Brain and Mind **1:** 353–384, 2000.

Toward a Biased Competition Account of Object-Based Segregation and Attention

SHAUN P. VECERA
Department of Psychology, E11 Seashore Hall, University of Iowa, Iowa City, IA 52242-1407, U.S.A. (E-mail: shaun-vecera@uiowa.edu)

(Received: 7 May 2000; in final form: 20 September 2000)

Abstract. Because the visual system cannot process all of the objects, colors, and features present in a visual scene, visual attention allows some visual stimuli to be selected and processed over others. Most research on visual attention has focused on spatial or location-based attention, in which the locations occupied by stimuli are selected for further processing. Recent research, however, has demonstrated the importance of objects in organizing (or segregating) visual scenes and guiding attentional selection. Because of the long history of spatial attention research, theories of spatial attention are more mature than theories of other visual processes, such as object segregation and object attention. In the present paper, I outline a biased competition account of object segregation and attention, following similar accounts that have been developed for spatial attention (Desimone and Duncan, 1995). In my biased competition account, I seek to understand how some objects can be segregated and selected over other objects in a complex visual scene. Under this account, there are two sources of visual information that allow an object to be processed over other objects: bottom-up information carried by the physical stimulus and top-down information based on an observer's goals. I use the biased competition account to combine many diverse findings from the object segregation and attention literatures into a common framework.

Key words: attention, perceptual organization, figure-ground segregation, object-based attention

The visual system receives thousands of environmental inputs at any given moment. Some of these inputs are relevant to current behavior and others are irrelevant. For example, the visual words printed on this page are relevant to the task of reading, but the perception of the desk on which this page lies is irrelevant to the task of reading. Because the visual system does not have the capacity to process all inputs simultaneously, there must exist attentional processes that help the visual system select some inputs (the words on the page) and filter out others (the desk).

Research on visual selective attention has suggested two theoretical accounts for selection (see Vecera and Luck, 2000, for an overview). Space-based or location-based accounts propose that stimuli are selected on the basis of spatial location by a spatial spotlight, a zoom lens, or a spatial gradient. For example, in a spotlight model, attention is directed to contiguous regions of the visual field; stimuli falling within the spotlight are processed further, and stimuli falling outside this region are processed little, if at all (Eriksen and Hoffman, 1973; Posner, 1980). A zoom lens account extends the spotlight account by proposing different levels of resolution, or focus, of the spotlight of attention (Eriksen and St. James, 1986). The spatial gradient approach incorporates aspects of both the spotlight and zoom lens views by hypothesizing that spatial attention is a fixed resource that is spread across the visual field (e.g., LaBerge and Brown, 1989). Evidence for spatial selection has come from a number of tasks, including spatial precuing tasks, in which observers are cued to a region in space by an abrupt luminance change, such as a flicker in the peripheral visual field. The standard validity effect in this line of research is that targets appearing the cued location are processed more efficiently than targets appearing at uncued locations (see Eriksen and Hoffman, 1973; Posner, 1980; Posner *et al.*, 1980).

In contrast with space-based accounts, object-based accounts of selective attention propose that stimuli are not selected on the basis of location alone. Instead, stimuli are selected as whole, organized “chunks,” and these chunks correspond to objects or shapes present in the visual environment. All of the visual features of an attended object are processed concurrently; features that belong to other, unattended objects are processed little, if at all (e.g., Duncan, 1984; Vecera and Farah, 1994). Evidence for object-based attention has come from several different tasks, with each task investigating the different parameters that allow the visual system to form objects from the sensory input. For example, some studies have demonstrated that stimuli that are grouped together by similarity (i.e., stimuli that are the same color) are selected simultaneously (Baylis and Driver, 1992). Other gestalt grouping cues, such as connectedness and good continuation, allow stimuli to be grouped and processed concurrently (e.g., Baylis and Driver, 1992; Kramer and Jacobson, 1991).

Because space-based selective attention has been studied longer and more intensively than object-based selection, theoretical accounts of and computational models of spatial attention are more mature than those of object attention (for recent theories and models of spatial attention, see Desimone and Duncan, 1995; Mozer and Sitton, 1998; Sandon, 1990; Sperling and Weichselgartner, 1995; Treisman, 1988; Wolfe, 1994). Because there are few theoretical accounts of object segregation and attention, in the present article, I outline a theoretical approach to object-based segregation and attention. My approach is to extend the biased competition account of visual search (Desimone and Duncan, 1995) to the initial segregation of objects from backgrounds and selection of these objects by attentional processes (also see Behrmann and Haimson, 1999; O’Craven *et al.*, 1999). Specifically, I will use biased competition to integrate diverse findings from the object segregation and attention literatures. My biased competition account should be viewed as a heuristic framework for (1) understanding the influences on object-based attention and segregation and (2) integrating the diverse literatures of object attention and segregation. One shortcoming of such an account is that it has a large amount of theoretical “wiggle room.” However, this theoretical wiggle room

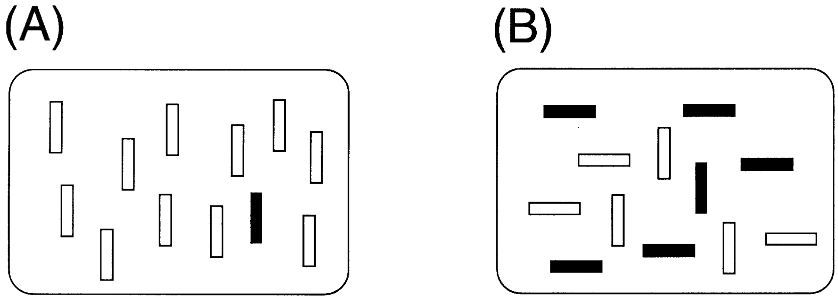

Figure 1. Sample visual search task in which observers search for a black vertical line. (A) Efficient feature search in which the target pops out from a homogeneous background. (B) Inefficient conjunction search in which the target does not pop out because the distractors share both black and vertical features with the target.

hopefully is offset by the account providing a conceptual structure that can be used to interpret a variety of experimental results.

In what follows, I outline my biased competition account of object segregation and selection and then review the empirical literature in light of this approach. I then discuss the biased competition account in conjunction with some of my previous research with a computational model of object segregation (Vecera and O'Reilly, 1998, 2000) that embodies the principles of the biased competition. The goal of my biased competition account is to try to explain object-based segregation and attention with a distributed processing system that relies on information in the external visual environment and on information relevant to an observer's current behavioral goals, motivation, and perceptual 'set.'

1. Biased Competition and Visual Search

The biased competition account of visual search was developed by Desimone and Duncan (1995; also see Cohen *et al.*, 1990; Cohen and Huston, 1994; Harter and Aine, 1984, for similar accounts) to explain the myriad results surrounding the visual search paradigm. Visual search refers to the collection of visual processes that allow us to "find what we are looking for" (e.g., Wolfe, 1998) by using spatial attention to combine the features of objects (e.g., Treisman, 1988; Treisman and Gelade, 1980). In a typical visual search task, several visual stimuli are present, and an observer searches for a specific target among the distracting stimuli (e.g., search for a black vertical bar in a field of distractors). Searching for a friend's face (the target) in a restaurant containing many people (the distractors) would be an example of "real world" visual search. There are two primary results from the visual search paradigm. First, targets that are highly salient (i.e., quite different from the background distractors) "pop out" and grab attention immediately and effectively, as when searching for a black vertical bar among white vertical bars (Figure 1A). These "feature searches" demonstrate efficient visual search in which the number of irrelevant distractors does not influence our time to search. Second, targets that are formed by a conjunction of features that are shared with the distractors require effortful search, as when searching for a black vertical bar among black horizontal bars, white vertical bars, and white horizontal bars (Figure 1B). Inefficient conjunction searches are highly dependent upon the number of distractors present; as the number of distractors increases, response times also increase (see Wolfe, 1998, for a comprehensive review of visual search).

How could these two results, which appear to be quite different, be explained by visual processes? Desimone and Duncan's (1995) biased competition model provides an answer. The biased competition view of visual search proposes two general sources for the control of attention: bottom-up sources that arise from sensory stimuli present in a scene and top-down sources that arise from the current behavioral goals. Visual search performance will require both bottom-up and top-down control sources to be considered and balanced against one another.[1]

In the biased competition account, the scene presented in a visual search task provides the bottom-up information that is searched through; this information indicates where objects are located and which features are present at each location. In efficient visual search (Figure 1A), the target may "pop out" from the distractors because of a possible bottom-up bias to orient attention to local inhomogeneities (e.g., Sagi and Julesz, 1984). The abrupt appearance of a new object or shape also captures attention (Yantis and Jonides, 1984; see Yantis, 1998, for a review), suggesting that abrupt onsets bias bottom-up attentional orienting. Not only do physically novel stimuli capture bottom-up attention, as when a unique target pops out of a display, but conceptually novel items capture attention in a bottom-up manner: A novel, upside down letter pops out of an array of familiar upright letters (e.g., Reicher *et al.*, 1976).

Under Desimone and Duncan's (1995) biased competition account of visual search, inefficient visual search may arise when there is no unique bottom-up stimulus characteristic to influence attentional allocation. Such a visual search (e.g., a conjunction search; Figure 1B) may depend on the top-down control of spatial attention in which items in a scene are examined in a sequential (or sequential looking) manner. A major source of top-down control in Desimone and Duncan's (1995) account is the target's description or identity. To be flexible, visual search must be sensitive to the goals of an observer. That is, an observer must be able to search for targets that may not be biased by the bottom-up input to the visual system. The target's description is based on the visual features of the target (e.g.,

[1] As one reviewer noted, the terms "bottom-up" and "top-down" are potentially confusing and unclear. I use these terms only for expository convenience because they are familiar to most who work in the field of cognitive neuroscience. Terms that may be more precise regarding the processing mechanisms implied are "stimulus driven" (bottom-up) and "goal driven" (top down). Also, discussing these components separately may give the impression that they operate at different times (i.e., information flows in a bottom-up direction first followed by top-down feedback). In interactive, distributed models, however, information flows in both directions simultaneously (see Vecera and O'Reilly, 2000, for discussion).

