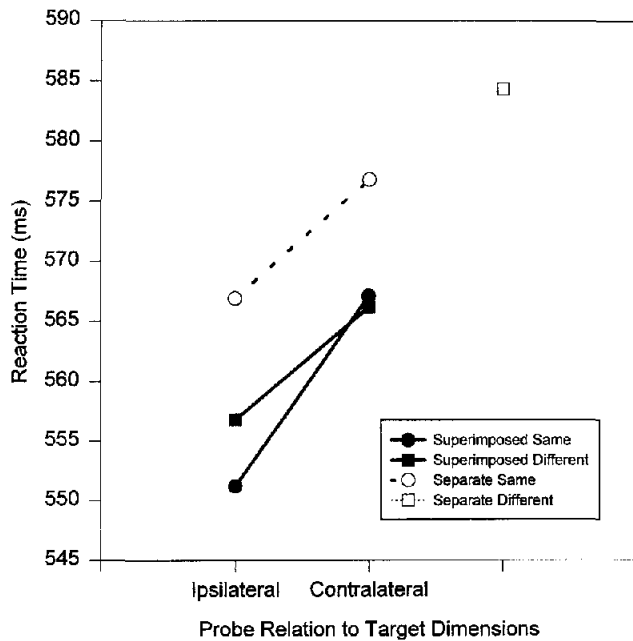


Figure 3. Mean reaction time (RT) values for the postdisplay probes. Ipsilateral refers to probes that occurred at the same position where both of the target properties were located. Contralateral refers to probes that occurred at the position (i.e., either left or right of fixation) in which neither of the target properties were located. The mean RT obtained in the separated different condition is not referred to the ipsilateral or the contralateral positions because in this condition one target property occurred to the left of fixation and the other target property occurred to the right of fixation.

both the target dimensions were found (i.e., the ipsilateral position) than in the display position in which neither of the target dimensions was located. RTs also appear to have been fastest in the superimposed display conditions, somewhat slower in the separated same-object condition, and slowest in the separated different-object condition. Note that there is not a contralateral–ipsilateral distinction on the separated different-object trials because one property was judged on each of the separated objects in this condition.

The probe RTs on correct property-judgment trials were analyzed in a 2×3 , repeated-measures ANOVA. The factors were probe position relative to the property judgments (ipsilateral and contralateral) and display condition (superimposed same-object, superimposed different-object, and separated same-object). The separated different-object condition was not included in this analysis because there was no ipsilateral–contralateral distinction in this condition.

The only significant effect was the main effect of probe position, $F(1, 23) = 10.30, p < .01$. RTs were faster when the probes appeared in the same position as the property judgments than when they appeared on the other side of fixation. We also compared the RTs in the four different conditions (i.e., superimposed same, superimposed different, separated same, and separated different) after collapsing over probe position. A significant main effect of condition was obtained in this analysis, $F(3, 69) = 10.80, p < .01$. Posthoc analyses revealed that the RTs elicited in the separated different-object condition were significantly slower than those obtained in the other three conditions.

Discussion

The main issue addressed in this study was whether object-based selection occurs from a grouped-array or a spatially invariant representation in the Vecera and Farah (1994) and Duncan (1984) tasks and displays. However, the answer to this question has implications beyond these studies because the Vecera and Farah data stand alone in suggesting attentional selection from internal spatially invariant representations. The RT data obtained from the postdisplay probes provide a relatively clear answer to this question. RTs were faster when the postdisplay probe appeared at a location previously occupied by an object that possessed both properties than when the probe appeared at the location of the object that possessed neither of the target properties on the separated trials. Furthermore, participants responded to the postdisplay probes more quickly on the superimposed trials when they appeared at the location previously occupied by both of the objects than when they appeared at the location occupied by the filler object.

Such a pattern of results would not be expected if participants had selected the target objects from an internal spatially invariant representation because, in this case, attention would not be differentially directed to locations in a representation of the visual field. If participants had directed attention to the locations of the target objects, as suggested by the grouped-array proposal, faster RTs would be ex-

pected when the probes occurred at locations formerly occupied by objects that contained the target properties than when they appeared at locations that had been occupied by objects that did not contain the target properties. The data were consistent with this latter proposal.

