Visuospatial attention: Beyond a spotlight model

KYLE R. CAVE and NARCISSE P. BICHOT
Vanderbilt University, Nashville, Tennessee

Much of the research in visual attention has been driven by the spotlight metaphor. This metaphor has been useful over many years for generating experimental questions in attention research. However, theories and models of visual selection have reached such a level of complexity that debate now centers around more specific questions about the nature of attention. In this review, the general question "Is visual attention like a spotlight?" is broken down into seven specific questions concerning the nature of visual attention, and the evidence relevant to each is examined. The answers to these specific questions provide important clues about why visual selection is necessary and what purpose attention plays in visual cognition.

For years, the study of visual attention has been dominated by "spotlight theories" claiming that information from one region of the visual field is selected by a mechanism analogous to a spotlight. This selection is assumed to occur while the eyes are stationary (covert attention) and is separate from the visual selection that is performed by directing gaze to a particular location. The idea of a spotlight as a metaphor for visual attention was suggested by Posner, Snyder, and Davidson (1980), although the concept may have originated before that, and ideas about rays from the eye selecting objects were discussed among the ancient Greeks (Zajonc, 1993). More recently, the idea of scale adjustment has been added, resulting in a zoom lens model (C. W. Eriksen & St. James, 1986; C. W. Eriksen & Yeh, 1985). Throughout this paper, the general concept of spotlight models will include zoom lens models.

The comparison between attention and a spotlight can be described with enough structural consistency and systematicity that Gentner and Jeziorski (1993) would probably classify it as an analogy, rather than just a metaphor. The comparison is detailed enough that it leads to some specific experimental predictions about visual attention, which have been the basis for a number of informative experiments. Despite the useful guidance that this metaphor has given to attentional research, however, attentional theories and models have now become complex and detailed enough that they have moved beyond the spotlight metaphor. One possible approach would be to follow the trend described by Gentner and Grudin (1985) and move to a new metaphor based on a more complex system. However, there is no obvious alternative system to

succeed the spotlight, and, besides, as Lewontin (1981) noted, "the price of metaphor is eternal vigilance." Thus, we start with the general question "Is attention like a spotlight?" and break it down into seven more specific questions about the nature of visual attention. These questions address basic and widely debated issues, such as location-versus object-based selection, whether attention moves in an analog fashion (i.e., passing over intermediate locations and taking longer for shifts that involve longer distances), the size and shape of the selected region (and, at a more basic level, whether attention works by facilitating targets or inhibiting distractors), and whether multiple noncontiguous regions of space can be selected simultaneously. In reviewing the experimental evidence relevant to answering each of these questions, we will consider evidence from studies using many different experimental techniques (e.g., spatial cuing involving detection or discrimination and with or without distractors, distractor interference, visual search, spatial probing, rapid serial visual presentation) and different measures (e.g., accuracy, reaction time). The results of some studies contradict one another, and, in many cases, differences are attributable to differences in the techniques used. Despite the complication arising from comparing studies with different techniques, however, each experimental technique has revealed different aspects of visual selection, and only by considering and integrating all of these data can we expect to have a complete picture of attention. After reviewing the relevant studies, we will explore what the results tell us about the nature of visual attention.

1. Is Selection Based on Location?

Of all the questions rolled together into the spotlight metaphor, the role of location in selection is the most basic. A spotlight illuminates anything and everything that is located within the region covered by its beam, and, thus, the spotlight metaphor suggests a mechanism that selects all the information from a particular location or region of the visual field and excludes information from all other locations or regions. Selecting by location is a

straightforward approach for a system, such as the visual system, in which the raw input arrives organized spatially. In addition to performing visual selection, a spotlight mechanism might also help in achieving position constancy. In some models (e.g., Kosslyn & Koenig, 1992; Olshausen, Anderson, & Van Essen, 1993), the region of the visual input that falls within an attentional window is copied into a higher level buffer, where it is available for high-level processing. One benefit of this transfer is that the information from the selected region is recoded in a location-independent reference frame, making it simpler to recognize an object regardless of where it appears in the visual field.

Location obviously plays *some* role in selection, as has been shown in countless cuing experiments. In one of the earliest, C. W. Eriksen and Hoffman (1974) showed that subjects responded more quickly to a stimulus if they knew where it would appear than if they did not. Other cuing experiments by Posner and his collaborators (Posner, Nissen, & Ogden, 1978; Posner et al., 1980) compared subjects' ability to detect stimuli appearing at expected and unexpected locations. At the beginning of each trial, a cue indicated which of two possible locations was likely to hold the upcoming stimulus. Response times (RTs) were faster for stimuli at the cued location.

In later cuing experiments using 10 different locations along the horizontal midline, Downing and Pinker (1985) found that RT increased steadily with distance between the expected location and the actual target location. Their data were somewhat unclear, but they suggested that the size of the cuing effect depended not on the physical distance between cue and stimulus, as one might expect, but on the "cortical distance" between the two points. Since more visual cortex is devoted to processing the center visual field, distance effects were much larger when the two points were near the center and were smaller when the two points were in the periphery. The largest effects occurred when the two points were on opposite sides of the vertical midline.

These distance effects were not universal, however. Hughes and Zimba (1985, 1987) found that responses were uniformly fast whenever cue and target were in the same hemifield and uniformly slow when in different hemifields. They found similar results whether they tested left versus right or top versus bottom. On the other hand, Rizzolatti, Riggio, Dascola, and Umiltà (1987) did find slower responses with longer distances within a quadrant, as Downing and Pinker (1985) would predict. They also found an additional cost when cue and target were on opposite sides of either the horizontal or the vertical midline, and, because their stimuli were located peripherally, they concluded that the effect was not due to cortical magnification. (Rizzolatti et al. claim that delays produced by cuing reflect the time associated with programming eye movements.) One important difference between these studies is that Downing and Pinker and Rizzolatti et al. marked each of their possible target locations with a square, whereas Hughes and Zimba's stimuli appeared on a rela-

tively empty screen. When Zimba and Hughes (1987) replicated Downing and Pinker's experiment with 10 squares marking the 10 possible target locations, they found distance effects as well. These experiments taken together suggest that RT increases with distance between cue and target only when the displays include landmarks.

The landmark squares may produce distance effects because they serve as distractors. When the squares are present, subjects cannot respond to just any stimulus they detect on the screen. They must be sure that the stimulus is a target and not part of one of the squares. They may be forced to focus their attention to the small area within the cued square. Without the squares, they are free to spread their attention over the entire screen (or at least one half of it). An alternative explanation is that attention is allocated not to locations but to objects and that the landmark squares are objects that can receive attention. This sort of explanation, which places attention at the relatively late stage of object representations rather than the early stage of spatially organized representations, will be explored more thoroughly below.

Posner et al. (1980) found that the RT advantage provided by a location cue is smaller for a discrimination task than for a simple detection task. Downing (1988) went a step further, carefully comparing cuing effects using four tasks: luminance detection, brightness discrimination, orientation discrimination, and form discrimination. She measured d' rather than RT and found that sensitivity generally dropped off as distance between the cue and the stimulus increased. Each of the 12 possible stimulus locations was marked by a square, and there was no control for midline crossings. Sensitivity dropped off more steeply in orientation discrimination and form discrimination than it did in luminance detection and brightness discrimination. It also dropped off more steeply when stimuli were packed closer together.

In addition to demonstrating differences in the way attention operates in detection and discrimination tasks, Downing's (1988) results also clearly demonstrated that the effects of spatial attention go beyond changes in decision-making bias. Downing clearly showed that attention produces an increase in perceptual sensitivity. More recent studies have also demonstrated that spatial attention allocated in response to a location cue affects perceptual sensitivity (Handy, Kingstone, & Mangun, 1996; Hawkins et al., 1990).

These cuing studies and other similar studies provide important data that must be explained by any complete account of visual attention. For instance, most of them (except Downing, 1988) used stimuli that appeared on a more or less blank display. There were usually no distractors to filter out. The fact that visual selection operates even in simple tasks with no distractors present suggests that it plays a role in early visual processing. Furthermore, cuing of form in a detection task does not produce facilitation like that produced by location, illustrating the special role for location in visual selection (Posner et al., 1980; but see Humphreys, 1981, for color cuing in

a discrimination task). However, data from cuing studies by themselves are unlikely to tell the whole story. Presumably, visual selection is most important in those situations in which distractors interfere with processing of visual targets. The importance of attention in suppressing competing information from distractors is supported by neurophysiological studies showing that attentional modulation of neural responses is greatest when target and distractor both fall in the receptive field of a neuron, thus competing for representation by that neuron (Luck, Chelazzi, Hillyard, & Desimone, 1997; Moran & Desimone, 1985). The most interesting results are thus likely to appear under conditions in which targets must be selected and distractors must be excluded. The different results between cuing studies with and without landmarks illustrate how performance can change drastically when distracting stimuli are added to the display.

Another limitation of these cuing studies comes from the suggestion built into the instructions that location is an important property. The visual system might be able to select from its input on the basis of a number of visual properties, such as color, motion, size, and stereoscopic depth. Thus, it is entirely possible that, had the subjects been instructed to select by color (e.g., "attend to the red stimulus"), without any expectation of target position, stimulus positions would play no role in the selection process. Alternatively, if all selection is mediated by location, as suggested by the spotlight metaphor, a subject who is to attend to the red stimulus must first determine where the red stimulus is located and then select the information from that location. In this case, location should play an important and measurable role even when selection by location is not suggested by the instructions.

The assumption that visual selection works by separating the information from one location or one region of the visual field and passing it on to higher level processing mechanisms can be justified partly by the fact that the retinal input to the visual system is organized by location (albeit retinal location) and that location plays a special role in the organization of visual representations. Navon (1978) argues that location dominates color, orientation, and other visual properties in a hierarchy of information organization (and that location is in turn dominated by time). Kubovy (1981) generally agrees with these points about the organization of visual information and goes on to claim that location (along with time) is an "indispensable attribute" by which other visual attributes are organized. However, in some cases, the assumption that selection is location-based may not arise from carefully considered theoretical or empirical grounds but merely because the idea feels natural, given our experience with eye movements. Eye movements constitute a selection mechanism that is based on location, and because we move our eyes all the time, many researchers undoubtedly find it natural to believe that attentional selection within a fixation is done in the same way. Part of the appeal of the spotlight metaphor may actually lie in its similarity to an eye-movement metaphor.

Several experiments have tested whether location affects selection even when location is irrelevant to the task and selection by location is not suggested by the instructions. Assuming that selection is based on location and that time and effort are required to shift attention from one location to another, then tasks involving multiple stimuli should be easier when the stimuli are at or near the same location. An early experiment by Hoffman and Nelson (1981) shows that the distance between two stimuli can affect their processing. Their subjects performed two tasks simultaneously. While searching for target letters among a rapidly changing array of four letters (similar to displays used by Schneider & Shiffrin, 1977), they also watched for a small U-shaped figure to appear near one of the letters and determined its orientation when it did. On trials in which they correctly identified which target letter appeared, they were more likely to correctly report the orientation of the U-shaped figure if it was positioned near the target letter. When they were wrong about the letter, they were more likely to be right about the orientation of the U-shaped figure if it was not near the target letter. These results indicate that two objects can be processed more quickly when they are close together. This pattern would be expected if some regions of the visual field were selected and/or others were inhibited at any given time. Whenever the area around one visual object is selected, other objects near it should benefit, and objects far from it should be at a disadvantage. Further evidence of this type comes from Hoffman, Nelson, and Houck (1983),[1] who instructed their subjects to search for target letters in an experiment in which the secondary task was to monitor four small dots to detect when one of them flickered. Subjects were more likely to detect the target letter when it was near the dot that flickered.