"black and vertical" in Figure 1), and in visual search experiments, the target description typically is based on an experimenter's instructions ("search for a black, vertical bar"). The experimenter's description of the target description can be conceptualized as creating a "template" that is stored in visual working memory for the duration of the task (e.g., Duncan and Humphreys, 1989; but see Woodman *et al.*, submitted). The target template acts to weight the incoming bottom-up stimulus information to allow attention to be biased toward one bottom-up input over another. In a less efficient conjunction search (Figure 1B), no single piece of bottom-up information is unique to the target item, so the bottom-up biases are less effective in guiding attention to the target. Top-down constraints are required to resolve the competition among the bottom-up inputs and bias attention toward the black, vertical line.

This brief overview of Desimone and Duncan's (1995) account highlights the two primary sources of attentional control – bottom-up and top-down sources. The biased competition model has been useful in describing a range of behavioral and neurobiological data from visual search experiments that rely on spatial attention, which suggests that the general approach of combining stimulus information and goal-related information may provide an accurate description of many attentional phenomena. How well would a biased competition approach explain the results from the emerging literature on object-based visual attention? I address this question in the subsequent sections by outlining a biased competition framework for object-based segregation and attentional control.

2. Biased Competition in Object-Based Segregation and Selection

Before reviewing the relevant findings from object-based segregation and attention research, I must be clear about my terms. *Object segregation* refers to the visual processes responsible for determining which visual features combine to form a single shape and which features combine to form other shapes. Object segregation is synonymous with perceptual organization, the term used in conjunction with the gestalt principles of visual organization (e.g., Wertheimer, 1923/1958). The ability to perform figure-ground segregation and distinguish foreground shapes ('figures') from background regions also involves object segregation processes (e.g., Rubin, 1915/1958). An example of object segregation appears in Figure 2A, which contains two perceptual groups that are formed by the gestalt principles of proximity and good continuation. The features (line segments) are organized into two distinct shapes – two lines, one straight and one squiggly. An example of figure-ground segregation appears in Figure 2B, in which the black symmetric region appears to be the foreground shape (the 'figure').

Object-based attention refers to the visual processes that select a segregated shape from among several segregated shapes.[2] Object segregation and object-based

[2] It is convenient to think of object-based attention as selecting one object among many objects. However, object-based selection also may arise from dividing attention among many objects simul-

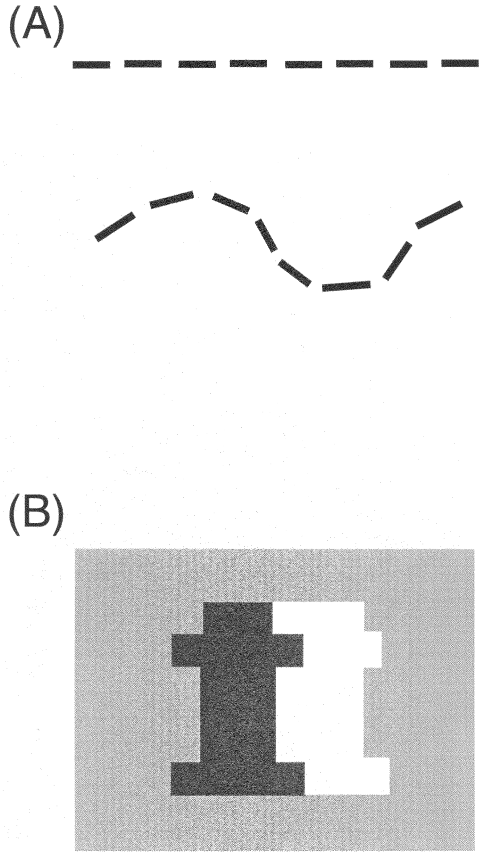

Figure 2. (A) An example of object segregation in which gestalt proximity and good continuation form two perceptual groups (two lines). The small straight lines of line 1 group together because they are closer to one another than the small lines of line 2. (B) An example of figure-ground segregation with a symmetry cue. The black symmetric region appears to lie in front of the asymmetric white region; the symmetric region appears to be more salient than the asymmetric region.

attention likely are interrelated – before a shape can be selected, the features of the shape first must be segregated from features of other shapes to some extent. In Figure 2A, before an observer could attend to the squiggly line, the features of that line must be grouped together and grouped separate from the features of the straight line. Object-based attention allows an observer to select either the straight line or the squiggly line with relatively little effort. However, if an observer wanted or needed to attend to both lines, object-based attention would be required. Object-based attention either would have to shift between the two lines or would need to be divided between the two lines. Either shifting or division of attention cause

taneously. For example, object-based costs in selection could be caused by dividing attention between two objects compared to focusing attention on one object.

performance to decline, and this is the basis of the object-based attentional effect in which attending to a single object is more efficient than attending to multiple objects. Similarly, in Figure 2B, object-based attention could be directed to either the black symmetric region or the white asymmetric region. Attentional selection is more efficient if attention is directed to a single region than to multiple regions.

Having defined the processes of object segregation and attention, I now turn to the key ideas behind the biased competition account I will rely on in the remainder of the paper. Because any visual scene contains many objects that compete with one another as they are being segregated and compete for attention, the visual system must be capable of allocating processing to one object or region over others. This allocation is achieved by biasing processing toward one object or region. These biases provide a resolution for the competition between objects or regions; this competition between objects occurs within both segregation processes and object attention processes. For example, the two regions in Figure 2B compete with one another for figural status (i.e., which region will be labeled 'figure'). Observers tend to perceive the symmetric black region as figure because symmetry acts as a bottom-up bias in the figure-ground competition and favors symmetric regions as figure (and asymmetric regions as ground). The biased competition account attempts to explain (1) how some objects or regions become more salient 'figures' or perceptual groups and (2) how some objects are selected over others.

In my biased competition account, as in Desimone and Duncan's (1995) account of visual search, there are two main sources for the biases that resolve the competition between objects. The first type of bias is a bottom-up bias; this bias arises from image cues or salient information in the external environment. The second type of bias is a top-down bias; this bias arises from task- or goal-relevant information. Both types of biases operate simultaneously and may cooperate or compete with one another.

Having outlined how competition between regions is resolved by biases, in the following sections I review empirical results that demonstrate that bottom-up and top-down information can control the bias of competition between objects or regions. As I will show in the following sections, the biased competition account that I outlined above provides a useful theoretical framework for understanding and integrating the empirical data on object segregation and attention (for previous reviews of these empirical data, see Driver and Baylis, 1998; Egeth and Yantis, 1997; Kanwisher and Driver, 1992).

2.1. BOTTOM-UP BIASES IN OBJECT SEGREGATION AND ATTENTION

In multi-object scenes, objects or regions compete with one another in two respects. First, there is a competition within object-based segregation processes and the segregated regions formed by segregation processes; the outcome of this competition is a perceptual group or figure that is more salient than other groups or figures. Second, there is a competition within object-based attentional processes;

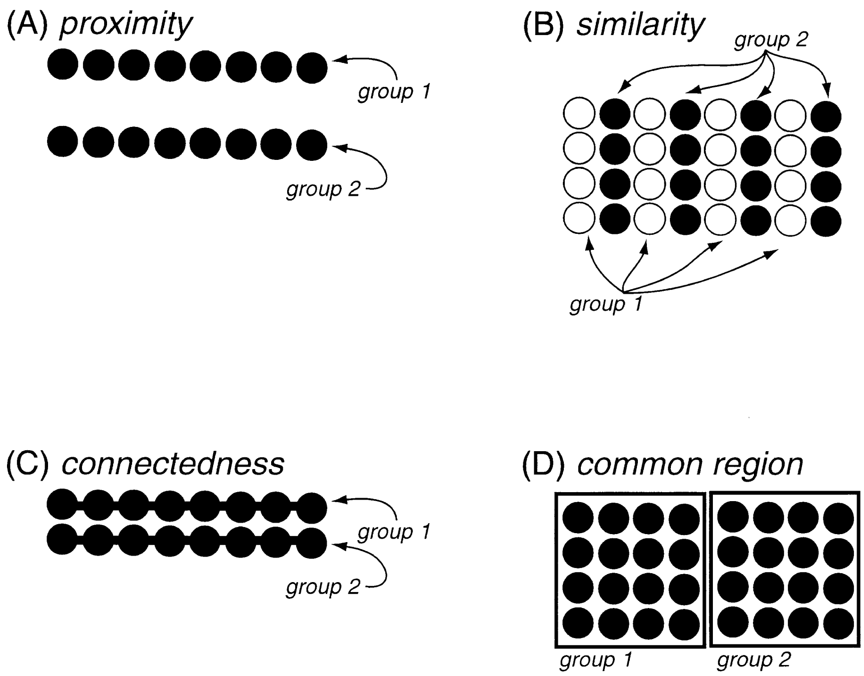

Figure 3. Some of the gestalt cues of perceptual organization. These cues are bottom-up heuristics that the visual system can use in determining which features should belong together and which features should be segregated apart from one another. (A) The proximity cue allows close features (the circles) to group to form horizontal rows of elements. (B) The similarity cue groups features on the basis of luminance such that features of similar luminance are grouped together. (C) The connectedness cue groups features that are physically connected to one another. (D) The common region cue groups features that are within a shared region.

the outcome of this competition is the selection of one perceptual group or figure over another. Although these two sources of competition are highly interrelated, they tend to be discussed as separate in the visual perception literature. The sources of bottom-up information, that is, the bottom-up biases, are discussed separately for segregation and attention.