The fact that we did not replicate all aspects of Vecera and Farah's (1994) property judgment results merits some discussion. As we previously indicated, Vecera and Farah obtained statistically equivalent same-object benefits when the objects were superimposed and when the box and the line were separated. They interpreted these data as evidence for attentional selection from an internal spatially invariant representation. Of course, this interpretation is predicted on the assumption that attention moves in an analog fashion, an assumption that has been questioned on methodological and empirical grounds (Eriksen & Murphy, 1987; Eriksen & Webb, 1989; Kwak et al., 1991; Sagi & Julesz, 1985; Sperling & Weichselgartner, 1995; Yantis, 1988).

Consistent with Vecera and Farah (1994), we found that participants were more accurate in performing the judgment task when both properties were located on a single object than when the properties were distributed between the box and the line. However, unlike Vecera and Farah, we did not find statistically equivalent same-object benefits in the superimposed and separated conditions. Instead, we obtained larger same-object benefits on the separated than on the superimposed trials. Given the analog movement of attention assumption, Vecera and Farah would interpret these results as evidence in favor of attentional selection from a grouped array.

The important question is Why is there a discrepancy in the relative size of the same-object benefit between the separated and superimposed trials in the two studies? One potential explanation is that there was insufficient power in the Vecera and Farah (1994) studies to detect an interaction between object type and object location. However, this appears unlikely given the number of participants employed in their studies and the non-significant trend in their data (see their Figure 2, p. 153) for a larger same-object benefit on the superimposed than on the separated trials.

A more likely explanation concerns the locations at which the objects were presented in the two studies. Vecera and Farah (1994) presented the box and line at fovea in the superimposed condition while presenting the objects to the left and right of fixation in the separated condition. We presented the box and line in the same positions as Vecera and Farah in the separated conditions. However, in an effort to ensure equivalent retinal eccentricity across conditions, we presented the superimposed box and line at the same peripheral locations used in the separated condition.

As can be seen in Figure 2, two results are relevant to our choice of stimulus locations in the separated and superimposed conditions. First, overall accuracy was higher in the separated than in the superimposed conditions. Second, the same-object effect is reduced in the superimposed as compared to the separated condition. These results, when viewed together, are consistent with reports of increased metacontrast masking in the periphery (Breitmeyer, 1984).

That is, it is likely that the reduction of accuracy and the same-object effect in the superimposed condition can be attributed to the increased interference between the box and the line in the superimposed, as compared to the separated, condition. Hereinafter, we refer to this explanation as the *masking hypothesis*. A smaller metacontrast masking effect would be expected in the Vecera and Farah (1994) studies given that superimposed objects were presented at fixation. In essence, the choice of the locations for the objects in the superimposed condition represents a trade-off between confounding location, and therefore acuity, between the superimposed and separated trials, as in the Vecera and Farah studies, and the potential for increased metacontrast masking when location is not confounded, as in our study.

However, there is another alternative explanation for the discrepancy in the relative size of the same-object benefit between the separated and superimposed trials in the Vecera and Farah (1994) and our study that merits discussion. That is, it is conceivable that the presence of the postdisplay probes in our study encouraged participants to adopt a space-based attentional selection strategy in an effort to rapidly locate the probe. In turn, such a space-based selection strategy might also have been employed during the identification of the properties of the box and the line. Indeed, the larger same-object benefit for the separated than for the superimposed trials is consistent with this hypothesis. Hereinafter this explanation is referred to as the *selection strategy modification hypothesis*.

In fact, in their Experiment 3, Vecera and Farah (1994) changed the task performed by the participants such that they were to make a simple RT response to the presence of a luminance increment that could occur on either the box or the line in the superimposed and separated trials. Another important aspect of their study was that a precue (i.e., a brief brightening of the box or the line either 150 or 400 ms prior to the occurrence of the luminance increment target) was employed to indicate the subsequent location of the luminance target. The precue indicated the subsequent target location with a validity of 80%. Under these conditions, in which the box and the line served only to indicate the potential location of the luminance target (i.e., participants did not judge the properties of the box in the line in this study), the RT difference between validly and invalidly cued locations was larger on the separated than on the superimposed trials. Such a pattern of results is consistent with a space-based selection strategy and suggests that the nature of the task (i.e., identify properties of the same or different objects as compared to detecting the occurrence of a spatially precued luminance target) has a strong influence on the mode of attentional selection.