A recent experiment by Kim and Cave (1995) relied on similar logic to investigate the allocation of spatial attention during feature and conjunction visual searches. They found that subjects were faster at responding to a small flashed black probe when it appeared at a location previously occupied by the target of a visual search display than when the probe appeared at a location formerly occupied by a distractor. The importance of location in visual search was further demonstrated in a study by Cave and Pashler (1995) in which each stimulus consisted of a series of frames, each containing a target digit of one color and a distractor digit of another color. The subjects' task was to name the highest digit of the target color. Subjects made fewer errors when successive targets appeared at the same location than when they appeared at different locations, apparently because they were selecting target objects by selecting their locations. Thus, when successive targets appeared at the same location, there was no need to select a new location, leaving more time for stimulus identification (see also Maljkovic & Nakayama, 1996).

These experiments show that, at least under some circumstances, two targets are easier to identify when they both appear at or near the same location. If selection is based on location, then it might also be harder to prevent

interference from distractors that are near a target. C. W. Eriksen and Hoffman (1972) used configurations of letters to show that distractors near a target (within about 1°) interfere with naming the target, whereas targets farther away do not. This interference disappears if the target location is cued at least 150 msec in advance. C. W. Eriksen and Hoffman (1973) and B. A. Eriksen & C. W. Eriksen (1974) argued that this interference occurred at the response stage, because the interference was much greater when the nearby distractors were associated with a different response. (In a later experiment, Gatti & Egeth, 1978, found Stroop interference from distractors as far away as 5° from target.)

Further evidence for selection by location can be found in the errors that subjects make when they select the wrong stimulus in a multielement display. Snyder (1972) found that when subjects misreported the shape of a stimulus selected by the criterion of color, they were likely to report the properties of a neighboring stimulus. However, the fact that subjects were also required to report the location of the target raises the possibility that location selection was suggested by the instructions. Tsal and Lavie (1988) found similar results with somewhat different methods. They argued that performance during error trials may not be representative of the mechanisms used during correct trials, and their experiments did not rely on error data. Immediately after a search task, their subjects reported all the stimuli they could from the display. They were more likely to report the stimuli near the search target than those that shared the search property (e.g., color) with the target. A subsequent study by Tsal and Lavie (1993) with tasks in which the target and nontarget stimuli appeared in successive frames produced similar evidence of selection by location. However, van der Heijden, Kurvink, de Lange, de Leeuw, and van der Geest (1996) have raised doubts about Tsal and Lavie's (1988) conclusions. As pointed out by Tsal and Lavie themselves, the location effects in their study could have been due to subjects' shifting fixation to a random array position before stimulus onset. Van der Heijden et al. repeated the experiment by Tsal and Lavie, this time ensuring that subjects were fixating at the center of the display by presenting a stimulus to be discriminated at that location. Subjects were more likely to report additional letters of the target color than additional letters adjacent to the target, indicating no special role for location in this task. Even if Tsal and Lavie's results can be attributed completely to eye movements, however, Cave and Pashler's (1995) and Kim and Cave's (1995) experiments demonstrate that search for a color target results in location-based selection even when the procedure includes no incentive and no time to program eye movements.

The special role of location in visual selection is also supported by investigations of the neural mechanisms underlying visual selection. In one set of such experiments, Luck, Hillyard, and colleagues (Luck, Fan, & Hillyard, 1993; Luck & Hillyard, 1995) have measured event-related potentials (ERPs) during visual search tasks in which targets were defined by features other than their location. In these experiments, visual search displays were followed by probes presented around the target or distractors, and attentional effects were assessed by measuring ERPs to the probe stimulus. The results of Luck et al. (1993) indicate that, during a color/shape conjunction search, probe stimuli at the target location elicit enlarged P1 and N1 components relative to those elicited by probes at a distractor location. These ERP effects were similar to those obtained in cuing experiments (see Mangun & Hillyard, 1995, for a review), suggesting a role of location in visual selection beyond cuing experiments. In a later experiment, Luck and Hillyard (1995) investigated whether the same attentional mechanisms are employed in both feature and conjunction visual searches. Their results indicate modulation of the N1 component even in feature searches, supporting a more general role for spatial selective attention. However, the P1 attentional effect was not found in the feature detection condition, suggesting that the process reflected by the P1 component may be required only when distractors are highly confusable with the target (see Mangun & Hillyard, 1991; for related results, see Luck et al., 1997, and Moran & Desimone, 1985).

Connor, Preddie, Gallant, and Van Essen (1997) investigated the effects of attention in area V4 using single-cell recordings in macaque monkeys. The results of this experiment are particularly important in assessing the importance of location in attention, because area V4 is generally regarded as belonging to the visual pathway concerned with representing objects rather than locations. Monkeys were required to attend to a ring outside a neuron's receptive field in order to detect a change in its circumference. While monkeys were attending to this ring, neural responses to task-irrelevant bars presented within the receptive field indicated attentional enhancement of neural responses to behaviorally irrelevant objects near the target. This result and the ERP findings mentioned earlier are consistent with behavioral observations that subjects are more accurate or faster at reporting objects near the target stimulus (e.g., Hoffman & Nelson, 1981; Kim & Cave, 1995; Snyder, 1972; Tsal & Lavie, 1988, 1993), because they all support the view that attention is allocated to the region of space occupied by a behaviorally relevant object. The fact that location affects selection even when it is not directly relevant to the task implies that location plays a special role in visual selection.

The discussion so far has assumed that the visual system limits its work by using some simple visual property, such as location, to eliminate a large part of the input early in the stream of processing. However, some researchers have argued that such a limitation is unnecessary. Despite such strong evidence for a role of location in visual selection, many researchers have suggested that visual selection is accomplished from spatially invariant object representations that are formed in the later stages of visual processing. These claims are built on the assumption that the visual system is powerful enough to classify

or identify many visual objects in parallel and that selection is necessary only in later stages, when visual information is used to guide responses. Under this view, selection operates not on raw, unprocessed, spatially organized visual input but on an abstractly organized symbolic representation of the objects identified in the visual input. Just what is selected from this representation depends on how it is organized, and the spatial layout does not determine the organization any more than other aspects of the stimulus. Location is just another property to use as an index in selecting from among these representations, and location affects selection only because it affects perceptual organization (Duncan 1981, 1984).

Some of the strongest evidence for object-based selection came from Duncan (1984), who presented subjects with small foveal displays consisting of two overlapping objects. Each of the objects could vary on two dimensions. Performance was similar when subjects reported either one or two dimensions of the same object, but it declined significantly when subjects reported two dimensions of two different objects. Duncan (1984) concluded that selection under these conditions is object-based, because of the cost in selecting both objects rather than just one. He further argued that this result could not be explained by a location-based account of selection since both objects occupied the same location within a 1° visual angle foveal region that would be selected by an attentional spotlight (C. W. Eriksen & Hoffman, 1972, 1973). A number of other studies using other methods have produced additional evidence for selection of objects (e.g., Baylis & Driver, 1993; Chen, 1998; Driver & Baylis, 1989; Egly, Driver, & Rafal, 1994; Harms & Bundesen, 1983; Kahneman, Treisman, & Gibbs, 1992; Treisman, 1982, 1992; but see Gibson, 1994).

Lavie and Driver (1996) replicated Duncan's (1984) object-based selection with a different set of superimposed stimuli, but they also demonstrated the importance of spatial selection in this task. Using a large display of two overlapping lines, they found an advantage for reporting two targets positioned on the same object over two targets on different objects, just as Duncan would predict. However, when subjects were cued to attend to one side of the display, the object-based attention effects disappeared. On the basis of this interaction between location-based and object-based attention, Lavie and Driver concluded that "selection ultimately takes place within a spatial medium, albeit under the influence of object-based factors."[2]

There is additional evidence that object-based selection with superimposed stimuli may be mediated by location-based selection. Cave and Kosslyn (1989) performed another experiment that required subjects to attend to a target object superimposed over a distractor object, but, in this test, the size of the objects varied from trial to trial. RTs increased with the change in size from one trial to the next, and the pattern of RTs was similar to that produced when subjects are adjusting the size of a mental image. A mental representation like those used in visual imagery might allow subjects to select precisely those locations occupied by contours of the target object without selecting contours of the distractor object. The drop in performance that comes from switching from one object to the other in Duncan's (1984) experiment could reflect the extra effort necessary to create a new image. In this instance, selection could be executed within an early, unanalyzed representation even though it is driven by a higher level, conceptual representation.

Vecera and Farah (1994) considered a similar explanation for Duncan's (1984) results. They raised the possibility that spatial attention may conform to an object's shape by selecting precisely the spatial locations occupied by the objects. They described this alternative mode as selection from a "grouped-array" representation. Thus, the poorer performance in Duncan's experiment when subjects made judgments about both objects may be attributable to a cost in activating a different group of locations rather than selecting a different object representation. Vecera and Farah set out to test this hypothesis by replicating Duncan's results and comparing them to a condition in which the two objects were on opposite sides of fixation. They hypothesized that if selection in Duncan's experiment was indeed location-based, there should be an increase in the effect size when the objects are placed apart from each other, due to the larger distance attention has to travel in this condition. However, they found no interaction between the two conditions and concluded that Duncan's results were due to visual selection from spatially invariant object representations, as postulated by Duncan.

However, the design and interpretation of Vecera and Farah's (1994) experiment is based, as stated by the authors, on the assumption that there is an increased cost associated with attention moving across larger spatial distances, which is far from certain in light of much evidence showing that attentional shifts are time-invariant (see Section 3). Kramer, Weber, and Watson (1997) noted this potential confound in Vecera and Farah's interpretations of their results and tested directly, using spatial probes similar to those in Kim and Cave (1995), whether selection in the Duncan (1984) and Vecera and Farah (1994) experiments was from grouped-array or spatially invariant object representations. In the separate-object condition, the two objects were on opposite sides of fixation; in the superimposed-object condition, both stimuli were presented on one side of fixation in an overlapping manner, and an irrelevant filler object was presented on the other side of fixation. Responses to a postdisplay probe showed that subjects in the separate-object condition responded faster to the probe when it appeared at the location previously occupied by an object whose properties were to be reported than when the probe appeared at the location of the object to be ignored. Furthermore, in the superimposed-object condition, subjects responded faster to the probe at the location of the objects than at the location of the filler. These results clearly show that attention is allocated to the locations of relevant objects, at least in this experiment, as acknowledged by Vecera (1997). The findings of this experiment are

further supported by Weber, Kramer, and Miller (1997), who combined the postdisplay probe technique with ERP recordings (e.g., Luck et al., 1993). Two partially overlapping objects (i.e., colored dashed outlines of shapes) were presented to the left or the right of fixation, and subjects were required to detect the presence of a combination of particular color and shape, whether the color and the shape both belonged to the same object or to different objects. On some of the trials, a small gray square probe that did not require a response was presented at one of the two potential stimulus locations on opposite sides of fixation. ERP analysis revealed that the N1 component elicited by the objects, and the P1 elicited by the probe when it was at the location of the objects, distinguished between same- and different-object conditions. Because, as mentioned earlier, the P1 and N1 components appear to be associated with spatial selection mechanisms, Weber et al. interpreted the differences in these components to indicate the spatial nature of the selection process.

One way to reconcile the distance effects on attention with an object-based account is to assume that distance affects the way visual objects are organized into groups and that attention is applied to representations of object groups. Evidence for this type of account comes from demonstrations of nonspatial grouping effects on attention. A classic example of nonspatial grouping was presented by Harms and Bundesen (1983), who studied the relation between selective attention and perceptual segregation by color in binary-choice RT tasks based on the interference paradigm of B. A. Eriksen and C. W. Eriksen (1974). Subjects were presented displays containing a central target letter alone or flanked on each side by a noise letter. The target was one of two letters, and the distractors could be one of a group of three that included the two target letters. Furthermore, each noise element could have the same color as or a different color than the central target element. The subjects' task was to identify the central target letter without any advance knowledge of the color of that element. The most important finding of the study was that noise letters interfered more with identification of the central target letter when they shared the target color.