Object segregation. There are at least two sources of bottom-up information influencing object segregation processes. The first sources of bottom-up information for object segregation are image cues, the cues that allow the visual system to determine if two features belong to the same object or different objects. The well-known gestalt principles of organization, shown in Figure 3, are bottom-up cues that are useful for visual segregation. Non-accidental properties (Lowe, 1985, 1987), properties which are unlikely to arise due to an observer's accidental viewpoint, also provide bottom-up cues that assist segregation. Also, visual scenes contain organization at many different levels. Specifically, objects can be segregated into smaller units that correspond to parts. There are bottom-up image cues, such as minima of curvature (i.e., places in which a contour points into an object) that allow objects to be segregated into parts (Hoffman and Richards, 1984; Hoffman and Singh, 1997). It appears that the visual system has bottom-up image cues that operate on entire objects (e.g., the gestalt cues) as well as operating below the object level. There are image cues that assist in decomposing an object into its parts (e.g., minima of curvature cues). The presence of these various cues for segregation demonstrates that information in the environment favors some segregation solutions over others. That is, although many segregations of a scene may be possible, the visual system uses bottom-up information to bias the segregation to conform to the principles outlined by the gestalt psychologists.

Although the gestalt principles can be viewed as bottom-up cues for object segregation, this does not imply that these cues may not be learned or acquired through experience. Many recent studies have suggested that bottom-up object segregation cues may arise from statistical regularities in the environment. The visual system becomes sensitive to these regularities through experience with visual scenes (see Mozer *et al.*, 1992, for a similar view). For example, in a visual scene, features that are the same color are also likely to move together and to be part of the same shape. Therefore, similarity of color is an environmental regularity that relates to individual objects – features that have the same color have a high probability of belonging to the same shape, so such features should be grouped with one another. Features that have different colors are less likely to belong to the same shape, so the tendency should be to segregate such features apart from one another (i.e., segregate them as belonging to different objects). This view of bottom-up object segregation cues emphasizes that bottom-up cues need not be innate to a viewer – learning can modify perceptual grouping if there are regularities in the environment (e.g., Behrmann *et al.*, 1998; Mozer *et al.*, 1992; Schyns and Murphy, 1994; Vecera and Farah, 1997), and it emphasizes that segregation is probabilistic, not an "all-or-none," process.

A second source of bottom-up information for object segregation is the information that allows one perceptual group or region to be favored over another. Not all perceptual groups are created equally, as shown in Figure 2. In Figure 2A, two perceptual groups have been created by gestalt grouping principles. However, neither object – the straight line or the squiggly line – appears more salient than the other. Contrast Figure 2A, which does not contain a bias for a perceptual group, with Figure 2B, which contains a strong bias favoring the symmetric region as the foreground figure. Thus, some bottom-up information causes one group or region to be more salient than another. Perhaps the best-known cues that work to bias processing based on bottom-up information are the gestalt principles of figure-ground segregation (see Pomerantz and Kubovy, 1986; Rock, 1975; Rubin, 1915/1958): Symmetric regions (Figure 2B) are more likely to be perceived as figure than asymmetric regions; smaller regions are more likely to be perceived as figure than larger regions; and convex (outwardly bulging) regions are more likely to be perceived as figure than concave regions. As in a visual search task in which a target pops out from distractors (Figure 1A), symmetry, area, and convexity may allow one region to pop out of a scene to be perceived as the figure; regions that are less likely to pop out may form background regions. The figure-ground principles are also likely to arise from environmental regularities and may be modified with experience.

Object attention. The attentional selection of objects is also constrained or biased by bottom-up information from the scene. Indeed, most of the object-based attention literature can be characterized as a search for the image grouping cues that influence attentional allocation in a bottom-up manner. Several studies, for example, have used the flanker task to determine the stimulus cues that guide attentional selection of objects. The flanker task was one of the earliest paradigms developed to study spatial selective attention (e.g., Eriksen and Eriksen, 1974; Eriksen and Hoffman, 1973). Observers are instructed to report a target letter appearing at a location (e.g., at the fixation point); the target letter is flanked by non-target letters which are either compatible with or incompatible with the response to the target. For example, observers are instructed to determine if a target letter is an H or a T; the target letter is either flanked by compatible letters (e.g., H flanked by other Hs) or by incompatible letters (e.g., H flanked by Ts). The flankers influence target processing: Flankers compatible with the target letter speed target identification and flankers incompatible with the target slow target identification.

The spatial selection involved in the flanker task can be biased by several object-grouping effects. Driver and Baylis (1989) reported that targets and flankers that moved together were processed as a single perceptual group (but see Kramer *et al.*, 1991, for a failure to replicate), suggesting that attention is biased to simultaneously process items moving in a common direction. These observations were extended to other bottom-up cues by Baylis and Driver (1992), who demonstrated that targets and flankers that were the same color were processed as a single group and that targets and flankers that were grouped through good continuation were processed simultaneously. Also, Kramer and Jacobson (1991) demonstrated that connectedness biased attentional selection in a flanker task. A target that was physically connected to the flankers was attended as a single unit or group; a target that was not connected to the flankers could be selectively attended with little influence from the surrounding flankers. As with other gestalt cues, connectedness cues bias attention to select items that are physically connected to one another.

A bottom-up bias over object selection also has been observed in spatial cuing tasks. Observers are given advance information about a target's location by a precue, and responses to targets are typically more accurate and faster when the cue and target appear at the same location (validly cued target) than at different locations (invalidly cued targets; see Posner, 1980; Posner *et al.*, 1980). Egly and colleagues used a spatial cuing task to study object-based attention (Egly *et al.*, 1994; also see Vecera, 1994). In this task, depicted in Figure 4, observers viewed two rectangles. The rectangles were perceptual groups formed by closure and

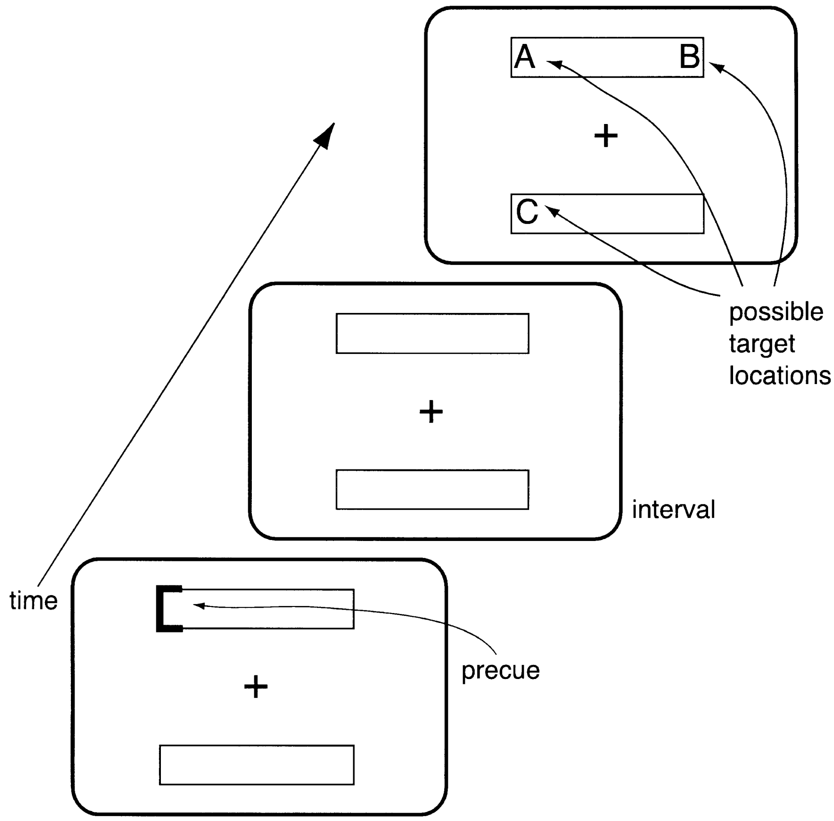

Figure 4. The spatial precuing task developed by Egly *et al.* (1994) to study object-based attention. Two rectangles appear, and the end of one is precued with a peripheral flash. After a delay, a target appears at one of three locations: (A) a validly cued target; (B) an invalidly cued target that appears in the cued object; (C) an invalidly cued object that appears in the uncued object. Observers are faster to detect invalidly cued targets appearing in the cued rectangle faster than those appearing in the uncued rectangle.

common region. One end of one of these shapes was cued, and a target appeared after the cue. Observers were instructed to detect the onset of the targets, and the targets could appear in one of three locations: at the spatially cued location (Figure 4A), in the opposite end of the cued object (Figure 4B), or in the uncued object (Figure 4C). Note that the targets that appear in the cued object are the same spatial distance from the cued region as the targets that appear in the uncued object. Using this task, Egly et al. reported that observers showed a spatial cuing effect; observers were faster to detect targets at the spatially cued location than at either of the uncued locations. More important, observers exhibited an object effect by detecting targets appearing in the cued object faster than targets appearing in the uncued object. These results indicate that closure and connectedness bias the allocation of spatial attention. When spatial attention is summoned to a cued location, attention can spread or move within a closed region more easily than between closed regions. The possibility that this task also may involve top-down information under a biased competition account is discussed later.

Object selection also is influenced by the common region grouping cue proposed by Palmer and Rock (1994). A recent study by Watson and Kramer (1999) demonstrated that two features of a single object were discriminated more easily than two features of different objects provided the object was composed of a single, homogenous color/texture. That is, object-based effects appear to depend upon an object being defined by a region that has common surface properties. If an object's surface properties are made heterogeneous by adding a different texture field between the two task-relevant features (such as a handle appears on a two-headed wrench; see Watson and Kramer, 1999), the single-object benefit is abolished; observers can discriminate features on different objects as easily as features on the same object. The color/texture inhomogeneity breaks common region, preventing this bottom-up cue from effectively guiding object selection.