The important question in this regard is whether the introduction of the postdisplay probe in our study encouraged participants to adopt a space-based selection strategy as in Vecera and Farah's (1994) Experiment 3, rather than adopt a spatially invariant selection strategy as Vecera and Farah suggested occurred in their studies in which the properties of the box and line were judged. We have several reasons to believe that the use of postdisplay probes in our study would not lead participants to adopt a different atten-

tional selection strategy than that which would be employed during the judgment of the properties of the line and the box in the absence of the probes.

First, the probes in our study occurred on a small subset of the trials rather than on the great majority of the trials as in the Vecera and Farah (1994) studies. Furthermore, the potential location of the postdisplay probes was not cued in advance as it was in Vecera and Farah's study. Given these differences in the design of the studies, it appears unlikely that participants would prepare to respond to the probes, using a space-based selection strategy, in our study. Second, participants identified the properties of the box and the line on all of the trials in our study. The box and the line served only to precue the location of the luminance targets in the Vecera and Farah study. Participants did not identify the properties of the box and the line in their study. Again, given these differences in the design of the studies, it seems unlikely that participants would have modified their selection strategy to prefocus attention on a specific area of space in preparation for the infrequently occurring (and spatially unpredictable) probes in our study.

Experiment 2

Although we believe that we have offered reasonably strong arguments to suggest that the presence of the postdisplay probes used in our study would not have modified the way that participants performed the property judgment task, we decided to err on the side of caution and test this hypothesis empirically. To that end, we ran an additional study in which the box and line were presented at fixation, as in the original Vecera and Farah (1994) study, rather than in the periphery as we had done earlier.

According to the masking hypothesis, equivalent magnitude in same-object benefits is expected on the superimposed and separated trials. This follows because the meta-contrast masking of the box and line should be reduced on the superimposed trials because they appear at fixation. Furthermore, according to the masking hypothesis, we also no longer expect the superimposed trials to be responded to less accurately than the separated trials, as is illustrated in Figure 2. Instead, given that the superimposed objects appear at fovea whereas the separated objects appear in the periphery, it seems likely that the accuracy of judging the object properties should be higher on the superimposed trials. The selection strategy modification hypothesis makes different predictions. If the probes serve to encourage participants to employ a space-based selection strategy in Experiment 1, presumably reflected in the larger same-object effect on the separated than on the superimposed trials, then we also expect participants to adopt a space-based selection strategy in the present study. In other words the selection strategy hypothesis predicts that the same-object effect is larger for the separated than for the superimposed trials regardless of whether the superimposed objects appear in the periphery or at fixation.

Method

Participants. Twenty-four undergraduate students between the ages of 19 and 27 participated in the study. Fifteen of the students were women. All of the students had normal or corrected-to-normal vision, which was confirmed with a Snellen vision test. The students were paid \$4 per hour for their participation.

Stimuli and apparatus. The stimuli and apparatus were identical to that employed in Experiment 1 with the following exceptions. First, the superimposed stimuli were presented at fixation rather than in the periphery, as had been done in Experiment 1. Second, the filler object that had been presented in the visual field opposite to that in which the superimposed stimuli were presented in Experiment 1 (see Figure 1) was not used in the present study. Third, the postdisplay probes, which were presented on 25% of the trials, could occur at three different locations in the present study. On the separated trials, in which the box and the line were presented in opposite visual fields, the probe could occur next to the box or the line in either the left or right visual field. Thus, on a separated same-object trial the probe could occur at either the position that contained the object with both target properties (i.e., ipsilateral to the target object) or at the position of the object that contained neither of the target properties (i.e., contralateral to the target object). On the separated different-object trials the probe could occur next to the box or the line. Finally, on the superimposed trials the probe, when present, occurred at fixation.

Procedure. The procedure was identical to that employed in the first study.

Design. The experiment was a within-subjects, $2 \times 2 \times 2$ design. The factors included display condition (separated and superimposed), object type (properties located on the same or different objects), and probe presence (present or absent). Within the probe factor, the probe could either appear ipsilateral or contralateral to the target object (i.e., the object that possessed both of the target properties) on the separated same-object trials. On the separated different-object trials the distinction between ipsilateral and contralateral positions is irrelevant because one target property is located to the right and the other target attributed is located to the left of fixation. The probe could also appear at fixation on the superimposed trials.