Using a similar paradigm, Baylis and Driver (1992) showed that noise letters with the target color interfered more than those of a different color, even if the different-color distractors were closer to the target. In another study, Driver and Baylis (1989) found that distractor letters interfered more with identification of a target letter when they were far from the target but moved with the target than when they were closer to the target but did not move with it. Grouping by coherent motion and grouping by color both had stronger influences than did spatial proximity.

It is important, however, to realize that the results of these grouping studies do not preclude the possibility that grouped objects, whether by color, coherent motion, or some other feature dimension, might have been selected by selecting the locations that they occupied. When deciding the level at which selection is likely to occur in a particular task, it is important to consider two different questions separately. First, at what level is it actually determined what information should be selected? Second, at what level is the selection actually accomplished? The decisions about what to select may be made using high-level, conceptual representations, but these decisions may drive mechanisms that separate selected from unselected information by operating on raw, spatially organized inputs at a lower level. In other words, although decisions about what to select may be made on the basis of many different factors, including perceptual grouping by features such as color and motion, the selection might still be accomplished with one or more spotlights.

A recent study by Kim and Cave (1999) addressed this issue by measuring RT to postdisplay probes in a task modeled after Harms and Bundesen's (1983). As in Harms and Bundesen's study, the primary task was the identification of a letter in the center of the display, which was flanked by two distractor letters, one on each side of the target letter, and one sharing the target's color and the other of a different color. Note that color was, as in Harms and Bundesen's study, irrelevant to the correct response. On half the trials, a small black probe was presented at a location previously occupied by one of the distractors; the stimulus onset asynchrony (SOA) between onset of the letters and onset of the probe was selected randomly for each trial to be either 60 or 150 msec. At the 60-msec SOA, subjects were faster at responding to the probe at the location previously occupied by the distractor with a color different than the target. At the 150-msec SOA, the pattern of RT was reversed, with an advantage for the probe at the location of the distractor with the same color as the target. These results demonstrate that the attention allocated to distractor locations varied as a function of their color with respect to that of the target. In the early SOA, attention is preferentially allocated to the location of the uniquely colored item (singleton), consistent with Theeuwes's (1991, 1992) findings that attention is captured by singletons, even if the feature that makes them a singleton is not relevant to the task. However, at the late SOA, the location of the distractor with the target color is preferentially selected, which implies perceptual grouping (based on color) mediated by spatial selection.

To ascertain whether the initial preferential allocation of attention to the singleton distractor location was due to attentional capture by a singleton, Kim and Cave (1999) performed an additional experiment using displays that did not contain a color singleton. They presented subjects with four letters around a central fixation point. For each trial, the target location was cued by a black outline box at that location. The target and distractor letters could each be either red or green. One of the distractors adjacent to the target had the same color as the target, and the other two distractors had a different color. The same SOAs as in the previous experiment were used. For both SOAs, responses were significantly faster to probes at locations previously occupied by the target and the distractor with the target color than to probes at nontarget-color distrac-

tor locations. Together, these two experiments suggest that the interference effects from color-grouped distractors observed in Harms and Bundesen (1983) and Baylis and Driver (1992) are due to spatial selection that favors the location of the grouped distractor as well as that of the target.

Perceptual grouping has been shown to affect visual selection in a number of other ways. For instance, Prinzmetal and Millis-Wright (1984) found more illusory conjunctions among the features of words and pronounceable nonwords than among the features of unpronounceable nonwords, presumably because the words and pseudowords formed perceptual units. A number of studies have specifically tested grouping by proximity, showing that visual stimuli are more likely to be selected together if they are positioned near one another to form a perceptual group. Kahneman and Henik (1977, 1981) had subjects report digits that were organized into two groups. The probability of correct report was generally the same for all digits within a group and was generally very different for digits in different groups. Treisman (1982) demonstrated that arranging elements into groups affected conjunction search but not feature search. Her results suggested that subjects examined an entire group all at once and that processing moved from one group to another serially. Prinzmetal (1981) showed that two features were more likely to be combined into an illusory conjunction if they were both part of the same group, even if the distance between the two was controlled. Treisman, Kahneman, and Burkell (1983) also controlled distance in a task that required subjects to read a word and find a gap in a line. Performance was better on both tasks when the word and the line were both part of the same perceptual object.

On one hand, there is solid evidence for location and distance effects on visual selection; on the other hand, there is solid evidence for effects of grouping and object organization. There could easily be multiple mechanisms for visual selection operating independently on location-based and object-based representations (see Vecera, 1997). However, it may still be possible to explain all of the results described here with a single mechanism. All of these results are consistent with spatial selection that is driven in part by the visual organization of the objects in the visual field. One might be tempted to assume that distance and location play no direct role in visual selection, but that these factors only determine how elements in the scene are organized into groups, which are then selected by attention. Indeed, if we assume that objects that are close together are always grouped together and that stimuli that appear at or near the same locations (e.g., Hoffman & Nelson, 1981; Kim & Cave, 1995; Tsal & Lavie, 1993) are organized into groups, then all the results, including the distance effects, might also be explained with selection of abstract representations of object groups.

Such a claim, however, is almost unfalsifiable. How do we distinguish between early selection by location and late selection by perceptual grouping? One important fac-

tor in comparing the two accounts is their relative simplicity and parsimony. The late selection account requires that the early stages of visual processing be powerful enough to identify many visual objects simultaneously. The early selection account makes many fewer demands on early visual processing, because selection is based on simple visual properties that are computationally easy to detect. In that sense, it would be simpler to attribute distance effects to early selection by location. However, the perceptual grouping experiments described above suggest that grouping plays a role in selection, even when distance is held constant. If there has to be a late selection mechanism sensitive to grouping, is it more parsimonious to attribute all selection to that mechanism or to postulate a separate mechanism operating at an earlier stage as well? Selection by location offers simplicity on one front but added complexity on another.

Although it is very difficult to distinguish between grouping by location and location-based selection, a recent study by Cave and Zimmerman (1997) provides strong evidence that the effects observed in probe studies (e.g., Kim & Cave, 1995) are indeed due to location-based selection mechanisms. In this study, the spatial probe technique was applied to measure changes in attentional allocation during a letter search task after practice. Subjects were slower at responding to the probe when it appeared at a distractor location near the target than when it appeared at distractor locations farther from the target. This response pattern, reminiscent of flanking inhibition observed in the activity of neurons (e.g., Schall et al., 1995), is clearly highly spatial in nature and is consistent with location-based selection but is not expected from grouping-by-location accounts in which location-independent representations are selected.

Shih and Sperling (1996) used a different technique to test for location-based selection of a target defined by color or size, producing another set of results that are difficult to explain by proximity-based grouping. Their subjects detected a target digit among multiple frames of superimposed letter arrays. In one condition, all stimuli within a frame were of the same color, with the color alternating in successive arrays. In the other condition, one stimulus in each array differed from the others in color. Subjects could better detect the target digit when it was the color they expected, but only when it was in a frame in which distractors were all of a different color. In this condition, the attended feature provided spatial information about the target. On the other hand, when all the characters in a single frame were of the same color, and the attended feature provided temporal but not spatial information, subjects were unable to use information about the color of the target to improve performance. The same observations were true when size information was used instead of color. Shih and Sperling concluded that subjects used color or size information to determine which location to select.

In summary, the results of a large number of attentional studies, including spatial cuing studies, neurophysiological studies, and studies that did not emphasize or encour-

age selection by location, show that location plays a crucial role in visual attention. Furthermore, some of the key studies demonstrating object-based selection have been shown to have strong spatial components, consistent with object selection that is accomplished or mediated by selection of locations (selection from grouped arrays). However, the data obtained to date do not rule out the possibility that both space-based and object-based mechanisms exist.

2. When Attention Shifts From One Location to Another, Does It Pass Over the Locations in Between?

A real spotlight cannot shift from one location to another without covering other locations along the way. Shulman, Remington, and McLean (1979) tried to show that attention shifts across the visual field in the same manner. They started each trial with a cue indicating where a probe was most likely to appear. They assumed that, once the cue appeared, attention would move from the fixation point to the cued location. Using a variety of delays between cue and probe, they tried to catch attention as it moved by occasionally probing at a location between the fixation and the cued location. The cued location was called the "far expected" location, because it was far from the fixation; the "in-between" location was called the "near expected" location, even though the probe was never expected at that location. They also occasionally probed two locations in the opposite hemifield from the one containing the cue: one at the location in that hemifield corresponding to the cued location (far unexpected) and the other corresponding to the in-between location (near unexpected).

The most striking pattern in the data is a general alerting effect. As SOA increased, RTs for all four probed locations dropped and then rose again after an SOA of about 400 msec. When this effect is removed, the two locations on the expected side showed gradually increasing facilitation (decreasing RTs), whereas the two on the other side showed increasing inhibition (increasing RTs). On the expected side, the facilitation for the near location jumped up sooner than for the far location, leading Shulman et al. (1979) to conclude that an attentional spotlight was passing over the near location on the way to the far location.

For many years, this study was the only one to demonstrate attentional facilitation traveling through intermediate points, and its conclusions have been questioned (C. W. Eriksen & Murphy, 1987; Yantis, 1988). Although the difference between RT for the near and far expected positions matches the predictions, the overall pattern across all four probed locations is difficult to reconcile with any sort of moving spotlight. As Yantis points out, the expected pattern of RTs varies depending on whether the hypothesized spotlight is focused narrowly or spread over a large area and whether it must be centered on the probe before a response can be made. No single set of assumptions seems to account for both locations on the unex-

pected side. RTs from the near unexpected location rise relatively little with SOA, suggesting a narrow beam that does not affect the uncued side much. However, RTs for the far unexpected location increase dramatically, as if a very large gradient were gradually moving farther and farther away from it.

The results from Shulman et al. (1979), taken as a whole, do not suggest a moving spotlight, but it is not clear what they do suggest. It is almost as if two different types of mechanisms were at work: one in the cued hemifield, and the other in the uncued hemifield. The side receiving a cue receives facilitation that starts at fixation and gradually spreads to the periphery. On the side without the cue, a strong inhibition gradually covers the periphery, but only a mild inhibition affects the center. These data are perplexing and demand further experiments.

Note that these results demonstrate attentional differences within a hemifield, contrary to Hughes and Zimba (1985, 1987). The possible probe locations were not marked, but the stimuli were delivered by four LEDs mounted in front of a CRT. If the unlit LEDs and the mounting were visible, they may have provided landmarks similar to the squares in Downing and Pinker's (1985) displays.

Murphy and C. W. Eriksen (1987) addressed the question of a moving spotlight with methods similar to those used in earlier experiments by C. W. Eriksen and his colleagues. Their subjects had to decide which of two letters appeared at a cued location, and the SOA between cue and letter varied from 0 to 175 msec. Rather than occasionally probing at uncued locations, Murphy and C. W. Eriksen added an extra distractor letter at one of the unprobed locations in every trial. As mentioned above, previous experiments had shown that when such a distractor was associated with the incorrect response for that trial, it raised the RT. Murphy and C. W. Eriksen reasoned that the interference from this distractor should increase if the spotlight passed over it on its way to the target. They found no evidence for increased interference, but they examined only the data from the 175-msec SOA. By that time, a moving spotlight might have passed the intermediate locations and arrived at the cued location. Their trials with intermediate SOAs should provide useful data, but they were not examined in a way that would answer this question.