Finally, bottom-up biases in attention exist for the segregation of an object into parts. Bottom-up image cues that allow an object to be decomposed into its parts, such as minima of curvature cues (Hoffman and Richards, 1984; Hoffman and Singh, 1997), influence attentional selection. Recent research from my lab demonstrates that observers are more accurate reporting features from a single part of an object than from multiple parts of an object (Vecera *et al.*, 2000, in press). These part-based attentional costs do not appear to be caused by selection with a simple spatially-based attention mechanism (Vecera *et al.*, 2000, in press) because changing the physical separation between the parts of an object influences attention very little, if at all. In some experimental paradigms, attention appears to select the parts themselves, not the visual locations occupied by the parts.

Information contained in a visual scene – bottom-up information – appears to both define perceptual groups and bias some groups to be more easily perceived than others. Once these segregation processes have operated, perceptual groups then bias the allocation of visual attention. In general, features or stimuli that group together based on the gestalt principles (Figure 3) bias attentional selection. Attention must obey those perceptual units formed by gestalt grouping processes, allowing attention to shift more easily within a group than between groups, for example.

A closer consideration of the object attention literature shows that these studies only have investigated how attention is biased by perceptual groups that are approximately equal in strength. That is, typically, two perceptual groups are presented to observers (e.g., Figure 4), and neither group is more salient than the other. A potentially interesting avenue for future research suggested by the biased competition account would involve the bottom-up "capture" of attention by salient perceptual groups. For example, symmetric regions are highly salient and easily segregate from background regions (Figure 2B). Another possible effect of a biased region, such as a symmetric region, is that it may capture attention in a bottom-up manner, much as peripheral flickers capture spatial attention automatically (e.g., Yantis, 1998). A bottom-up capture of attention by strong perceptual groups may explain data that show more accurate processing of high spatial frequency targets on figures than on grounds (Wong and Weisstein, 1983) and better memory for a figure's contours than for a ground's contours (Driver and Baylis, 1996).

Object segregation and attention are not guided entirely by bottom-up biases, however. Behavior would be severely limited if humans only recognized, attended, and acted upon objects defined by bottom-up criteria. The segregation and attention systems must be modulated by information that is relevant to current behavior or goals. This modulation comes in the form of top-down inputs that can bias the competition among perceptual groups. In the next section, I review some of the sources of top-down information that may influence or guide object segregation and attention.

2.2. TOP-DOWN BIASES

In many cases, visual scenes do not contain a single perceptually salient region that is relevant to some current behavior, such as searching for a coffee cup on a cluttered desk. Instead, there may be multiple regions or objects that have equal salience or, as a worst-case, there may be an object or event that is more salient than the object relevant to a goal (e.g., the coffee cup), such as when an error message abruptly appears on your computer monitor. Top-down sources of information will be needed either to bias attention to one of many equally salient objects or groups or to overcome a salient but behaviorally-irrelevant object or group. What types of top-down information can influence object segregation and attention, and what evidence is there for this top-down influence? There are at least three sources of top-down information in both object segregation and object selection: (1) object recognition processes, (2) perceptual "set" processes, and (3) endogenous spatial attention processes. I discuss each of these sources in turn.

Top-down biases from object recognition processes. Object representations stored in long-term visual memory correspond to familiar objects, such as a familiar face or a word. Behavioral studies have shown that familiar objects can provide top-down feedback to object segregation and object attention processes. In visual segregation, familiar shapes form more salient perceptual regions than less familiar regions. For example, in figure-ground segregation, Peterson and her colleagues (see Peterson, 1994, 1999, for reviews) have shown that a familiar region is more likely to be labeled the 'figure' than a less familiar region (see Figure 5A). This preference for familiar regions as figure is stronger when the stimulus appears in its canonical upright orientation than when the stimulus is rotated 180° away from upright. The difference between upright and 180° rotated displays suggests an influence from object representations on figure-ground processes because object representations are thought to stored in an orientation-specific format (e.g., Joli-

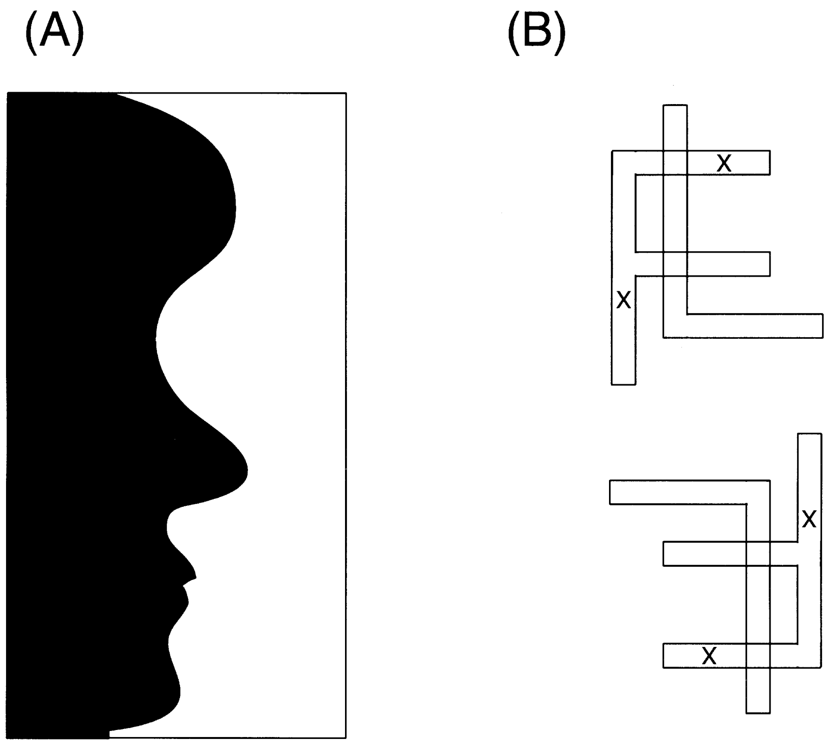

Figure 5. Sample stimuli that have been used to study familiarity effects in (A) figure-ground segregation and (B) object-based attention. In figure- ground segregation, upright familiar regions (the black region in Panel A) are more likely to be perceived as 'figure' than unfamiliar regions (the white region in Panel A). In object attention, familiar objects, such as upright letters, are selected and processed faster than less-familiar objects, such as upside down letters.

coeur, 1985; Tarr, 1995). Although Peterson (1994, 1999) does not explain her behavioral results with top-down feedback from object representations to figure-ground processes, top-down feedback is sufficient to explain her findings (Vecera and O'Reilly, 1998, 2000).

There also is some relevant evidence indicating that object familiarity may influence object attention in a top-down manner. My own research (Vecera and Farah, 1997) suggests that familiar objects (upright letters) are selected by object attention more rapidly than less-familiar objects (rotated letters). In our study, observers viewed displays that contained two overlapping transparent letters, such as those shown in Figure 5B, and were asked to determine if two small Xs were on the same shape or on different shapes. This task requires observers to first segregate the two regions from one another and then to orient attention to the Xs, which are either on the same object or on different objects. Our results suggested an object-based attention effect in this task: Observers were faster to respond when the Xs were on the same object than when they were on different objects. There also was an effect of object familiarity; reaction times were faster to the familiar upright letters than to the less-familiar, rotated letters. Thus, object attention appears to be influenced by shape familiarity in a top-down manner (see Vecera and Farah, 1997, for a fuller discussion). Current research in my lab is attempting to extend these findings by precuing an object's shape on a trial-by-trial basis. We are finding that advance information about an object biases object-based attention toward that object. Observers have an attentional priority for cued shapes compared to uncued shapes. We hypothesize that these results are due to a top-down influence from object representations (perhaps stored in visual working memory) to earlier object segregation and attention processes that organize and select from the visual scene.

Perceptual set. Another mechanism for top-down bias signals in object segregation and attention comes from "perceptual set." Perceptual set loosely refers to the expectancies or goals held by an observer. An observer's perceptual set typically is established by the task-relevant instructions provided by an experimenter in a laboratory setting. In visual search tasks, for example, observers are instructed to search for a specific target; under some theories of visual search, this target is held in some memory store as a "target template" (e.g., Bundesen, 1990; Desimone and Duncan, 1995; Duncan and Humphreys, 1989). In the context of object segregation and attention, perceptual set may be a more general form of a "target template" because perceptual set refers to several possible types of information that observers can use to guide their behavior. Admittedly, the notion of perceptual set is vague and, therefore, the mechanisms that underlie perceptual set may be difficult to study. Nevertheless, there are several studies that have shown that instructions provided by an experimenter can influence the manner in which a stimulus is processed, supporting the notion that top-down information can bias bottom-up object segregation and attention.