Results

The percentage of correct responses for the property judgments on the superimposed and separated trials are presented in Figure 4. These data were analyzed in a three-way, repeated-measures ANOVA with display condition (separated and superimposed), object type (properties located on the same or different objects), and probe presence (present or absent) as factors. As in Experiment 1, participants were more accurate when both of the target properties appeared on a single object than when one of the target properties occurred on the box and the other target property occurred on the line, $F(1, 23) = 39.80, p < .01$. Participants performed the property judgment task more accurately when the postdisplay probe was absent than on the 25% of the trials on which the probe was present, $F(1, 23) = 101.80, p < .01$.

Consistent with the masking hypothesis, we failed to find

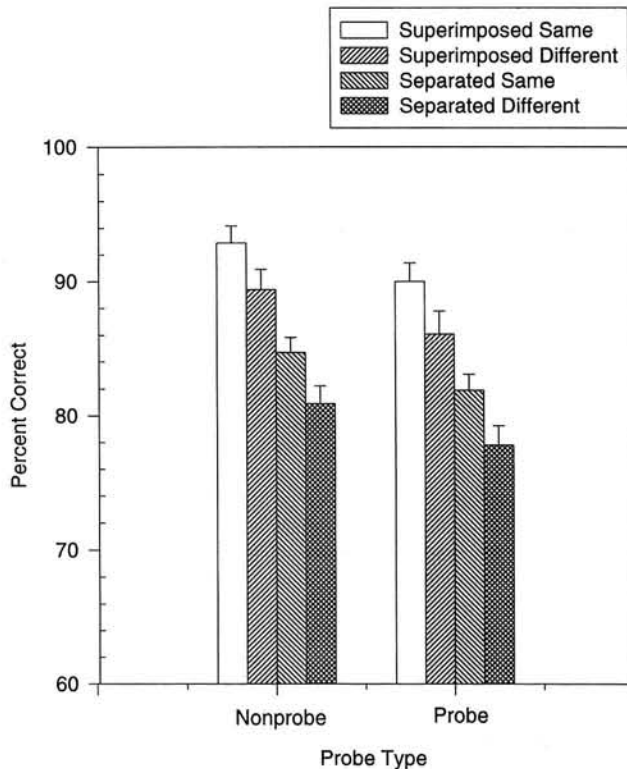


Figure 4. The percentage of correct responses for the property judgment task when the superimposed objects were presented at fixation. The small bars represent standard errors.

a difference in the size of the same-object effect on the superimposed and separated trials, $p > .47$. As was expected given that the superimposed objects appeared at fixation, accuracy was higher for the superimposed than for the separated trials, $F(1, 23) = 103.10$, $p < .01$. Thus, it appears that the larger same-object effect obtained in Experiment 1 (see Figure 2) was not due to participants modifying their attentional selection strategy as a result of the infrequent presentation of the postdisplay probe but instead can be attributed to the metacontrast masking of the superimposed objects when presented in the periphery.³

Given that we have replicated Vecera and Farah's (1994) finding of equivalent same-object effects for the superimposed and separated trials, the important question is whether we still obtain faster probe RTs on the separated trials when the probe appears at the position of the object that possessed both of the target properties as compared to the position of the object that possessed neither of the target properties. The answer to this question is affirmative. The mean RTs on the trials in which the probe appeared ipsilateral and contralateral to the object that possessed both target properties were 566 and 575 ms, respectively. This difference was statistically significant, $F(1, 23) = 46.10$, $p < .01$.

As in Experiment 1, the separated different-object probe RTs were somewhat longer than the RTs elicited in the other conditions. The mean RT in this condition was 588

ms. The mean RTs on the superimposed same and different-object trials were 548 and 546 ms, respectively. A comparison of the average of the same and different-object probe RTs for the superimposed trials (i.e., 547 ms), with the average of the ipsilateral and contralateral probe RTs for the separated same-object trials (i.e., 571 ms) and the separated different-object trials (i.e., 588 ms), was significant, $F(2, 46) = 6.40$, $p < .01$. Post hoc comparisons revealed that RTs elicited on the superimposed trials were faster than those obtained on the separated same-object trials that in turn were faster than the RTs obtained on the separated different-object trials. It would appear reasonable to assume that the fast RTs obtained on the superimposed trials were due, in part, to the fact that these probes occurred at fixation.