Further evidence for discrete, rather than analog, attentional shifts is provided by a series of experiments by Chastain (1992a, 1992b). In one experiment, Chastain (1992a) used a dual-task paradigm to test whether attention passes over intermediate locations before reaching its final destination. Targets could appear at four locations around a fixation point, and a probe that could appear halfway between fixation and a target location was used to measure attentional allocation at such intermediate locations. Chastain (1992a) measured sensitivity in discriminating target and probe during both central and peripheral cuing paradigms. He found that at no SOA between target and probe presentation was sensitivity greater

at a probe location between target and fixation than it was at the probe location diametrically opposite from fixation, suggesting that attention was shifted from fixation to target location in a discrete manner.

Chastain (1992b) used a different approach to arrive at the same conclusion. Two locations 14° apart were cued, one 200 msec after the other. The target could appear after either cue, and it was twice as likely to appear at the location of the immediately preceding cue. The target was presented at a variety of intervals after the first or second cue, and, as expected, there was a tradeoff in accuracy at the two cued locations. To determine whether attention shifted in an analog manner, covering the area between the cued locations, the average accuracy in target identification at the two cued locations was compared with baseline accuracy. Chastain (1992b) calculated the mean accuracy at the two cued locations during the attentional shift, when attention should be moving from one to the other. The mean remained well above baseline, and it did not vary with target–cue interstimulus interval. These results argue against an attentional spotlight "passing over" intermediate locations, which would have predicted a decrease in average accuracy at the two cued locations at some point during the attentional switch while the spotlight is at a location between the two cued locations.

Although there is still disagreement about whether intermediate locations are activated as attention shifts from one location to another, evidence to date appears to suggest that attentional resources are allocated to the new location as they are simultaneously deallocated at the old location, without passing over regions in between. Given current understanding of the physiology of visual attention (Motter, 1994; Schall et al., 1995), an attentional shift could plausibly be implemented in the brain by decreasing the activity in a pool of neurons with receptive fields covering the location previously selected while increasing the activity in a pool of neurons with receptive fields covering the new location to be attended.

3. How Is the Time Required to Shift Attention Related to the Distance Between the Old and New Positions?

Not only does a moving spotlight pass over the locations between its origin and destination but, assuming that it always moves at a constant velocity, it requires more time to move over longer distances. Tsal (1983) tested whether the beam of the attentional spotlight also requires more time for longer shifts. In this experiment, each trial began with a cue indicating that the target would appear at one of three locations to the left of fixation or one of three to the right. All six locations were along the horizontal midline. After a varying SOA, a target letter (either X or O) appeared at the cued location, and the subject vocally identified the target. This experiment differed from most cuing experiments in that subjects performed a discrimination rather than a simple detection. For each of the three fixation–target distances, RT decreased steadily

as SOA increased, up to a point, and then seemed to asymptote. In other words, when subjects had more time to prepare after the cue, they required less time to respond after the target appeared. Tsal assumed that subjects used this preparation time to move the attentional spotlight from the fixation point to the cued location. In his view, the time at which the RT asymptotes corresponds to the time that the spotlight arrives at the target location. As expected, the RT asymptotes later when the distance between fixation and target is longer.

The researchers who criticized Shulman et al. (1979) also criticized Tsal's (1983) study. Yantis (1988) pointed out that most cuing studies include a condition with a cue that indicates that a target is about to appear but provides no information about its location. Such a cue produces a general alerting that temporarily lowers RT for targets at all locations. Comparing these control trials with the location cue trials reveals which effects are location-specific and can therefore be attributed to selection by location. Tsal did not include a nonspatial cue; thus, it is impossible to determine from his data whether the cue is producing facilitation for a particular location or a general alerting. Of course, if it is a general change in alertness, there is still a problem in explaining why it produces different effects at different times for locations at different distances from fixation.

Yantis (1988) also showed how a model in which attention "jumps" from one location to another could produce data similar to those of Tsal (1983). He assumes that the time necessary for the jump varies from trial to trial. For longer SOAs, the probability will be higher that attention has jumped to the target location before the target appears, and the mean RT from many trials averaged together should be lower. Tsal's steadily decreasing RT functions could reflect a steadily increasing proportion of trials in which jumps had been completed. In Yantis's (1988) model, the functions for peripheral locations asymptote later because subjects take longer to detect cues at these locations, so the attention shifts begin later. Tsal argued against delays in cue detection with a control experiment in which subjects reported peripheral cues just as quickly as more central cues. Yantis (1988) rejects this control, although he does not say why.

C. W. Eriksen and Murphy (1987) raise similar objections to Tsal's (1983) experiment and add others as well. They point out that other processes besides attention shifts and cue detection might vary with eccentricity. They also explain that different subjects would probably vary somewhat in the speed with which the attentional spotlight traveled across the visual field. There might also be variations for a single subject from trial to trial. There would also be variation in other aspects of the task, such as identifying the cue and initiating the attention shift. When the data are averaged together, the result should be a negatively accelerating function, rather than one that drops off steadily and then flattens out. They also question the data from the control experiment in which the time to report the cue is constant regardless of eccentricity. Given

these points, it is rather puzzling to imagine what sort of process could produce the data that Tsal obtained, and C. W. Eriksen and Murphy offer no clues to solve this mystery.

C. W. Eriksen and Murphy (1987) offer their own jumping spotlight model that is generally consistent with Tsal's data. They reject Tsal's (1983) assumption that subjects cannot respond to a target until the spotlight reaches its location; instead, they assume that processing can proceed without attention but at a slower rate. A stimulus will be identified more quickly when there is more time to focus attention at its location. Stimuli in the periphery are harder to identify, and, thus, they have more room to benefit from attention.

Murphy and C. W. Eriksen's (1987) experiment described in Section 2 also addressed the question of distance effects. Their task differed from Tsal's (1983) in many ways; however, like Tsal's, their targets appeared at one of three distances from fixation, and they varied the SOA between cue and target. RT decreased with longer SOAs, but, unlike Tsal's data, the RT functions flattened out gradually, and the rate at which they leveled off was about the same, regardless of distance from fixation. There was no indication that the spotlight arrived at near locations faster than at far locations.

In a subsequent study, C. W. Eriksen and Webb (1989) cued multiple locations in eight-letter circular displays while varying the number of cued locations, as well as the compactness of the cued region (i.e., the number of intervening noncued locations). In one experimental condition, locations were cued prior to the presentation of the search display; in the other condition, one location was cued prior to display presentation and another location was cued simultaneously with display presentation. Subjects' RT revealed a cost in increasing the number of cued locations, supporting a serial search strategy. However, there was no evidence of effects associated with distance between cued locations, implying that shifting attention across longer distances does not require a longer amount of time.

Remington and Pierce (1984) also performed experiments similar to Tsal's (1983) experiment and also reached a conclusion that differed from his. They identified the same problems in Tsal's experiment that Yantis described, and they corrected them by using central, rather than peripheral, cues and by including a condition with targets at an uncued location to control for general alerting. They used a simple detection task. The only major difference from the original Posner paradigm was that, in some sessions, the two possible target locations were 2° from fixation, and, in some sessions, they were 10°. Like Tsal, they used a variety of SOAs between cue and target to measure the time course of attentional facilitation. A constant velocity spotlight would predict that the facilitation would appear at the near locations before the far locations, but, in fact, the facilitation developed at the same rate, regardless of distance. One might conclude from this result that the spotlight jumps from location to location, skip-

ping the intermediate locations. Remington and Pierce, however, were convinced of the sliding nature of the spotlight by the results from Shulman et al. (1979). They concluded that the spotlight does slide from one location to another but that its velocity is adjusted so that all shifts take the same amount of time. They justified this conclusion by pointing to hand movements and saccades for which velocity varies with distance. Given the questions raised about the conclusion from Shulman et al., however, it seems more plausible that attentional facilitation jumps from one location to another at a constant rate and without passing through the intervening locations.

Sagi and Julesz (1985) approached this question in a different way. In each trial of their experiment, subjects saw two tachistoscopically presented letters, each either a T or an L rotated by a certain amount. The task was to decide whether the two letters were the same, either both Ts or both Ls. Sagi and Julesz varied the distance between the two letters while keeping each letter a constant distance from fixation. They chose these stimuli because they believed that the similarities between the two letters would force subjects to focus their attention first on one and then on the other, and they offered evidence that subjects were indeed using such a serial strategy. They found that accuracy diminished when the two targets were very near each other, which would be expected because of lateral masking. Otherwise, distance had little effect on accuracy. There was no evidence for a spotlight that took more time to move over longer distances. In a later study, Kwak, Dagenbach, and Egeth (1991) also found that RT in a *same–different* matching task with rotated Ts and Ls did not increase with separation between letters. Kwak et al.'s study differed from the earlier study by Sagi and Julesz in that they avoided using postmasks to rule out potential interpretational complications, they used more subjects, and they presented additional evidence for serial processing.

Finally, Sperling and Weichselgartner (1995) provided additional evidence for time-invariant attentional shifts using a different approach. In one variant of the experiment, they presented subjects with two streams: one of numerals presented foveally, and another of letters presented parafoveally. The attentional task required the subjects to report the earliest possible digit in the numeral sequence following the detection of a target letter in the parafoveal stream. The motor task required them to release a response key upon detection of the target letter. A preliminary experiment showed that RT in the motor task was unaffected by a concurrent attention task, whereas the reverse was not true. Thus, the motor task RT was used to index target detection. In a subsequent experiment, the distance between the letter and the numeral streams was manipulated, and motor RT was subtracted from attentional RT to obtain a measure of attentional shift latency from periphery to fovea independent of target detection time (which increases with retinal eccentricity). The results revealed no effect of distance on attentional shift latency. Interestingly, the presentation of noise symbols

between the two streams at the 7° separation did not affect the measure of attentional shift, consistent with the conclusion of the previous section that attention does not travel trough space like a real spotlight.

All in all, there is little evidence that the spotlight "slides" across the visual field in an analog fashion. Along with evidence that attention does not pass over intermediate locations during attentional switches it seems that if there is something like a spotlight of attention, it probably does not move like a spotlight but instead "jumps" from location to location. As mentioned in the previous section, such discrete attention shifts seem to be generally consistent with the physiological evidence.

4. What Is the Size of the Selected Area and How Much Flexibility Is There in Adjusting That Size?

Because visual stimuli can vary greatly in size, the size of a visual selection spotlight might need to be adjusted to accommodate them. It would also be useful to adjust the size of the spotlight according to the expected location of an upcoming stimulus. If it were known that the stimulus would appear at one exact spot, then a small spotlight around that spot would most effectively eliminate interference. If the stimulus were expected somewhere within a large region, then a large spotlight could cover the entire region. Also, the spotlight might be a mechanism for size constancy, just as it might produce position constancy. If the spotlight encompasses a small object, the selected information could be scaled up to a standard size. Likewise, a large selected object could be scaled down. Only a few studies have addressed the role of size and area in visual selection.

Experiments by Larsen and Bundesen (1978) and by Cave and Kosslyn (1989) measured RTs when subjects expected a stimulus to appear at one size and it actually appeared at a different size. Larsen and Bundesen's subjects judged whether a letter was rightside up or upside down. Cave and Kosslyn's subjects judged whether a rectangle was a square. In one sense, these studies are analogous to those studies in which a location is cued before the stimulus appears. Rather than starting each trial with an expectation concerning the stimulus location, these subjects started with an expectation concerning its size. In both studies, RT increased with the ratio between the expected and the actual sizes, as if a selection mechanism were adjusted in an analog fashion from one size setting to another.[3]

These two studies provide evidence against an earlier conclusion by Egeth and Berninger (reported in Egeth, 1977, pp. 301–303). In each trial of Egeth's experiment, subjects saw two digits, often of different sizes. The time to decide that two digits were the same increased with size ratio, but the time to decide that two digits did not match decreased with size ratio. Egeth decided that comparing stimulus pairs was not the best way to attack this question. He then tested RTs to naming a series of digits presented individually and found subjects were just as good whether the digits were all the same size or of different sizes. He concluded that subjects do not prepare in advance for stimuli of a particular size.