Perceptual-set effects have been studied in perceptual segregation tasks. The approach of these studies is to present observers with a bi-stable stimulus (i.e., a stimulus that supports two perceptual interpretations, such as a Necker cube), and then bias the observers to see one of the interpretations of the stimulus. The bias is generally created through the experimenter's instructions. Peterson and Hochberg (1983) developed a procedure for experimentally manipulating and studying perceptual-set (which Peterson and Hochberg, 1983, referred to as "intention" – an observer's intention to perceive a certain region; also see Peterson *et al.*, 1991). In one study, Peterson and Hochberg (1983) instructed observers to try to perceive a region as being in front of another region in a potentially ambiguous stimulus. For example, in Figure 6A, observers were instructed to try to perceive either (1) surface X as a square that occludes another square (region Y in Figure 6A) or (2) surface Y as a "wing chair" in front of a square wall (region X in Figure 6A). Observers could perceive either interpretation; thus, observers' intentions influenced the organization of an ambiguous, bi-stable stimulus. However, it is easier to see region X as an occluder if point 1 in Figure 6A is fixated; it is easier to see

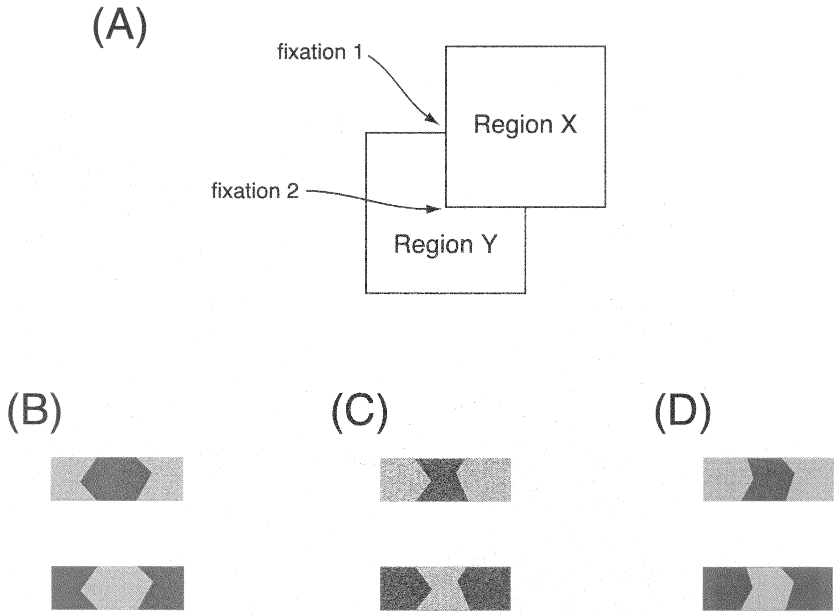

Figure 6. Stimuli in which perceptual set can be manipulated to show top-down effects in object segregation and attention. (A) A display from Peterson and Hochberg's (1983) studies. The experimenter instructs the observer to try to perceive region X as lying in front of region Y or region Y as lying in front of region X. (B) Stimuli used by Baylis and Driver (1993) to the influence of perceptual set on object-based attention. If observers are instructed to perceive the black region, then the top panel involves attending to one object (the central black region) and the bottom panel involves attending to two objects (the flanking black regions). (C) Stimuli used by Gibson (1994) to show that a reversal of convexity can reverse the object attention effects observed by Baylis and Driver (1993). (D) Ambiguous stimuli that do not contain convexity cues from Baylis (1994). See text for additional details.

region Y as a wing chair if point 2 is fixated. Thus, Peterson and Hochberg's (1983) results show that intention alone does not affect the organization of a stimulus. Bottom-up stimulus-level information, namely the intersection fixated (point 1 or 2 in Figure 6A), also influences organization.[3] Peterson and Hochberg's (1983) results are relevant for my biased competition account for two reasons. First, these results show that perceptual set or intention can bias perceptual organization in a top-down manner. Second, these results suggest that *both* top-down intention cues and bottom-up image cues influence observers' final organization and interpretation of a scene. As I outlined earlier, the biased competition account proposes both bottom-up and top-down control over object segregation and attention. Peterson and Hochberg's (1983) results further suggest that these control sources may

[3] Fixation position can be viewed as a bottom-up cue in these studies because fixating a region makes the fixated region more salient in terms of retinal acuity. This higher degree of retinal resolution may bias bottom-up processes toward the fixated region.

operate simultaneously with one another. A cautionary note is in order, however, in using Peterson and Hochberg's (1983) results to support the biased competition account because perceptual reversals (i.e., perceptual biases) in the bi-stable stimuli used by Peterson and Hochberg (1983) may occur at the level of object representation, not object segregation (also see Shulman. 1992). The stimuli in this study may have required minimal segregation because there were few shapes that easily stood out as 'figure.' Further studies should be conducted using bi-stable stimuli that require perceptual segregation, such as ambiguous figure-ground stimuli (see Vecera and O'Reilly, 2000).

In addition to influencing object segregation, perceptual set also influences object-based attention. A now-classic example comes from Neisser and Becklen (1975), who had observers view two spatially overlapping films that were played simultaneously. One film contained two sets of hands playing a "hand game" and the other film contained a basketball game. Because the films were spatially overlapped and because both films were approximately equally salient, there were few, if any, bottom-up cues to favor one film clip over the other. However, Neisser and Becklen (1975) instructed observers to attend to one film or both films. Not only could observers use this instruction to monitor the films, observers also showed an object-based attentional effect; observers were better able to monitor a single film (e.g., the basketball game only) than to divide their attention across both film clips. The experimenter's instructions allowed observers to selectively attend one of the two films, and attention was directed more effectively to a single event (i.e., film clip) than to multiple events.

Similar results have been demonstrated by Baylis (1994; Baylis and Driver, 1993). Observers viewed figure-ground stimuli that contained three regions (Figure 6B) in which the central region was one color (e.g., red) and the two flanking regions were another color (e.g., green). Observers were asked to attend to the two contours shared by the central region and the two adjacent regions and to report which apex was above the other. To manipulate whether observers attended one or two objects, Baylis and Driver (1993; Baylis, 1994) manipulated perceptual set by instructing each observer to pay attention to either the color red or the color green. On some trials, the central region was red, and observers attended to a single object; on other trials, the two flanking regions were red and observers attended to two objects. Baylis and Driver (1993) showed that observers were faster to compare the apices when one object was attended than when two objects were attended. This object-based effect occurred even though the apices were identical between the central and flanking regions. Thus, object-based attention can be influenced by perceptual set: Observers responded faster and more accurately if perceptual set, based on the experimenter's instructions, biased selection of a single region than multiple regions.

Although Baylis and Driver's (1993) results appear consistent with a top-down biasing influence from perceptual set, there is a potential difficulty with their findings. This difficulty resulted in a theoretical exchange in the literature regarding the nature of perceptual set, stimulus-level cues, and object-based attention (see Baylis, 1994; Gibson, 1994). My current proposed biased competition account can provide an explanation of this theoretical exchange through the combination of both bottom-up and top-down sources of information.

This difficulty with Baylis and Driver's (1993) initial results was pointed out by Gibson (1994). The confound in the Baylis and Driver stimuli is visible in Figure 6B. Note that the stimuli in Figure 6B contain a bottom-up cue to segregation; specifically, the central region was always convex and the two adjacent regions were always concave. Gibson (1994) correctly noted that convexity is a salient determinant of figure-ground segregation and, therefore, Baylis and Driver's (1993) results could have been caused by easier segregation of the central region, not object-based attention to a single region. To argue his point, Gibson (1994) conducted a study in which he reversed the convexity in the stimuli by making the two flanking regions convex and the central region concave (Figure 6C). Interestingly, Gibson (1994) found that reversing the convexity also reversed the so-called object attention effect: Observers who viewed the stimuli in Figure 6C were faster to compare the apices when the two flanking regions were perceived as figure than when the central region was perceived as figure, a reversal of Baylis and Driver's (1993) results. Gibson's (1994) results suggest that salient bottom-up information, such as convex regions, may grab or capture attention more effectively than less salient bottom-up information. Object-based attention can be applied to multiple regions more effectively than to a single region, provided there is stimulus information (e.g., convexity) that favors the multiple-region interpretation of the scene over the single-region interpretation. In subsequent research, however, Baylis (1994) demonstrated an object-based effect in displays that had equal convexity between the central region and adjacent regions (Figure 6D). Thus, all else being equal, when object-based attention is biased by perceptual set, it is easier to select a single region than multiple regions.

Under the biased competition account, the exchange between Gibson (1994) and Baylis (1994) is more than a minor disagreement about a stimulus confound. What the Baylis/Gibson exchange clearly shows is that both bottom-up information and top-down perceptual set information can influence object attention. Specifically, if bottom-up information (e.g., convexity) favors a two-object interpretation of a scene, attention may more easily select two objects than one object (Gibson's, 1994, results). However, bottom-up information alone does not guide object attention because equating all bottom-up cues allows perceptual set to influence object attention, which selects a single region more easily than two regions (Baylis', 1994, results).

Top-down biases from spatial attention. The various studies that have used perceptual set to bias observers raise an important question: What is the mechanism that underlies perceptual set? Perceptual set can be likened to an expectancy effect or a task demand imposed by the experimenter, but what are the visual processes involved in expectation or task demand? One possibility is that perceptual set involves the voluntary (or endogenous) control of spatial attention. That is, observers may voluntarily allocate their spatial attention to a region on the basis of instructions provided by the experimenter. For example, in Baylis and Driver's (1993) studies, observers instructed to perceive the red region as figure may shift spatial attention to the red region, biasing this region to be perceived as figure.

This account of perceptual set does find some empirical support in studies that show that overt attention – eye fixation location – can influence the interpretation of a scene (e.g., Peterson and Gibson, 1994a; Peterson and Hochberg, 1983). More important, this interpretation of perceptual set suggests that spatial attention more generally may provide another top-down biasing signal in the biased competition account. I now turn to studies that support a top-down bias from spatial attention in both object segregation and object attention. Recent evidence suggests that voluntary spatial attention may bias some forms of object based-segregation and attention.

Simple spatial precues appear to exert top-down effects on object segregation and attention. Spatial attention can be viewed as a top-down signal because object segregation and attention are thought to be preattentive processes that operate before the level of spatial attention (for this view, see Julesz, 1984; Neisser, 1967; Treisman, 1988). Further, the top-down role assigned to these cues is supported by comparing predictive precues that invoke voluntary, endogenous spatial orienting with nonpredicitve precues that invoke reflexive, exogenous spatial orienting. In the case of object segregation, Driver and Baylis (1996) demonstrated that certain forms of spatial attention can influence figure-ground segregation. Observers performed a contour-matching task in which they had to match the contour of one of the regions from the figure-ground display; in previous experiments, Driver and Baylis (1996) found that observers matched the contour of the perceived figure faster than the contour of the perceived ground. In two experiments, Driver and Baylis (1996) investigated the role of spatial attention in figure-ground assignment. Observers viewed ambiguous figure-ground displays in which there were no bottom-up cues that could bias segregation. Prior to the figure-ground display, a spatial precue appeared. The spatial precue either did not predict or did predict the region that observers would need to match. Driver and Baylis (1996) found that predictive spatial precues influenced figure-ground segregation: Observers were faster to match the spatially precued region only when the cue was predictive, suggesting that voluntary (endogenous) spatial attention influenced segregation. Non-predictive spatial precues did not influence segregation. Because only the predictive precue influenced figure-ground segregation, Driver and Baylis (1996) reasoned that voluntary, or endogenous, control of spatial attention influenced segregation. These results are similar to Peterson and Hochberg's (1983) "intention" effects discussed previously. The advantage of Driver and Baylis' (1996) technique, however, is that the "intention" comes in the form of a spatial precue that may be a better-understood manipulation in terms of the visual processes evoked by a precue and that is presented on a trial-by-trial basis.