Discussion

The results obtained in our two studies described earlier—in particular the fact that the postdisplay probes were responded to more rapidly when they occurred at a location previously occupied by an object (box or line) that possessed both of the target properties—provide strong support for the hypothesis that attentional selection occurred on the basis of grouped-array rather than spatially invariant representations. Thus, the target objects appear to have been selected on the basis of the activation of particular regions of a representation of the visual field.

It is also important to note that our results with the postdisplay probe technique are consistent with previous reports of spatial attention effects in paradigms in which evidence has been obtained for object-based selection. For example, Egly, Driver, and Rafal (1994) performed a series of studies in which a precue was used to indicate the potential location of a luminance target at the end of one of two rectangles that were presented side-by-side. Most of the time the target appeared at the cued (valid) location. However, on a small percentage of trials the target appeared at an invalid (uncued) location in either the cued or the uncued rectangle.

Despite the fact that the invalid locations in the cued and uncued rectangles were the same distance from the cued location, RTs were longer when the target appeared in the

³ The main aim of Experiment 2 was to determine if the different-magnitude object effects (i.e., same-minus-different-object accuracy data) obtained in the separated and superimposed conditions in Experiment 1 were abolished when the superimposed objects appeared at fixation. Indeed, this was the case. However, a comparison of the object effects between Experiments 1 and 2 indicates a slight increase in the object effect from Experiment 1 to Experiment 2 on the superimposed trials (3% in Experiment 1 and 4% in Experiment 2) and a slight decrease in the object effect on the separated trials (5% in Experiment 1 and 4% in Experiment 2). In an effort to determine whether these changes were reliable, we performed between-experiment analyses on the magnitude of the object effects separately for the superimposed and separated conditions. Neither of these effects were significant, $ps > .35$.

invalid location on the uncued than on the cued rectangle. This difference in RTs between the invalid locations was interpreted as evidence for object-based selection (i.e., faster processing of the target if it appeared in the attended rectangle). Interestingly, however, RT was also increased, albeit not to the same extent as when the target appeared in the uncued rectangle, when the target appeared in the invalid location in the cued rectangle. This result was interpreted as the cost incurred by the requirement to reorient attention (see also Vecera, 1994).

Kramer and Jacobson (1991) reported a similar pattern of effects in a paradigm in which participants were instructed to respond to a centrally located target and ignore surrounding distractors. The influence of the distractors on target processing was greatly reduced when the target and distractors were components of different objects: an object-based selection effect. However, distractor effects were also diminished when the distractors were moved farther away from the target even when they were the components of a single object: a spatial attention effect.

The results from both the Egly et al. (1994; see also Egly, Rafal, Driver, & Starrveveld, 1994; Vecera, 1994) and Kramer and Jacobson (1991) studies as well as the probe data obtained in our studies support the notion of object-based selection from grouped arrays. That is, in three different paradigms that have been employed to examine object-based attentional selection, the data appear most compatible with the proposal that selection of objects is mediated by their location in the visual field. How such object-based attentional selection is accomplished awaits further study: Does it occur in coordination with an object-recognition process by the same mechanism proposed to underlie traditional notions of space-based selection, albeit with a more flexible focus than either the spotlight or zoom-lens models, or by a separate attentional selection mechanism (Farah, Wallace, & Vecera, 1993; Humphreys, Olson, Romani, & Riddoch, 1995; Kramer & Watson, 1995)?⁴

Conclusions

At first glance the conclusion that object-based attentional selection is accomplished by the activation of the spatial locations occupied by an object may appear at odds with object-recognition theories that suggest that objects are represented in memory in an object-centered coordinate system in which object parts are specified relative to the main axis of the object (Marr, 1982; Marr & Nishihara, 1978) or that simple object parts are defined in relation to other object parts (Lowe, 1987). Indeed, both psychophysical and neurophysiological studies have obtained data that is consistent with object-centered representations (Palmer, Rosch, & Chase 1981; Perrett et al., 1991; Warrington & Taylor, 1973). We wish to make clear that we are not disputing the existence of object-centered or high-level, viewer-centered representations (see Perrett, Oram, Hietanen, & Benson, 1994, for an interesting review of this position) in memory. Instead, what we are arguing is that the data obtained thus

far suggest that the initial selection of objects from the visual environment appears to occur on the basis of the direction of attention to specific spatial locations, that is, those locations occupied by an object.