The effects measured in Larsen and Bundesen's (1978) and Cave and Kosslyn's (1989) experiments were rather subtle and might have been missed in Egeth's (1977) experiments. Egeth only used ratios ranging from 1:1 to 4:1, whereas Larsen and Bundesen and Cave and Kosslyn used ratios ranging from 1:1 to 9:1. From these last two studies, it appears that subjects can, at least under some circumstances, set themselves to process stimuli of a particular size and that the more a stimulus deviates from the chosen size in either direction, the longer it takes to identify it.

These experiments suggest that the visual system can select information at a particular size or scale, just as it can select information from a particular location. However, there is little indication that the mechanisms responsible for the size effects described here are the same mechanisms underlying the location effects described earlier. These different effects may result from different mechanisms operating at different stages of processing. No distractors were present in these experiments; thus, there was no need for the filtering that seems to be important in many other selection experiments. More work is needed to determine whether or not these size effects come from adjusting the size of the same "attentional spotlight" that is at work in experiments manipulating location.

LaBerge (1983) demonstrated size effects of a different sort by using two tasks in conjunction: one to control the area over which attention was allocated, and the other to probe for the effects of this allocation at different locations. In the first task, the subjects either judged whether a five-letter noun was a proper noun or judged whether or not the middle letter of a five-letter string came from a certain range of the alphabet. Occasionally, instead of the regular word or letter string, the stimulus would be a string consisting of some combination of four #s and a single test character (7, T, or Z). The RT on these catch trials was a measure of the focus of attention at each of the five possible locations of the test character. When subjects were set to read the entire word, the position of the test character in the catch trials made little difference. When they were set to focus on the single letter in the middle, RT was fastest for characters in the middle and slowest for characters on the end.

This experiment shows how attention can be focused narrowly or spread over a larger area, and it raises the question of whether the benefits of attention are weakened when it is spread. In LaBerge's (1983) experiment, there does not appear to be any such tradeoff. An experiment by Egeth and Shwartz (reported in Egeth, 1977, pp. 294–297) suggests otherwise, however. They measured RT in a letter discrimination task. Responses were slower for a letter in the center of the display when the letter could have appeared at a number of locations within a large area than when the letter could only have appeared at the center location. Responses were slower when attention was spread over a larger area.

The same questions motivated a study by C. W. Eriksen and St. James (1986). As in many of C. W. Eriksen's experiments, the target was a letter that appeared along with seven distractor letters in a circle around the fixation point. From one to three contiguous positions would be cued before the letters appeared (or, in some trials, very soon thereafter), and the target would always appear in one of these cued positions. RTs increased with the number of cued positions. C. W. Eriksen and St. James argue that when more positions are cued, attention must be spread out over a larger area, and fewer resources are dedicated to each location. However, more cued locations also mean that more distractors appear at cued locations, and these cued distractors may interfere with identification of the target. The authors are not very concerned about this alternative explanation because they can offer evidence that the subjects do not process each of the cued locations individually and serially. This is easy to believe because the distractor letters are always very different from the target letters. Nonetheless, at the end of their paper, C. W. Eriksen and St. James admit that "part of the increase in RT with an increase in cued positions must be attributed to a change in the discriminative difficulty of the task," and it is indeed unclear how much of the effect should be attributed to the extra attentional demands.

Castiello and Umiltà (1990) followed up on C. W. Eriksen and St. James's (1986) experiments in a cuing task similar to that of Posner et al. (1980). They presented subjects with box outlines that cued the location of a subsequent stimulus to be detected, which always appeared at the center of the box (see Section 5). The RT benefits from the cue decreased with increasing cue box size, confirming C. W. Eriksen and St. James's view that efficiency of processing decreases with increasing area of attentional focus.

Taken together, the studies by LaBerge (1983), Egeth (1977), and C. W. Eriksen and St. James (1986) suggest that the size of the spotlight can be adjusted according to the demands of the current task, and the study by Castiello and Umiltà (1990) suggests that there may be a cost associated with spreading the spotlight over a larger area.

5. Once an Area Is Selected, Is Everything Within That Area Processed Equally, or Does the Central Region Receive More Thorough or Effective Processing?

Behind this question lies a deeper one: What purpose does the spotlight serve in visual processing? It might be an obligatory step in processing, so that a stimulus would always have to be selected by the spotlight before it could be passed on to higher level processing stages and trigger a response. Such an obligatory spotlight would have to move quickly and efficiently to capture all the important parts of a complex scene. It would probably be small enough to select only one object or a few objects. Because one object or a few objects would need to

be fully processed while the remaining objects are completely excluded, the spotlight would have a sharp, well-defined edge. Such a sharply defined edge would be expected if the spotlight were responsible for position and size constancy.

On the other hand, the spotlight may not exert such complete control over visual processing. It might only facilitate the analysis of certain regions. Objects that were located outside the area of such a facilitatory spotlight might still be processed and identified, although not as efficiently. If there is not such a clear distinction between selected and unselected objects, the spotlight might not have such a sharply defined edge. Facilitation could be strong in the center of the selected region and drop off gradually with distance. Such an attentional gradient could cover a large area. Because it is not as crucial to have the most important visual object always in the center of the spotlight, this facilitatory spotlight might move more slowly when it changes position. A spotlight that was merely a facilitator could still play an important role in filtering out distractors, but additional selection mechanisms would be required at later stages of processing.

Either the obligatory or the facilitatory spotlight can be made to fit much of the data on attention shifts. For instance, consider Downing and Pinker's (1985) finding that RT increases with distance between cued location and target location. Before the target appears, the spotlight will be at the cued location. With an obligatory spotlight, the subject cannot respond to the target until the spotlight has moved to its location, even when the target is a simple spot of light. Assuming a sliding spotlight such as Tsal's (1983), the longer RT with longer distances reflects time necessary to move the spotlight.[4]

Downing and Pinker's (1985) data also fit with a spotlight that is merely facilitatory and that moves more sluggishly. Such a spotlight would cover a large area, with a wide range of effectiveness between its center and its edge. The spotlight would be centered on the cued location. Because that location received the most facilitation, targets there would trigger very fast responses. Targets farther from the probe would receive less facilitation and produce longer RTs. If spotlight shifts require a fair amount of time, then it is better to leave the spotlight positioned as it is, even when targets are far from the center, and to make do with little attentional facilitation. Of course, Downing and Pinker's subjects might choose to readjust the spotlight for some distances, or they may be able to move it as they are identifying the target. Their data might be the result of some combination of facilitation and moving time.

The obligatory spotlight requires that long moves take more time, and, as we noted above, there is little evidence for such a pattern. Also, C. W. Eriksen & Murphy (1987, p. 304) point out that Shulman et al.'s (1979) data from the cued side are not consistent with an obligatory spotlight (although Yantis, 1988, pp. 204–205, fails to see a conflict). However, their data, taken as a whole, are difficult to reconcile with any single model of attention,

and, without any replications, it might be unwise to draw conclusions from that study alone. Besides, Remington and Pierce (1984) found a different pattern that did not clash so harshly with an obligatory spotlight.

More direct evidence comes from a study by LaBerge and Brown (1989) using a basic paradigm similar to one described earlier by LaBerge (1983). Each trial included two displays presented quickly in succession. The first had a target character in the center, with many distractors lined up horizontally on both sides, ensuring that attention was focused narrowly on the center of the screen when the second display appeared. The second display consisted of a single target character with a single distracting flanker character on either side. This three-character combination appeared somewhere along the horizontal line previously occupied by the first central target and its line of flankers. This task required attention to both the first display and the second display. RT increased with distance between the second target and center. This rise might be because the spotlight must be moved over to the second target, or it might be because the target appeared farther out on the attentional gradient and received less facilitation.

LaBerge and Brown (1989) tried to distinguish between the two explanations by varying the similarity between the second target and its two flanking distractors. They assumed that the nature of the distractors could not affect the speed of a moving obligatory spotlight. Thus, the distractor type should not affect the slope of the RT function when plotted by distance. However, the added difficulty that came from distractors very similar to the target might interact with the amount of facilitation from a stationary attentional gradient. The slope did indeed increase with the similarity between target and distractors, leading LaBerge and Brown to choose a gradient over a small moving spotlight. The moving spotlight might be salvaged by assuming that the spotlight is focused more narrowly and travels more slowly when the distractors are very similar to the target, but fitting that account with data from other experiments would probably be difficult.

Additional evidence for an attentional gradient comes from studies by C. W. Eriksen and his colleagues mentioned earlier. One example is C. W. Eriksen and St. James (1986), in which subjects identified a target letter (C or S) that appeared among distractor letters (A, H, or N). The target always appeared within a cued region, and one of the uncued locations also always contained a C or an S. When the distractor was different from the target, subjects took longer to identify the target. This interference increased if the distractor was closer to the cued region. The closer the distractor was to the attended region, the more difficult it was to ignore, just as would be expected with a gradient spotlight.

Castiello and Umiltà's (1990) follow-up study to C. W. Eriksen and St. James's (1986) study also found evidence for some type of attentional gradient. In their second experiment, Castiello and Umiltà (1990) cued the upcoming stimulus location with a box at one of two different sizes. The detection stimulus was presented at one of five locations—one at the center of the box, and four around the center—and these locations were the same across the two box size conditions. In the small box condition, the four eccentric stimulus positions fell outside the box, whereas, in the large box condition, they fell inside. Stimuli presented off the center of a box resulted in slower responses than stimuli presented at the center of the box only in the small box condition. These results could be interpreted as attention being directed precisely to the confines of the boxes. If the attended area had such a definite edge, however, responses to the eccentric locations should be faster with the large box. In fact, however, they were faster with the small box, when they were outside the cued area, suggesting that attention is uniformly distributed within the selected region and then falls off gradually from the edge of that region. Nonetheless, studies by Downing and Pinker (1985), Downing (1988), and Henderson and Macquistan (1993) all found graded attention effects from an exogenously cued location, which is more consistent with a gradient of attentional allocation.

Further evidence for attentional gradients can be found in electrophysiological studies (for a review, see Mangun & Hillyard, 1995). For example, Mangun and Hillyard (1987) examined ERP signals as subjects attended to one of three stimulus locations above fixation and reported the occasional appearance of target stimuli among a series of stimuli. They found progressive declines in the amplitude of evoked P1 and N1 components as the distance between evoking stimuli and the locus of attention increased. In a subsequent study, Mangun and Hillyard (1988) found a significant correlation between stimulus detectability and these ERP components. Interestingly, the later N2 component was elicited only by stimuli at the attended location and did not exhibit a graded amplitude decline, suggesting that attention may become more narrowly focused in late stages of visual processing.

An interesting aspect of attentional gradient measurements has been the consistently narrower estimates using signal detection measures, relative to RT measures. This discrepancy has led Handy et al. (1996) to ask whether these gradient estimate differences arise because of methodological confounds in signal detection measures or because RT and signal detection measures tap into different aspects of information processing. Using an experimental paradigm designed to allow accurate signal detection and RT measurements, they found that differences in gradient widths can be accounted for by differences in the perceptual requirements of signal detection and RT studies. Most signal detection experiments use near-threshold stimuli or discriminations, whereas RT experiments generally employ stimuli well above detectability threshold. Thus, it appears that as perceptual demands increase, attention is more narrowly distributed, possibly to increase resources at the target location to compensate for the increased difficulty in detecting the target or to decrease the possibility of false alarms from nonstimulus locations.

The data suggesting an attentional gradient might still be explained with a sharp-edged spotlight if we assume that subjects are imprecise in positioning the spotlight. If the exact location of the spotlight varies from trial to trial, the results averaged over many trials would mimic an attentional gradient, even without the spotlight moving during the trial. Barring such imprecision in location selection, the evidence suggests a mechanism that can facilitate processing of information over a wide area of the visual field. A stimulus need not fall into the center of the selected area to be fully identified and to trigger a response. The facilitation is strongest at the center of the chosen area and falls off with distance from the center. The steepness with which it declines can be adjusted depending on the nature of the task, the similarity between the distractors and the target, and the distance between the distractors and the target (see Downing, 1988). This facilitation can be applied to any region where it is needed, of course, but not quickly, so that there is little to be gained by deploying it if stimuli are simple and responses will be fast without the facilitation.