Spatial attention also influences object selection. Perhaps one of the most straightforward tasks that shows how a spatial precue can influence object selection is the cued detection task developed by Egly *et al.* (1994) that was discussed earlier (see Figure 4). Recall that the target usually appears at the cued location (Figure 4A), and when it appears at an uncued location it may appear within the cued object (Figure 4B) or in the other, uncued object (Figure 4C). The important result is that observers are faster to respond to targets appearing in the uncued end of the cued rectangle than at either end of the uncued rectangle. This object selection may involve a top-down component in my biased competition account in addition to the bottom-up cues of closure and common region that define the two rectangles as separate objects. The spatial precue may allow spatial attention to bias attention in a top-down manner: Because the objects (the two rectangles) are identical, there is no bottom-up bias to favor attention to one object over the other (as in the two lines in Figure 2A). Thus, if there was no precue, each object should have an equal chance of being attended, and presumably object attention would select randomly between the two objects. The spatial precue acts to bias object attention toward the cued object, allowing observers to respond faster to targets appearing in the cued object than in the uncued object.

The top-down role of spatial attention has been demonstrated more recently in a series of studies reported by Lavie and Driver (1996), who investigated the combined roles of spatial attention and object attention by combining spatial precues with an object-based attention task. Specifically, observers viewed two overlapping lines (objects) and discriminated features that appeared on the same line (same object) or different lines (different objects). In their final study, Lavie and Driver (1996) precued the adjacent ends of the *different* objects; the spatial precue summoned attention to a subregion of the display that contained features from the different objects. The spatial precue abolished object-based attention. Observers were just as fast to report features on the same object as features on different objects. Because spatial attention was restricted to a subregion of the display, the entirety of both objects was no longer in a spatially attended region. Lavie and Driver (1996) concluded that object attention only occurs within a spatially attended region.

Although Lavie and Driver's results are consistent with an influence of top-down input from spatial attention, their explanation does not appeal to the mechanisms of the biased competition account that I outlined earlier. However, the biased competition account can readily explain Lavie and Driver's (1996) results, without the need to suggest that object attention only occurs within a spatially attended region. One additional piece of information is needed to explain Lavie and Driver's (1996) results with the biased competition account: The spatial precues used in their study were highly predictive. On 70% of the trials, observers were precued to attend to different objects. The top-down inputs from spatial attention were strong (i.e., highly predictive) and may have biased observers to adopt a spatial orienting strategy in which the objects were ignored. Recent results from my lab (Gilds and Vecera, submitted) suggest that less-predictive spatial precues, in which the top-down inputs are weaker than those used by Lavie and Driver (1996), may not abolish object-based attention. In my research, the bottom-up object segregation cues can compete against non-predictive spatial precues, allowing object attention effects to be observed. In Lavie and Driver's (1996) research, the top-down cues were strong enough to override the bottom-up object segregation cues. My findings suggest that both bottom-up and top-down factors influence object attention and that the *strengths* of the top-down inputs also must be considered in the biased competition account.

2.3. SUMMARY

Both bottom-up, stimulus-driven cues and top-down, goal-driven cues can influence object segregation and attention. The organization or segregation of a scene and the allocation of object-based attention within this scene is dependent upon the cooperation and competition of bottom-up and top-down cues. Presumably, both sources of information influence behavior in important ways. For example, our visual systems need to be sensitive to bottom-up cues because these cues contain information regarding salient information in the external world. To maintain flexible behavior that is not entirely stimulus driven, however, visual processing must be modified by an observer's goals.

In an attempt to keep the focus this paper relatively narrow, I have presented only behavioral results that support a biased competition account of object segregation and attention. Importantly, however, there are many recent results from cognitive neuroscience that appear to further support this account, such as findings that demonstrate neurons in primary visual cortex are sensitive to both figure-ground relations (e.g., Lamme, 1995; Zipser *et al.*, 1996) and object-based attention (e.g., Roelfsema *et al.*, 1998). Instead of attempting to review the neurophysiological and neuropsychological results relevant to the biased competition account, I now turn to a discussion of my own computational modeling efforts that are consistent with the biased competition account I have outlined. My goal in the following section is to show that the principles of biased competition can be implemented in a simple neural network model. The model is relatively narrow in scope, primarily explaining findings from figure-ground segregation. However, there are general computational principles that allow the model to guide theorizing about object segregation and attention more generally.

3. A Model for Biased Competition in Figure-Ground Segregation

How does object familiarity influence figure-ground segregation? This question has guided some of my recent computational modeling efforts (Vecera and O'Reilly, 1998, 2000). The model that I present in this section is useful because it embodies the principles of the biased competition account reviewed above, demonstrating that these principles are sufficient to explain a range of data on object attention and segregation. The model not only simulates existing behavioral data, but it also makes predictions for future studies (see Vecera and O'Reilly, 2000). In general, neural network modeling provides a useful research tool because these models can explain the cognitive processes – potential computational mechanisms – that underlie global behavior (Mozer and Sitton, 1998).

As reviewed earlier, several behavioral studies demonstrate that a region that depicts a familiar object is more likely to be assigned as the 'figure' than a region that does not depict a familiar object (see Peterson, 1994, 1999, for reviews). This "familiarity effect" is reduced for stimuli in which the familiar region is upside down, suggesting that object representations influence figure-ground segregation. Although these empirical results themselves are straightforward and not debated, the interpretation of these results is more contentious. One account of familiarity effects in figure-ground segregation has been proposed by Peterson and her colleagues (Peterson, 1994, 1999; Peterson and Gibson, 1993, 1994a, 1994b; Peterson *et al.*, 1991). Peterson's account hypothesizes that object recognition occurs prior to figure-ground segregation. Objects are preliminarily identified by a "prefigural" recognition process that operates on the luminance contours present in an image. The prefigural account stands in contrast to most hierarchical accounts of perceptual segregation and object recognition that place segregation processes prior to recognition processes (e.g., Biederman, 1987; Kosslyn, 1987; Marr, 1992; Palmer and Rock, 1994); under these hierarchical accounts, object representations receive input from segregation processes. Peterson (1994, 1999) proposes that prefigural recognition processes receive input from edge (contour) representation processes, not from regions that have been assigned as 'figure' by figure-ground segregation processes. Once prefigural object recognition has occurred, the outputs of the prefigural recognition system influence figure-ground segregation (see Peterson, 1999, for the most explicit discussion of her prefigural account).

An alternative account of the familiarity effects in figure-ground segregation is an interactive account that my colleagues and I have proposed (Vecera and Farah, 1997; Vecera and O'Reilly, 1998, 2000). Following the principles of parallel distributed processing (PDP) neural network models, we hypothesized that top-down feedback from object representations to an earlier figure-ground segregation process could explain the empirical results presented by Peterson and colleagues. We developed a straightforward neural network model that did not require "prefigural" recognition processes; instead, figure-ground segregation received top-down feedback from object representations. This interactive model explained a range of previously published behavioral data (Vecera and O'Reilly, 1998) and made accurate predictions of the results from new studies (Vecera and O'Reilly, 2000).

The interactive model of figure-ground segregation is relevant to the biased competition account of object segregation and attention. Although not originally intended to explain a wide range of object segregation and attention data, our interactive model embodies the components of biased competition and provides specific computational mechanisms for the more general account that I discussed above. I now turn to a more detailed discussion of the model to illustrate the processes required for biased competition in object segregation and attention.

3.1. THE MODEL

The network we developed to investigate top-down effects in figure-ground segregation was based on a model developed and investigated by Sejnowski, Hinton, and colleagues (see Kienker *et al.*, 1986; Sejnowski and Hinton, 1987). We adopted this model as a framework for our simulations because it is capable of exhibiting some basic figure-ground organization. Our model can be viewed as an implementation of several general principles of figure-ground segregation. These general principles are meant to be noncontroversial and to capture the requirements of any system that performs object segregation:

- *Principle 1*: Figure-ground segregation selects one interpretation (i.e., segregation solution) in a scene at any given time.
- *Principle 2*: Figures tend to be contiguous, connected regions.
- *Principle 3*: Segregation, like attention, has a selective function in that it favors one region (the figure) over another (the ground). This principle illustrates the close relationship between object segregation and object attention – both have a selective function.

Similar principles have been formulated for other visual processes, such as spatial attention (see Mozer and Sitton, 1998, for an excellent example). These three general principles of object segregation can be translated into "rules" implemented in the model. These rules are soft constraints on the model's operation, not rigid rules:

- *Rule 1*: The locations occupied by the 'figure' should be activated (the model makes no predictions regarding the state of the 'ground'). This activation corresponds to the interpretation of figure and ground stated in Principle 1.
- *Rule 2*: Connections among each figure unit and its neighbors (Figure 7C) should allow these units to form connected figure regions; these connections form the basis of the cooperation stated in Principle 2.
- *Rule 3*: Competition occurs between the two possible interpretations of the boundaries shared by two regions; one consequence of this competition is to activate the figure relative to the ground. This rule implements the selective function of segregation stated in Principle 3.