⁴ It should be noted that although the data obtained in our studies show that spatial attention is used to select a target over a distractor (either a distractor that can serve as a target on other trials, as was the case in the separated same-object trials when the probe fell on the location of the object that possessed neither of the target attributes, or a distractor that is never task relevant, as was the case with the filler object employed in Experiment 1) that occurs in a different location, our studies are not diagnostic with respect to whether spatial attention can select a target object over a superimposed distractor object. In both of our studies the probe RT differences between same- and different-object trials in the superimposed object conditions were small and nonsignificant (i.e., 552 and 556 ms for Experiment 1 and 548 and 546 ms for Experiment 2 for the same- and different-object trials, respectively). However, this is not particularly surprising given that the probe was presented midway between the contour of the box and the line on the superimposed trials. Although the choice of this probe location, between the contour of the two objects, precluded the examination of differential spatial selection of the two objects on the superimposed trials, it also ensured that the contour of the box and the line would not mask the postdisplay probe. We are currently in the process of systematically investigating the selective processing of superimposed and partially occluded objects.

References

- Baylis, G., & Driver, J. (1992). Visual parsing and response competition: The effect of grouping factors. *Perception and Psychophysics*, *51*, 145–162.
- Baylis, G., & Driver, J. (1993). Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 451–470.
- Breitmeyer, B. G. (1984). *Visual masking: An integrative approach*. New York: Oxford University Press.
- Broadbent, D. E. (1982). Task combination and selective intake of information. *Acta Psychologica*, *50*, 253–290.
- Cave, K. (1994). *Measuring the allocation of spatial attention*. Manuscript submitted for publication.
- Downing, C. J., & Pinker, S. (1985). The spatial structure of visual attention. In M. I. Posner & O. S. M. Martin (Eds.), *Attention and Performance XI* (pp. 171–187). Hillsdale, NJ: Erlbaum.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, *113*, 501–517.
- Duncan, J. (1993). Similarity between concurrent visual discriminations: Dimensions and objects. *Perception and Psychophysics*, *54*, 425–430.
- Egry, R., Driver, J., & Rafal, R. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, *123*, 161–177.
- Egry, R., Rafal, R., Driver, J., & Starrveveld, Y. (1994). Covert orienting in the split brain reveals hemispheric specialization for object-based attention. *Psychological Science*, *5*, 380–383.
- Eriksen, C., & Murphy, G. (1987). Movement of attentional focus