6. What Shape Is the Region That Is Selected, and How Much Flexibility Is There in Adjusting This Shape?

A spotlight always has a circular shape, but the region occupied by a visual object is often more complex. Multipart objects can have any configuration whatsoever, and if a target object is surrounded or partially obscured by distractors, it may be difficult to select a region that includes the target and none of the distractors unless it is possible to select regions of arbitrary shape. Some models suggest that there may be flexibility in the shape of the selected area (Logan, 1996), but this possibility has not been been explored much. In a study by Podgorny and Shepard (1983), subjects first viewed a 3×3 grid with four or five of the cells shaded. Next, a dot appeared, and the subject judged whether or not the dot appeared in the shaded area. Podgorny and Shepard calculated a "compactness rating" for each pattern of shaded cells and found that responses were faster when the shaded area was more compact. In other words, subjects responded more quickly when the shaded area was a simple shape, such as a square or rectangle. This advantage might come because it is harder to focus attention on an oddly shaped region. Alternatively, the RT differences may not reflect attentional processes. Rather than allocating their attention only to the shaded area, subjects may attend to the whole grid, because the dot can appear anywhere within it. The longer RTs for complex shapes may reflect the higher demands on spatial memory and extra precision required in localization of the dot.

C. W. Eriksen and St. James (1986) concluded that the spotlight is probably circular or oval-shaped. As described earlier, their subjects identified one target letter that appeared along with seven distractor letters in a circle. The target only appeared at a position that had been previously cued. In one of their experiments, they included one condition in which half of the positions (all contiguous) were cued and another condition in which all eight were cued. Because performance was only slightly better when half of the positions could be ruled out, they concluded that the shape of the spotlight must make it difficult to focus attention on only half of the circle. A later study by Pan and C. W. Eriksen (1993) indicated that subjects cannot attend to two noncontiguous target object locations without processing intervening distractor positions, adding more evidence that the spotlight does not have a complex shape.

In contrast to the C. W. Eriksen and St. James (1986) study, an experiment by Juola, Bouwhuis, Cooper, and Warner (1991) raises the possibility that attention is much more flexible than a circular or oval-shaped spotlight. In an experimental paradigm similar to that of Egly and Homa (1984), they presented subjects with three concentric rings around fixation and cued one of them before the presentation of a target letter at 1 of 24 candidate positions, 8 on each ring; all other positions were occupied by distractor letters. They found that subjects responded more quickly to the target when it appeared on the cued ring (i.e., valid trials) than when it did on an uncued ring (i.e., invalid trials). This effect was observed for all three rings, and the fact that it was observed for the largest and middle-size rings argues against a spotlight (or zoom lens) model of attentional allocation in which attention would be stronger, or at least just as strong, at foveal locations than at peripheral locations. Thus, Juola et al. argued that attention can be allocated in a specific size and shape (i.e., a ring in this experiment). The data can also be explained using a spotlight that scans locations around the cued ring serially, although Juola et al. showed that their model of attentional allocation accounted for more of the variance in RT than did a model of serially scanning spotlight.

In order to consider the size and shape of the attentional window, one must first assume that attention operates by facilitating processing in a selected region of the visual field. However, different selection methods are possible, especially when the task does not involve attending to a cued location. In some circumstances, attention might operate more by suppressing distractor information than it does by enhancing target information. Cepeda, Cave, Bichot, and Kim (1998) have investigated how spatial attention is allocated when subjects search for a target of a given color among distractors of another color. To measure the allocation of attention in visual space, they used the dual-task probing technique of Kim and Cave (1995) and measured RT to small, postdisplay black squares. To determine whether attention operates by facilitation of the target location (as implied in the spotlight metaphor) or by inhibition of distractor locations, the probe was presented at blank locations between stimuli in addition to the usual target or distractor locations. Comparing probe RTs at stimulus locations and blank locations is complicated by potential differences in masking, but Cepeda et al. devised stimuli to equalize masking across all locations. The entire screen was covered

by a grid of horizontal and vertical lines, and target and distractor stimuli were produced by changing the color of line segments in four regions of the grid to the target or distractor color. The subjects were instructed to determine the orientation of the line segments of the target color. On some trials, a probe appeared either at the location of one of the search stimuli or at a location between two of the search stimuli. In either case, the probe was surrounded by lines from the grid, and a control experiment confirmed that masking was equal at all locations. When the probe appeared, subjects pressed a button.

Probes at target locations elicited faster responses than did probes at distractor locations. More importantly, probes at distractor locations elicited significantly slower responses than did probes at blank locations. Thus, selection in this search task was accomplished via the inhibition of distractor location information. This evidence for attentional inhibition by Cepeda et al. (1998) is consistent with the previously mentioned study by Cave and Zimmerman (1997) showing attentional distribution during a letter search reminiscent of flanking inhibition. Of course, attention allocated in response to spatial cues may still take the form of facilitation of the cued location.

In summary, there are at least some circumstances in which attention takes the form of inhibition focused at distractor locations. Further experiments are necessary to find the conditions that determine whether attention appears as focused facilitation, a ring of facilitation, focused inhibition, or some other complex spatial pattern. If there are multiple selection mechanisms operating at different processing levels to control different types of interference and overload (Luck, Hillyard, Mouloua, & Hawkins, 1996), then we might expect the spatial pattern of attentional effects to vary considerably, depending on which mechanism is triggered by the demands of the current task. These questions about the shape of the selected area lead to one final question: whether noncontiguous regions can be selected simultaneously.

7. Can Two or More Separate Regions be Selected At Once Without Selecting the Area in Between?

Often, two different potential targets will be separated by a distractor object, or a target object will be partially obscured by a distractor. In these circumstances, it would be useful to select two noncontiguous regions while excluding information from the area between them. In fact, early studies by M. L. Shaw and P. Shaw (1977) and M. L. Shaw (1978) presented accuracy and RT data, respectively, suggesting the possibility that subjects simultaneously allocate different amounts of attention to different locations in the visual field in order to optimize performance. However, these data were also consistent with the selection of only a single location on each trial, with the probability of a location being selected on a given trial matching the probability of target occurrence at that location.

This possibility led Posner et al. (1980) to try a modification of the basic cuing paradigm in which they cued two different potential target locations. Each trial had a primary cue that was valid on 65% of the trials and a secondary cue that was valid on 25%. The secondary cue produced a benefit only when its location was adjacent to the location of the primary cue. The authors concluded that the facilitation from attention could not be split across two noncontiguous regions. C. W. Eriksen and Yeh (1985) tried a different cuing experiment in which subjects searched for one of two target letters among distractor letters, and they reached the same conclusion.

This conclusion is further supported by ERP evidence presented by Heinze et al. (1994). Their subjects compared two shapes that either were at two contiguous locations or were separated by another shape. Spatial attention was measured using the P1 component of event-related brain potentials to probe stimuli that appeared after the stimuli to be compared. They found attentional enhancement of the P1 component for probes appearing at locations occupied by either of the targets to be compared and also for probes at the location in between the two targets. They concluded that the intervening region was selected along with the two targets, although, as they point out, the P1 component is an indirect measure and may not reflect all aspects of spatial attention.

In contrast to these studies, Castiello and Umiltà (1992) used a different approach to demonstrate attention to noncontiguous locations in a cuing paradigm, on the basis of the assumption that the benefits of attention will be diminished when spread over a larger area. However, their experiments, like other cuing studies exploring this issue, leave open the possibility that subjects may attend to one location on some trials and another location on other trials, producing mean RTs that resemble a split of attention (see also Egly & Homa, 1984; Juola et al., 1991). Castiello and Umiltà (1992) argue against this strategy by presenting the RT distributions and comparing variances between different conditions, but their evidence does not completely exclude the possibility of attention to different locations on different trials.

Kramer and Hahn (1995) proposed an explanation for the discrepancy between the studies, indicating a single attentional spotlight and those supporting split attention. They suggested that some studies may have failed to find evidence for split attention because stimuli in these studies were presented as sudden onsets, which in some circumstances will "automatically" capture attention (Theeuwes, 1995; for a review, see Yantis, 1995). The attentional capture of sudden-onset distractor stimuli may make it difficult or impossible for subjects to maintain their attentional focus on previously cued locations. To test their hypothesis, Kramer and Hahn used an experimental paradigm modeled after that of Pan and C. W. Eriksen (1993), who used response competition to show that subjects are unable to selectively ignore stimuli that were located between two cued locations. Kramer and Hahn cued (with boxes) two target locations separated by two distractor locations. Subjects determined whether the letters presented inside the two cue boxes were the same

while ignoring intervening distractor letters. In one condition, the letters were presented as sudden onsets; in the other condition, they were revealed by the removal of segments of a figure-eight premask (i.e., nononset condition). As predicted, distractor letters interfered with performance in the sudden-onset condition (as in Pan & C. W. Eriksen, 1993) but not in the nononset condition, leading Kramer and Hahn to conclude that subjects could simultaneously attend to the two cued locations.

All of the above studies have used some form of location cuing to investigate split attention. A recent study by Bichot, Cave, and Pashler (1999) investigated, using two different methods, subjects' ability to simultaneously attend to multiple noncontiguous locations in visual search. In one set of experiments, subjects attended to red digits presented in multiple frames with green digits, similar to the paradigm used in Cave and Pashler (1995). Accuracy was no better when red digits appeared successively than when pairs of red digits occurred simultaneously, implying split attention to the two locations simultaneously. Different tasks involving oriented grating stimuli produced the same result. Another experiment demonstrated split attention with an array of spatial probes, similar to one of the techniques used in the studies by Kim and Cave (1995) and Tsal and Lavie (1993). When the probe at one of two target locations was correctly reported, the probe at the other target location was more often reported correctly than any of the probes at distractor locations, including those between the targets. Together, these experiments provide strong converging evidence that when two targets are easily discriminated from distractors by a basic property, spatial attention can be split across both locations. Perhaps these results are not surprising when one considers Cepeda et al.'s (1998) finding that, during a visual search, selection is primarily mediated by suppression of information from distractor locations. Furthermore, unlike in Kramer and Hahn (1995) where location cuing was used, subjects were able to split their attention even though all stimuli were presented as sudden onsets. Apparently, the inherent featural differences between Bichot et al.'s targets and distractors were robust enough to override the effects of sudden onset. Overall, evidence from several studies suggests that attention can be divided across noncontiguous locations, but only under the right circumstances.

Whatever Attention Is, What Is It Doing?

There are still many more questions than answers about the nature of visual attention and the validity of the spotlight metaphor. There does seem to be a location-based selection mechanism operating early in visual processing. It probably requires the same amount of time to shift between locations regardless of distance, and there is little evidence that it affects the areas between its starting point and its destination. The size of the region to be selected can probably be adjusted up or down. Selecting a larger area may result in some loss of resolution or efficiency in some tasks, but not in others. The selected area

is probably not sharply defined. Rather, there seems to be a gradient of activation, strongest in the center or the selected area and falling off toward the edges. There are few clues as to whether the shapes of selected areas can be adjusted; however, given their gradient nature, complex shapes are probably not possible. Under some circumstances, attention is focused inhibition, rather than focused facilitation. Furthermore, it is sometimes possible to select stimuli at two or more noncontiguous locations.

Further progress in understanding visual selection will require that we formulate our questions carefully. We must ask specific questions about what factors control selection and how they are manipulated. When does location determine what information is selected? Do other properties play a role in selection as well, and, if so, under what circumstances? How is attention shifted from one location to another? What is the relationship between selection by location and selection by size or scale? When is focused inhibition used rather than focused facilitation, and how is it directed?