These rules of figure-ground segregation are implemented in our interactive model, which is depicted in Figure 7A. The model has three processing layers. The first four layers process the boundaries contained in the image; these boundary

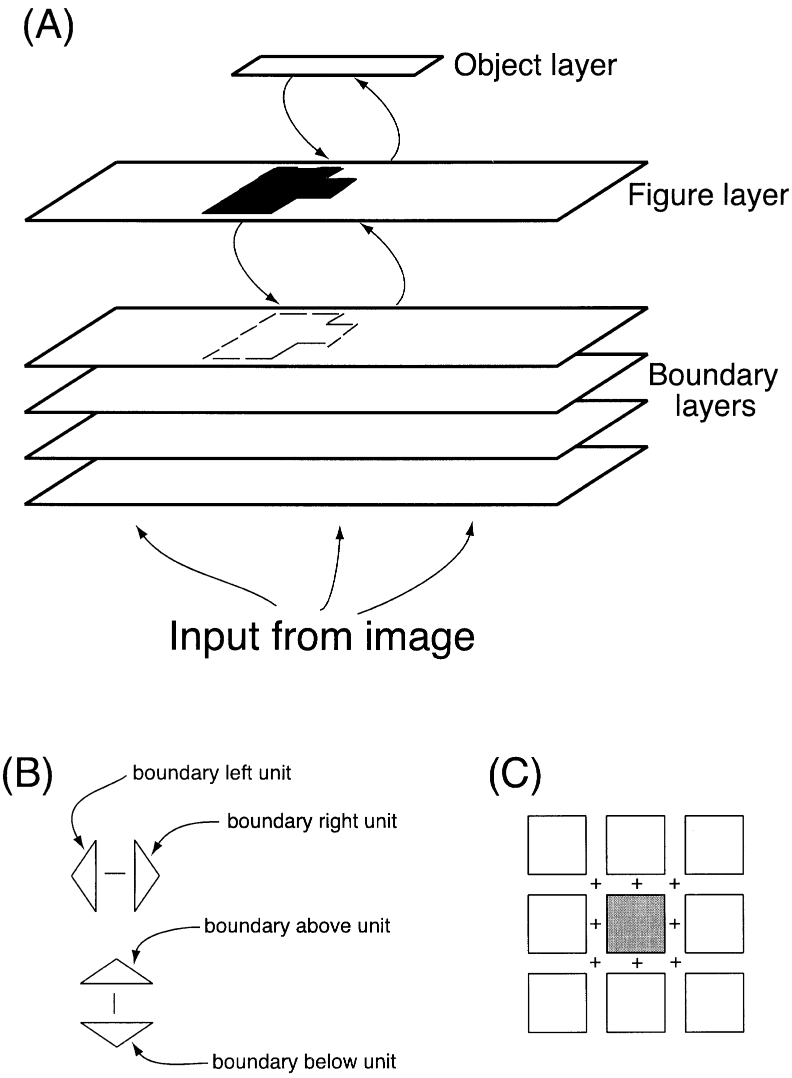

Figure 7. (A) The Vecera and O'Reilly (1998, 2000) interactive model of figure-ground segregation. (B) Inhibitory connections that implement competition among opposing boundary units. (C) Excitatory connections that implement cooperation among a figure unit and its 8 nearest neighbors.

layers correspond to simple image features (edges). These layers are followed by a layer that represents which surfaces are 'figure.' Finally, there are object representations that code for familiar shapes. Thus, consistent with a biased competition account, the figure-ground units receive two types of information: bottom-up information from the edges present in the visual field and top-down information from internally-stored object representations. Note that although the network can be divided into these distinct processing layers, this hierarchical organization is blurred by the interactive processing in the network. As soon as information begins to propagate within the network the information flow is bi-directional.

In addition to the network's general configuration, there are two patterns of connectivity that are relevant for understanding how biased competition can arise from an interactive PDP model. These patterns of connectivity are depicted in panels B and C of Figure 7. The first pattern of connectivity is the competition that occurs between the boundary layers. Each of the four boundary layers codes for one of four types of edges that could be present in a visual scene; each edge, or boundary, corresponded to a simple visual contour and the side of that contour on which the figural region is most likely to be located. For example, activation of a *Boundary Left* unit signaled the presence of an edge that had the figural region to its left. Similarly, activation of a *Boundary Right* unit signaled the presence of a vertical edge that had the figural region to its right. Activation of either a *Boundary Above* or *Boundary Below* unit signaled the presence of a horizontal edge that had the figural region above or below the edge, respectively. Importantly, opposing pairs of boundary units were mutually inhibitory, as illustrated in Figure 7B. A Boundary Left unit at one location in the visual field inhibited the Boundary Right unit at the same location because the figure cannot lie to both the left and right of the edge; the figure must lie to *either* the left or the right. The Boundary Above and Boundary Below units are also mutually inhibitory. These inhibitory connections implement the bottom-up competition that arises between boundaries that are shared by the two regions in a figure-ground display. These inhibitory connections implement, in part, the "selectivity rule" (Rule 3 above) in which boundaries or regions are activated selectively over others.

The second relevant pattern of connectivity in the network, shown in Figure 7C, is the cooperation among figure units. Each of the Figure units was excitatorily connected to its eight nearest neighbors in order to implement the constraint that the figural region tends to be connected and continuous (Rule 2 above). Figure units are influenced by (1) bottom-up input from boundaries, (2) lateral cooperation among figure units, and (3) top-down inputs from object representations. The consequence of these three influences is to activate figural regions (Rule 1 above).

3.2. BIASED COMPETITION IN THE MODEL

Our model fits well with the biased competition account of object segregation and attention in several respects. The figure-ground display presented to the model provides the bottom-up information that must be segregated. The input provided in our original simulations was ambiguous in that neither region contained any bottom-up cues for segregation (Vecera and O'Reilly, 1998). That is, the area, convexity, symmetry, and so forth were identical between the two regions. The two regions in the display compete for segregation – that is, the two regions compete because only one can eventually become the foreground figure.

Figure 8. Stimuli used by Vecera and O'Reilly (2000) to manipulate both bottom-up information (smaller area of the white region) and top-down information (familiarity of the black region).

Because the bottom-up input is insufficient to resolve this competition, there must be top-down inputs to bias the bottom-up competition. In our model, the top-down inputs came from object representations: One of the two regions in the figure-ground display corresponded to a familiar object that was represented by one of the object units. If the stimulus display appeared in its normal, upright orientation, this top-down input biased segregation – the model assigned the familiar region to be the 'figure' more often than the unfamiliar region. This preference for the familiar region was reduced if the display was upside down. The difference between upright and upside down displays demonstrates that orientation-dependent object representations can influence figure-ground segregation in a top-down manner.

A more interesting case of biased competition arises in a later simulation study that we conducted with our model (Vecera and O'Reilly, 2000). Although this follow-up simulation was aimed at responding to a critique of the model (Peterson, 1999), this simulation shows the combined effects of both bottom-up and top-down sources of information. We asked if the top-down information from object representations could override biased bottom-up information. Under a biased competition account, not only do different regions compete to be 'figure,' but different cues may also compete. We studied a condition in which the bottom-up information favored the unfamiliar region as figure but the top-down information favored the familiar region as figure. Figure 8 provides an illustration of this scenario; the familiar region (the face profile in black) is larger than the unfamiliar region. Because size, or area, is a well-known gestalt cue for segregation, bottom-up cues favor the unfamiliar white region in Figure 8 and top-down cues favor the familiar black region. The results of this simulation found that both bottom-up and top-down sources of information can influence segregation. Specifically, when the unfamiliar region was smaller than the familiar region, the model had an overall bottom-up bias to choose the smaller, unfamiliar region as figure. However, the model continued to show a preference for upright familiar regions than for upside down familiar regions. That is, the bottom-up bias was different between the upright and upside down displays, thereby showing a top-down influence from object representations (Vecera and O'Reilly, 2000).

Although I have discussed our simulation studies briefly, there are three main points illustrated by our model: (1) competition occurs between regions to determine which region should be the figure; (2) this competition is biased by top-down feedback from object representations; and (3) this competition also is biased by bottom-up image cues, such as size. Thus, the model appears to embody the general principles of the biased competition account of object segregation that I outlined earlier in this article. Of course, there are several important points that deserve discussion. For example, our simulation results to date do not bear directly on object-based attention. Although the model could, in principle, explain biased competition in object attention, simulation results would make any explanation more complete. Also, there are a range of potential top-down biasing cues that could influence segregation, and our model explores only one of these cues (object familiarity). Other cues, such as spatial attention or perceptual set, should be examined. Finally, the interactive PDP model I have presented may be only one of many possible models that might be consistent with the biased competition account I have outlined. There is a well-known computational problem with interactive models: such models often reach suboptimal solutions that are inconsistent with a coherent interpretation of a visual scene (see Hinton and Lang, 1985; Mozer *et al.*, 1992; Vecera and O'Reilly, 1998). Models that do not have the difficulties associated with interactive constraint satisfaction (e.g., Mozer, 1999; Mozer and Sitton, 1998; Mozer *et al.*, 1992) may explain the mechanisms of biased competition. Such models should be explored to determine if they are compatible with the general approach of biased competition in object segregation and attention.

4. Conclusions

Any complex visual scene is likely to have multiple interpretations. Some objects may be more salient than others. Some objects may be more relevant for a current goal than others. Object salience and goal relevance characterize the two main influences on object segregation and attention under the biased competition account

I have proposed here. In this review, I have followed the visual search literature and developed a biased competition account of object segregation and attention in an attempt to organize the various segregation and attention results under a common framework. In general, the objects or regions that ultimately guide behavior arise from biases that emerge from either the environment (bottom-up biases) or from the observer's goals (top-down biases).

The account I have presented here may be advantageous for several reasons. First, my biased competition account provides a framework for integrating the multitude of behavioral studies on object segregation and object-based attention, which highlights the similarities between segregation and object attention processes. Second, the biased competition account can suggest new experimental studies, such as the attentional capture by regions defined as 'figure' by segregation processes. Third, the general principles of biased competition can be instantiated concretely in PDP models, and the principles of PDP models illustrate how biased competition may characterize different visual processing tasks (e.g., visual search, object attention, etc.). The theoretical usefulness of the biased competition account I have presented will be determined by this account's ability to accommodate the results from relevant neurobiological studies (e.g., Lamme, 1995; Roelfsema *et al.*, 1998; Zipser *et al.*, 1996) and future studies.