- across the visual field: A critical look at the evidence. *Perception and Psychophysics*, 42, 299–305.
- Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception and Psychophysics*, 40, 225–240.
- Eriksen, C. W., & Webb, J. (1989). Shifting of attentional focus within and about a visual display. *Perception and Psychophysics*, 45, 175–183.
- Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 583–597.
- Farah, M., Wallace, M., & Vecera, S. (1993). "What" and "where" in visual attention: Evidence from neglect syndrome. In I. Robertson & J. Marshall (Eds.), *Unilateral neglect: Clinical and experimental studies* (pp. 123–138). Hillsdale, NJ: Erlbaum.
- Hillyard, S. A., Anillo-Vento, L., Clark, V. P., Heinze, H. J., Luck, S. J., & Mangun, G. R. (1995). Neuroimaging approaches to the study of visual attention: A tutorial. In A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), *Converging operations in the study of visual selective attention* (pp. 107–138). Washington, DC: American Psychological Association.
- Humphreys, G. W., Olson, A., Romani, C., & Riddoch, M. J. (1995). Competitive mechanisms of selection by space and object: A neuropsychological approach. In A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), *Converging operations in the study of visual selective attention* (pp. 365–393). Washington, DC: American Psychological Association.
- Jolicoeur, P., Ullman, S., & MacKay, M. (1991). Visual curve tracing properties. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 997–1022.
- Jonides, J. (1983). Further toward a model of the mind's eye. *Bulletin of the Psychonomic Society*, 21, 247–250.
- Kahneman, D., & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual Organization* (pp. 181–211). Hillsdale, NJ: Erlbaum.
- Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & R. Davies (Eds.), *Varieties of attention* (pp. 29–62). New York: Academic Press.
- Kahneman, D., Treisman, A., & Gibbs, B. (1992). The reviewing of object files: Object specific integration of information. *Cognitive Psychology*, 24, 175–219.
- Kim, M. S., & Cave, K. (1995). Spatial attention in visual search for features and feature conjunctions. *Psychological Science*, 6, 376–380.
- Kramer, A. F., & Jacobson, A. (1991). Perceptual organization and focused attention: The role of objects and proximity in visual processing. *Perception and Psychophysics*, 50, 267–284.
- Kramer, A. F., & Watson, S. E. (1995). Object-based visual selection and the principle of uniform connectedness. In A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), *Converging operations in the study of visual selective attention* (pp. 395–414). Washington, DC: American Psychological Association.
- Kwak, H., Dagenbach, D., & Egeth, H. (1991). Further evidence for a time-independent shift of the focus of attention. *Perception and Psychophysics*, 49, 473–480.
- LaBerge, D., & Brown, V. (1989). Theory of attentional operation in shape identification. *Psychological Review*, 96, 101–124.
- Lowe, D. (1987). Three-dimensional object recognition from single two dimensional images. *Artificial Intelligence*, 31, 355–395.
- Luck, S., Fan, S., & Hillyard, S. (1993). Attention related modulation of sensory-evoked brain activity in a visual search task. *Journal of Cognitive Neuroscience*, 5, 188–195.
- Mangun, R., & Hillyard, S. (1990). Electrophysiological studies of visual selective attention in humans. In A. Scheibel & A. Wechsler (Eds.), *Neurobiology of higher cognitive function* (pp. 271–295). New York: Guilford Press.
- Marr, D. (1982). *Vision*. San Francisco: Freeman.
- Marr, D., & Nishihara, H. (1978). Representation and recognition of the spatial organization of 3-dimensional shapes. *Proceedings of the Royal Society of London, B(200)*, 269–294.
- Palmer, S., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In J. Long & A. Baddley (Eds.), *Attention and performance IX* (pp. 135–151). Hillsdale, NJ: Erlbaum.
- Perrett, D., Oram, M., Harries, M., Bevan, R., Hietanen, J., Benson, P., & Thomas, S. (1991). Viewer-centered and object-centered encoding of heads by cells in the superior temporal sulcus of the rhesus monkey. *Experimental Brain Research*, 86, 159–173.
- Perrett, D., Oram, M., Hietanen, J., & Benson, P. (1994). Issues of representation in object vision. In M. Farah & G. Ratcliff (Eds.), *The neuropsychology of high level vision* (pp. 33–61). Hillsdale, NJ: Erlbaum.
- Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, 109, 160–174.
- Remington, R., & Pierce, L. (1984). Moving attention: Evidence for time-invariant shifts of visual selective attention. *Perception and Psychophysics*, 35, 393–399.
- Sagi, D., & Julesz, B. (1985). Fast noninertial shifts of attention. *Spatial Vision*, 2, 141–149.
- Schneider, W. (1988). Micro Experimental Laboratory: An integrated system for IBM-PC compatibles. *Behavior Research Methods, Instrumentation, and Computers*, 20, 206–217.
- Shulman, G., Remington, R., & McLean, J. (1979). Moving attention through visual space. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 522–526.
- Sperling, G., & Weichselgartner, E. (1995). Episodic theory of the dynamics of spatial attention. *Psychological Review*, 102, 503–532.
- Treisman, A., Kahneman, D., & Burkell, J. (1983). Perceptual objects and the cost of filtering. *Perception and Psychophysics*, 33, 527–532.
- Tsal, Y. (1983). Movements of attention across the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 523–530.
- 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., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General*, 123, 146–160.
- Warrington, E., & Taylor, A. (1973). The contribution of the right parietal lobe to object recognition. *Cortex*, 9, 152–164.
- Yantis, S. (1988). On analog movements of attention. *Perception and Psychophysics*, 43, 203–206.

Received May 12, 1995

Revision received October 13, 1995

Accepted January 4, 1996 ■