Answers to these questions will provide important clues to the role that visual selection plays in higher level visual processing. The most valuable payoff from studying visual selection will be a more complete understanding of visual cognition in general. Below are four different conceptions of the spotlight's role in visual processing. These four explanations are not mutually exclusive. Future research should be aimed at assessing these alternatives (and possibly others not yet conceived).

Spotlight as attention window. One way to simplify identification of a visual object is to identify the region occupied by that visual object and to cleanly and completely separate information in this region from the rest of the input. The spatial layout of that information can be preserved as it is scaled to a standard size and shifted to a standard location. This raw visual input can then be passed on to higher level mechanisms to be analyzed and identified. Such a selection mechanism provides a computational advantage, because most complex visual processing is limited to the small selected region, making it unnecessary to duplicate hardware across all different locations in the visual field (Kosslyn & Koenig, 1992; Olshausen et al., 1993). This mechanism is also appealing because its role in visual processing is well defined, making it easy to understand and implement. However, the data indicate that the spotlight does not have the well-defined edge that would be expected from such a simple attention window. This explanation might be saved by assuming that the spotlight is positioned imprecisely on each trial, but such a poorly aimed spotlight seems of little use in object recognition. Another alternative is to assume that the output of the spotlight is fed into a processing buffer that analyzes its center more thoroughly and precisely than its edges. If these alternatives prove untenable, and there is not an attention window with a sharp edge separating selected from unselected information, it implies that visual recognition is fairly robust and does not require the complete elimination of interference

from distractors. Even though the facilitation from an attentional gradient may be weak and imprecise, in many cases it may be enough to cause activation of a target representation to inhibit activation of distractor representations in a winner-take-all competition.

Spotlight as distributed resource. C. W. Eriksen accounts for his results by assuming that the spotlight works by allocating processing resources over the selected areas of the visual field. The concept of attention as a resource to be allocated goes back to Kahneman (1973). If attention is allocated to a large area, then fewer resources can be dedicated to any single location, and the entire area will be processed somewhat less efficiently than if a smaller area had been selected. Thus, spreading attention over a large area entails a cost, even if no distractors appear within that area. This cost appears in the experiment by Egeth and Shwartz (see Egeth, 1977) described in Section 4, but not in LaBerge's (1983) study.

Spotlight as filter. Obviously, visual processing becomes more difficult when more objects are present in the visual field. A likely reason for the spotlight's existence, and perhaps the only reason, is to block information from distractor objects that would otherwise overload the later stages of visual processing. If this were the spotlight's only mission, then we might expect to see its effects only when distractors are present or at least expected by the subject. Yet some of the earliest demonstrations of the spotlight were in cuing experiments in which the target was a simple dot and almost no distracting stimuli were present. Of course, subjects might be in the habit of deploying the spotlight filter even in those situations in which it is not necessary. On the other hand, filtering might be useful in this task if scanning larger regions for dots increases the probability of false alarms. Also, if the spotlight is necessary for filtering and not resource allocation, it becomes difficult to explain the cost associated with large attended regions found by Egeth and Shwartz (see Egeth, 1977). Because they used a letter discrimination task rather than simple dot detection, the chance of false alarms seems much less. However, subjects do not necessarily use their attention mechanism optimally, especially in these artificial laboratory tasks. For instance, when the target can appear within a large area, some subjects may occasionally focus their attention out toward the edge of the cued area.

If the main purpose of visual attention is to filter out distractors, then the patches of distractor inhibition suggested by Cepeda et al. (1998) should be very effective. This strategy of selection will prevent distractors from interfering with target processing but will not hamper the detection of new stimuli that suddenly appear in the visual field. In most natural scenes, however, there are so many distractor objects that most of the visual field may be inhibited, producing a pattern that looks more like a spotlight of facilitation. If this type of selection occurred simultaneously in many visual maps representing the input at different scales, it would produce an attentional gradient (see Cave, in press, for details).

Spotlight for action. Regardless of whatever processing limitations there might be in vision, there are obvious limitations on the type or responses that can be produced at any one time. For instance, gaze can only be directed to a single location of the visual field. The bottleneck in selecting a response has led many to suggest that a principal role for attention is to reduce the influence of irrelevant information on motor actions (for a review, see van der Heijden, 1990). This view is supported by evidence from physiological studies (for a review, see Desimone & Duncan, 1995) and behavioral studies (Cave & Pashler, 1995; C. W. Eriksen & Hoffman, 1973). However, one should be careful not to confuse the concept of "selection for action" with late selection theories of visual attention. In other words, even though selection plays an important role in action, it does not have to be implemented exclusively in late stages of brain processing in areas that have strong motor characteristics. In fact, neural evidence suggests that attentional effects are present in nearly all visual or visual-related brain areas (for reviews, see Duncan, Humphreys, & Ward, 1997, and Maunsell, 1995). It appears that visual stimuli compete for representation by neurons at multiple levels of visual processing, possibly to influence motor output. Such multilevel competition has been implemented in a recent model of visual selection by Cave, Kim, Bichot, and Sobel (1999).

Multiple attention systems. By pursuing the questions posed here about the nature of attention, we are collecting evidence that will help us to distinguish from among these different accounts of attention. A final concern in this endeavor is the possibility of multiple attention systems. If visual processing involves a number of separate stages, then each stage will have its own computational limits and may have its own selection mechanism or mechanisms. On the practical side, this means that when comparing results from different experimental paradigms, one cannot assume that the same selection mechanism is at work in all of them. Thus, combining results across different experimental paradigms can be tricky. On the theoretical side, we must remember that once we assemble the full picture of selection in vision, we may find that each of the accounts listed here is part of the picture.

REFERENCES

Baylis, G. C., & Driver, J. (1992). Visual parsing and response competition: The effect of grouping factors. *Perception & Psychophysics*, **51**, 145-162.

Baylis, G. C., & Driver, J. (1993). Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 451-470.

Bichot, N. P., Cave, K. R., & Pashler, H. (1999). Visual selection mediated by location: Feature-based selection of noncontiguous locations. *Perception & Psychophysics*, **61**, 403-423.

Castiello, U., & Umiltà, C. (1990). Size of the attentional focus and efficiency of processing. *Acta Psychologica*, **73**, 195-209.

Castiello, U., & Umiltà, C. (1992). Splitting focal attention. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 837-848.

CAVE, K. R. (in press). The FeatureGate model of visual selection. *Psychological Research*.

CAVE, K. R., KIM, M.-S., BICHOT, N. P., & SOBEL, K. V. (1999). *Visual selection within a hierarchical network: The FeatureGate model*. Manuscript submitted for publication.

CAVE, K. R., & KOSSLYN, S. M. (1989). Varieties of size-specific visual selection. *Journal of Experimental Psychology: General*, **118**, 148-164.

CAVE, K. R., & PASHLER, H. (1995). Visual selection mediated by location: Selecting successive visual objects. *Perception & Psychophysics*, **57**, 421-432.

CAVE, K. R., & ZIMMERMAN, J. M. (1997). Flexibility in spatial attention before and after practice. *Psychological Science*, **8**, 399-403.

CEPEDA, N. J., CAVE, K. R., BICHOT, N. P., & KIM, M.-S. (1998). Spatial selection via feature-driven inhibition of distractor locations. *Perception & Psychophysics*, **60**, 727-746.

CHASTAIN, G. (1992a). Analog versus discrete shifts of attention across the visual field. *Psychological Research*, **54**, 175-181.

CHASTAIN, G. (1992b). Time-course of sensitivity changes as attention shifts to an unpredictable location. *Journal of General Psychology*, **119**, 105-111.

CHEN, Z. (1998). Switching attention within and between objects: The role of subjective organization. *Canadian Journal of Experimental Psychology*, **52**, 7-16.

CONNOR, C. E., PREDDIE, D. C., GALLANT, J. L., & VAN ESSEN, D. C. (1997). Spatial attention effects in macaque area V4. *Journal of Neuroscience*, **17**, 3201-3214.

DESIMONE, R., & DUNCAN, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, **18**, 193-222.

DOWNING, C. J. (1988). Expectancy and visual-spatial attention: Effects on perceptual quality. *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 188-202.

DOWNING, C. J., & PINKER, S. (1985). The spatial structure of visual attention. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI: Mechanisms of attention* (pp. 171-187). Hillsdale, NJ: Erlbaum.

DRIVER, J., & BAYLIS, G. C. (1989). Movement of visual attention: The spotlight metaphor breaks down. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 448-456.

DUNCAN, J. (1981). Directing attention in the visual field. *Perception & Psychophysics*, **30**, 90-93.

DUNCAN, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, **4**, 501-517.

DUNCAN, J., HUMPHREYS, G., & WARD, R. (1997). Competitive brain activity in visual attention. *Current Opinion in Neurobiology*, **7**, 255-261.

EGETH, H. (1977). Attention and preattention. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 11, pp. 277-320). New York: Academic Press.

EGLY, 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.

EGLY, R., & HOMA, D. (1984). Sensitization in the visual field. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 778-793.

ERIKSEN, B. A., & ERIKSEN, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, **16**, 143-149.

ERIKSEN, C. W., & HOFFMAN, J. E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. *Perception & Psychophysics*, **12**, 201-204.

ERIKSEN, C. W., & HOFFMAN, J. E. (1973). The extent of processing of noise elements during selective encoding from visual displays. *Perception & Psychophysics*, **14**, 155-160.

ERIKSEN, C. W., & HOFFMAN, J. E. (1974). Selective attention: Noise suppression or signal enhancement? *Bulletin of the Psychonomic Society*, **4**, 587-589.

ERIKSEN, C. W., & MURPHY, T. D. (1987). Movement of attentional focus across the visual field: A critical look at the evidence. *Perception & 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 & Psychophysics*, **40**, 225-240.

ERIKSEN, C. W., & WEBB, J. M. (1989). Shifting of attentional focus within and about a visual display. *Perception & Psychophysics*, **45**, 175-183.

ERIKSEN, C. W., & YEH, Y.-Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception & Performance*, **11**, 583-597.

GATTI, S. V., & EGETH, H. E. (1978). Failure of spatial selectivity in vision. *Bulletin of the Psychonomic Society*, **11**, 181-184.

GENTNER, D., & GRUDIN, J. (1985). The evolution of mental metaphors in psychology: A 90-year retrospective. *American Psychologist*, **40**, 181-192.

GENTNER, D., & JEZIORSKI, M. (1993). The shift from metaphor to analogy in Western science. In A. Ortony (Ed.), *Metaphor and thought* (pp. 447-480). Cambridge: Cambridge University Press.

GIBSON, B. S. (1994). Visual attention and objects: One versus two or convex versus concave? *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 203-207.

HANDY, T. C., KINGSTONE, A., & MANGUN, G. R. (1996). Spatial distribution of visual attention: Perceptual sensitivity and response latency. *Perception & Psychophysics*, **58**, 613-627.

HARMS, L., & BUNDESEN, C. (1983). Color segregation and selective attention in a nonsearch task. *Perception & Psychophysics*, **33**, 11-19.

HAWKINS, H. L., HILLYARD, S. A., LUCK, S. J., MOULOUA, M., DOWNING, C. J., & WOODWARD, D. P. (1990). Visual attention modulates signal detectability. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 802-811.

HEINZE, H.-J., LUCK, S. J., MÜNTE, T. F., GÖS, A., MANGUN, G. R., & HILLYARD, S. A. (1994). Attention to adjacent and separate positions in space: An electrophysiological analysis. *Perception & Psychophysics*, **56**, 42-52.

HENDERSON, J. M., & MACQUISTAN, A. D. (1993). The spatial distribution of attention following an exogenous cue. *Perception & Psychophysics*, **53**, 221-230.

HOFFMAN, J. E., & NELSON, B. (1981). Spatial selectivity in visual search. *Perception & Psychophysics*, **30**, 283-290.

HOFFMAN, J. E., NELSON, B., & HOUCK, M. R. (1983). The role of attentional resources in automatic detection. *Cognitive Psychology*, **51**, 379-410.

HUGHES, H. C., & ZIMBA, L. D. (1985). Spatial maps of directed visual attention. *Journal of Experimental Psychology: Human Perception & Performance*, **11**, 409-430.

HUGHES, H. C., & ZIMBA, L. D. (1987). Natural boundaries for the spatial spread of directed visual attention. *Neuropsychologia*, **25**, 5-18.

HUMPHREYS, G. W. (1981). Flexibility of attention between stimulus dimensions. *Perception & Psychophysics*, **30**, 291-302.

JUOLA, J. F., BOUWHUIS, E. E., COOPER, C., & WARNER, B. (1991). Control of attention around the fovea. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 125-141.

KAHNEMAN, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

KAHNEMAN, D., & HENIK, A. (1977). Effects of visual grouping on immediate recall and selective attention. In S. Dornic (Ed.), *Attention & performance VI* (pp. 307-332). Hillsdale, NJ: Erlbaum.

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., & GIBBS, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, **24**, 175-219.

KIM, M.-S., & CAVE, K. R. (1995). Spatial attention in visual search for features and feature conjunctions. *Psychological Science*, **6**, 376-380.

KIM, M.-S., & CAVE, K. R. (1999). *Perceptual grouping via spatial attention in a focused-attention task*. Manuscript submitted for publication.

KOSSLYN, S. M., & KOENIG, O. (1992). *Wet mind: The new cognitive neuroscience*. New York: The Free Press.

KRAMER, A. F., & HAHN, S. (1995). Splitting the beam: Distribution of attention over noncontiguous regions of the visual field. *Psychological Science*, **6**, 381-386.

KRAMER, A. F., WEBER, T. A., & WATSON, S. E. (1997). Object-based attentional selection—Grouped-arrays or spatially-invariant representations? Comments on Vecera and Farah (1994). *Journal of Experimental Psychology: General*, **126**, 3-13.

KUBOVY, M. (1981). Concurrent-pitch segregation and the theory of indispensable attributes. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 55-98). Hillsdale, NJ: Erlbaum.

KWAK, H.-W., DAGENBACH, D., & EGETH, H. (1991). Further evidence for a time-independent shift of the locus of attention. *Perception & Psychophysics*, **49**, 473-480.

LABERGE, D. (1983). Spatial extent of attention to letters and words. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 371-379.

LABERGE, D., & BROWN, V. (1989). Theory of attentional operations in shape identification. *Psychological Review*, **96**, 101-124.

LARSEN, A., & BUNDESEN, C. (1978). Size scaling in visual pattern recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **4**, 1-20.

LAVIE, N., & DRIVER, J. (1996). On the spatial extent of attention in object-based visual selection. *Perception & Psychophysics*, **58**, 1238-1251.

LEWONTIN, R. (1981). On constraints and adaptation. *Behavioral & Brain Sciences*, **4**, 245.

LOGAN, G. D. (1996). The CODE theory of visual attention: An integration of space-based and object-based attention. *Psychological Review*, **103**, 603-649.

LUCK, S. J., CHELAZZI, L., HILLYARD, S. A., & DESIMONE, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology*, **77**, 24-42.

LUCK, S. J., FAN, S., & HILLYARD, S. A. (1993). Attention-related modulation of sensory-evoked brain activity in a visual search task. *Journal of Cognitive Neuroscience*, **5**, 188-195.

LUCK, S. J., & HILLYARD, S. A. (1995). The role of attention in feature detection and conjunction discrimination: An electrophysiological analysis. *International Journal of Neuroscience*, **80**, 281-297.

LUCK, S. J., HILLYARD, S. A., MOULOUA, M., & HAWKINS, H. L. (1996). Mechanisms of visual-spatial attention: Resource allocation or uncertainty reduction? *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 725-737.

MALJKOVIC, V., & NAKAYAMA, K. (1996). Priming of pop-out: II. The role of position. *Perception & Psychophysics*, **58**, 977-991.

MANGUN, G. R., & HILLYARD, S. A. (1987). The spatial allocation of visual attention as indexed by event-related potentials. *Human Factors*, **29**, 195-212.

MANGUN, G. R., & HILLYARD, S. A. (1988). Spatial gradients of visual attention: Behavioral and electrophysiological evidence. *Electroencephalography & Clinical Neurophysiology*, **70**, 417-428.

MANGUN, G. R., & HILLYARD, S. A. (1991). Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 1057-1074.

MANGUN, G. R., & HILLYARD, S. A. (1995). Mechanisms and models of selective attention. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind: Event-related brain potentials and cognition* (pp. 41-85). Oxford: Oxford University Press.

MAUNSELL, J. H. R. (1995). The brain's visual world: Representation of visual targets in cerebral cortex. *Science*, **270**, 764-769.

MORAN, J., & DESIMONE, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, **229**, 782-784.

MOTTER, B. C. (1994). Neural correlates of feature selective memory and pop-out in extrastriate area V4. *Journal of Neuroscience*, **14**, 2190-2199.

MURPHY, T. D., & ERIKSEN, C. W. (1987). Temporal changes in the distribution of attention in the visual field in response to precues. *Perception & Psychophysics*, **42**, 576-586.

NAVON, D. (1978). On a conceptual hierarchy of time, space, and other dimensions. *Cognition*, **6**, 223-228.

OLSHAUSEN, B. A., ANDERSON, C. H., & VAN ESSEN, D. C. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *Journal of Neuroscience*, **13**, 4700-4719.

PAN, K., & ERIKSEN, C. W. (1993). Attentional distribution in the visual field during *same–different* judgments as assessed by response competition. *Perception & Psychophysics*, **53**, 134-144.

PODGORNY, R., & SHEPARD, R. N. (1983). Distribution of visual attention over space. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 380-393.

POSNER, M. I., NISSEN, M. J., & OGDEN, W. C. (1978). Attended and unattended processing modes: The role of set for spatial location. In H. J. Pick & I. J. Saltzman (Eds.), *Modes of perception* (pp. 137-157). 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.

PRINZMETAL, W. (1981). Principles of feature integration in visual perception. *Perception & Psychophysics*, **30**, 330-340.

PRINZMETAL, W., & MILLIS-WRIGHT, M. (1984). Cognitive and linguistic factors affect visual feature integration. *Cognitive Psychology*, **16**, 305-340.

REMINGTON, R., & PIERCE, L. (1984). Moving attention: Evidence for time-invariant shifts of visual selective attention. *Perception & Psychophysics*, **35**, 393-399.

RIZZOLATTI, G., RIGGIO, L., DASCOLA, I., & UMILTÀ, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia*, **25**, 31-40.

SAGI, D., & JULESZ, B. (1985). Fast noninertial shifts of attention. *Spatial Vision*, **1**, 141-149.

SCHALL, J. D., HANES, D. P., THOMPSON, K. G., & KING, D. J. (1995). Saccade target selection in frontal eye field of macaque: I. Visual and premovement activation. *Journal of Neuroscience*, **15**, 6905-6918.

SCHNEIDER, W., & SHIFFRIN, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, **85**, 1-66.

SHAW, M. L. (1978). A capacity allocation model for reaction time. *Journal of Experimental Psychology: Human Perception & Performance*, **4**, 586-598.

SHAW, M. L., & SHAW, P. (1977). Optimal allocation of cognitive resources to spatial location. *Journal of Experimental Psychology: Human Perception & Performance*, **3**, 201-211.

SHIH, S.-I., & SPERLING, G. (1996). Is there feature-based attentional selection in visual search? *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 758-779.

SHULMAN, G. L., REMINGTON, R. W., & McLEAN, J. P. (1979). Moving attention through physical space. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 522-526.

SNYDER, C. R. R. (1972). Selection, inspection, and naming in visual search. *Journal of Experimental Psychology*, **92**, 428-431.

SPERLING, G., & WEICHSELGARTNER, E. (1995). Episodic theory of the dynamics of spatial attention. *Psychological Review*, **102**, 503-532.

THEEUWES, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics*, **50**, 184-193.

THEEUWES, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, **51**, 599-606.

THEEUWES, J. (1995). Perceptual selectivity for color and form: On the nature of the interference effect. In A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), *Converging operations in the study of visual selective attention* (ch. 10, pp. 297-314). Washington, DC: APA.

TREISMAN, A. M. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 194-214.

TREISMAN, A. [M.] (1992). Perceiving and re-perceiving objects. *American Psychologist*, **47**, 862-875.

TREISMAN, A. [M.], KAHNEMAN, D., & BURKELL, J. (1983). Perceptual objects and the cost of filtering. *Perception & Psychophysics*, **33**, 527-532.

TSAL, Y. (1983). Movements of attention across the visual field. *Jour-*

nal of Experimental Psychology: Human Perception & Performance, **9**, 523-530.

Tsal, Y., & Lavie, N. (1988). Attending to color and shape: The special role of location in selective visual processing. *Perception & Psychophysics*, **44**, 15-21.

Tsal, Y., & Lavie, N. (1993). Location dominance in attending to color and shape. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 131-139.

van der Heijden, A. H. C. (1990). Visual information processing and selection. In O. Neumann & W. Prinz (Eds.), *Relationship between perception and action* (pp. 203-226). Berlin: Springer-Verlag.

van der Heijden, A. H. C., Kurvink, A. G., de Lange, L., de Leeuw, F., & van der Geest, J. N. (1996). Attending to color with proper fixation. *Perception & Psychophysics*, **58**, 1224-1237.

Vecera, S. P. (1997). Grouped arrays versus object-based representations: Reply to Kramer et al. (1997). *Journal of Experimental Psychology: General*, **126**, 13-18.

Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General*, **123**, 146-160.

Weber, T. A., Kramer, A. F., & Miller, G. A. (1997). Selective processing of superimposed objects: An electrophysiological analysis of object-based attentional selection. *Biological Psychology*, **45**, 159-182.

Yantis, S. (1988). On analog movements of visual attention. *Perception & Psychophysics*, **43**, 203-206.

Yantis, S. (1995). Attentional capture in vision. In A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), *Converging operations in the study of visual selective attention* (ch. 2, pp. 45-76). Washington, DC: APA.

Zajonc, A. (1993). *Catching the light: The entwined history of light and mind*. New York: Bantam.

Zimba, L. D., & Hughes, H. C. (1987). Distractor–target interactions during directed visual attention. *Spatial Vision*, **2**, 117-149.

NOTES

1. Hoffman and Nelson (1981) used an inconsistent mapping of stimuli to responses, which Schneider and Shiffrin (1977) claimed would require controlled processing. Hoffman et al. (1983) used a consistent mapping, which, according to Schneider and Shiffrin, should result in automatic processing over time.

2. The object selection in Duncan's (1984) and Lavie and Driver's (1996) experiments could also be explained by a mechanism capable of selecting different locations in three-dimensional space, if subjects perceive the two superimposed objects as being in different depth planes. Lavie and Driver offer a detailed argument rejecting this explanation.

3. Both Larsen and Bundesen (1978) and Cave and Kosslyn (1989) concluded that there was also a second, slower size scaling process that was related to visual imagery and that came into play only when subjects knew in advance what shape was likely to appear.

4. One might object that when subjects must respond to a probe that can occur anywhere within a large area, as in Downing and Pinker's (1985) experiment, they must spread their spotlight out over a large area in order to be able to detect it. There is an alternative, however. There may be a simple preattentive mechanism that can detect any stimulus, such as Downing and Pinker's probe, that differs from its surroundings. This preattentive mechanism would not actually identify what the stimulus is but instead would direct the spotlight to that location.