Acknowledgements

This paper was prepared with partial support from grant BCS 99-10727 from the National Science Foundation. Thanks to Steve Luck, Mike Mozer, Randy O'Reilly, and Pieter Roelfsema for discussions that have shaped the ideas in this paper. Thanks also to Maureen Marron, George Graham, Geoff Woodman, and an anonymous reviewer for comments on a previous version of the paper.

References

Baylis, G.C., 1994: Visual attention and objects: Two-object cost with equal convexity, *Journal of Experimental Psychology: Human Perception and Performance* **20**, 208–212.

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

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

Biederman, I., 1987: Recognition-by-components: A theory of human image understanding, *Psychological Review* **94**, 115–147.

Behrmann, M. and Haimson, C., 1999: The cognitive neuroscience of visual attention, *Current Opinion in Neurobiology* **9**, 158–163.

Behrmann, M., Zemel, R. and Mozer, M.C., 1998: Object-based attention and occlusion: Evidence from normal subjects and a computational model, *Journal of Experimental Psychology: Human Perception and Performance* **24**, 1011–1036.

Bundesen, C., 1990: A theory of visual attention, *Psychological Review* **97**, 523–547.

Cohen, J.D., Dunbar, K. and McClelland, J.L., 1990: On the control of automatic processes: A parallel distributed processing account of the Stroop effect, *Psychological Review* **97**, 332–361.

Cohen, J.D. and Huston, T.A., 1994: Progress in the use of interactive models for understanding attention and performance, in C. Umiltá and M. Moscovitch (eds.), *Attention and Performance XV*. Cambridge, MA: MIT Press, pp. 453–476.
Desimone, R. and Duncan, J., 1995: Neural mechanisms of selective visual attention, *Annual Review of Neuroscience* **18**, 193–222.
Driver, J. and Baylis, G.C., 1989: Movement and visual attention: The spotlight metaphor breaks down, *Journal of Experimental Psychology: Human Perception and Performance* **15**, 448–456.
Driver, J. and Baylis, G.C., 1996: Edge-assignment and figure-ground segmentation in short-term visual matchingm, *Cognitive Psychology* **31**, 248–306.
Driver, J. and Baylis, G.C., 1998: Attention and visual object segregationm, in R. Parasuraman (ed.), *The Attentive Brain*. Cambridge, MA: MIT Press, pp. 299–325.
Duncan, J., 1984: Selective attention and the organization of visual information, *Journal of Experimental Psychology: General* **113**, 501–517.
Duncan, J. and Humphreys, G.W., 1989: Visual search and stimulus similarity, *Psychological Review* **96**, 433–458.
Egly, R., Driver, J. and Rafal, R.D., 1994: Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects, *Journal of Experimental Psychology: General* **123**, 161–177.
Eriksen, B.A. and 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. and Hoffman, J., 1973: The extent of processing of noise elements during selective encoding from visual displays, *Perception & Psychophysics* **14**, 155–160.
Eriksen, C.W. and St. James, J.D., 1986: Visual attention within and around the field of focal attention: A zoom lens model, *Perception & Psychophysics* **40**, 225–240.
Egeth, H.E. and Yantis, S., 1997: Visual attention: Control, representation, and time course, *Annual Review of Psychology* **48**, 269–297.
Egly, R., Driver, J. and Rafal, R.D., 1994: Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects, *Journal of Experimental Psychology: General* **123**, 161–177.
Gibson, B.S., 1994: Visual attention and objects: One versus two or convex versus concave? *Journal of Experimental Psychology: Human Perception and Performance* **20**, 203–207.
Gilds, K.S. and Vecera, S.P., submitted: *The Influence of Spatial Attention on Object-based Selection*. Manuscript submitted for publication.
Harter, M.R. and Aine, C.J., 1984: Brain mechanisms of visual selective attention, in R. Parasuraman and D.R. Davies (eds.), *Varieties of Attention*. Orlando, FL: Academic Press, pp. 293–321.
Hinton, G.E. and Lang, K.J., 1985: Shape recognition and illusory conjunctions, in *Ninth Annual Joint Conference on Artificial Intelligence*. Los Altos, CA.: Morgan-Kaufmann, pp. 252–259.
Hoffman, D.D. and Richards, W.A., 1984: Parts of recognition, *Cognition* **18**, 65–96.
Hoffman, D.D. and Singh, M., 1997: Salience of visual parts, *Cognition* **63**, 29–78.
Jolicoeur, P., 1985: The time to name disoriented natural objects, *Memory & Cognition* **13**, 289–303.
Julesz, B., 1984: A brief outline of the texton theory of human vision, *Trends in Neurosciences* **6**, 41–45.
Kanwisher, N.G. and Driver, J., 1992: Objects, attributes, and visual attention: Which, what, and where, *Current Directions in Psychological Science* **1**, 26–31.
Kienker, P.K., Sejnowski, T.J., Hinton, G.E. and Schumacher, L.E., 1986: Separating figure from ground with a parallel network, *Perception* **15**, 197–216.
Kosslyn, S.M., 1987: Seeing and imagining in the cerebral hemispheres: A computational approach, *Psychological Review* **94**, 148–175.
Kramer, A.F. and Jacobson, A., 1991: Perceptual organization and focussed attention: The role of objects and proximity in visual processing, *Perception & Psychophysics* **50**, 267–284.

Kramer, A.F., Tham, M.-P. and Yeh, Y.-Y., 1991: Movement and focused attention: A failure to replicate, *Perception & Psychophysics* **50**, 537–546.
LaBerge, D. and Brown, V., 1989: Theory of attentional operations in shape identification, *Psychological Review* **96**, 101–124.
Lamme, V.A.F., 1995: The neurophysiology of figure-ground segregation in primary visual cortex, *Journal of Neuroscience* **15**, 1605–1615.
Lavie, N. and Driver, J., 1996: On the spatial extent of attention in object-based visual selection, *Perception & Psychophysics* **58**, 1238–1251.
Lowe, D.G., 1985: *Perceptual Organization and Visual Recognition*. Boston: Kluwer.
Lowe, D.G., 1987: Three-dimensional object recognition from single two-dimensional images, *Artificial Intelligence* **31**, 355–395.
Marr, D., 1982: *Vision*. San Francisco: Freeman.
Mozer, M.C. and Sitton, M., 1998: Computational modeling of spatial attention, in H. Pashler (ed.), *Attention*. Hove, UK: Psychology Press, pp. 341–393.
Mozer, M.C., Zemel, R.S., Behrmann, M. and Williams, C.K.I., 1992: Learning to segment images using dynamic feature binding, *Neural Computation* **4**, 650–665.
Neisser, U., 1967: *Cognitive Psychology*. New York: Appleton-Century-Crofts.
Neisser, U. and Becklen, R., 1975: Selective looking: Attending to visually specified events, *Cognitive Psychology* **7**, 480–494.
O'Craven, K.M., Downing, P.E. and Kanwisher, N., 1999: fMRI evidence for objects as the units of attentional selection, *Nature* **401**, 584–587.
Palmer, S.E. and Rock, I., 1994: Rethinking perceptual organization: The role of uniform connectedness, *Psychonomic Bulletin & Review* **1**, 29–55.
Peterson, M.A., 1994: Object recognition processes can and do operate before figure-ground organization, *Current Directions in Psychological Science* **3**, 105–111.
Peterson, M.A., 1999: What's in a stage name? Comment on Vecera and O'Reilly (1998), *Journal of Experimental Psychology: Human Perception and Performance* **25**, 276–286.
Peterson, M.A. and Gibson, B.S., 1991: The initial identification of figure-ground relationships: Contributions from shape recognition processes, *Bulletin of the Psychonomic Society* **29**, 199–202.
Peterson, M.A. and Gibson, B.S., 1993: Shape recognition inputs to figure-ground organization in three-dimensional displays, *Cognitive Psychology* **25**, 383–429.
Peterson, M.A. and Gibson, B.S., 1994a: Object recognition contributions to figure-ground organization: Operations on outlines and subjective contours, *Perception & Psychophysics* **56**, 551–564.
Peterson, M.A. and Gibson, B.S., 1994b: Must figure-ground organization precede object recognition? An assumption in peril, *Psychological Science* **5**, 253–259.
Peterson, M.A., Harvey, E.M. and Weidenbacher, H., 1991: Shape recognition contributions to figure-ground organization: Which routes count? *Journal of Experimental Psychology: Human Perception and Performance* **17**, 1075–1089.
Peterson, M.A. and Hochberg, J., 1983: Opposed-set measurement procedure: A quantitative analysis of the role of local cues and intention in form perception, *Journal of Experimental Psychology: Human Perception and Performance* **9**, 183–193.
Pomerantz, J.R. and Kubovy, M., 1986: Theoretical approaches to perceptual organization, in K.R. Boff, L. Kaufman and J.P. Thomas (eds), *Handbook of Perception and Human Performance*, volume II. New York: Wiley, pp. 36.1–36.46.
Posner, M.I., 1980: Orienting of attention, *Quarterly Journal of Experimental Psychology* **32A**, 3–25.
Posner, M.I., Snyder, C.R.R. and Davidson, B.J., 1980: Attention and the detection of signals, *Journal of Experimental Psychology: General* **109**, 160–174.
Reicher, G.M., Snyder, C.R.R. and Richards, J.T., 1976: Familiarity of background characters in visual scanning, *Journal of Experimental Psychology: Human Perception and Performance* **2**, 522–530.

Roelfsema, P.R, Lamme, V.A.F. and Spekreijse, H., 1998: Object-based attention in the primary visual cortex of the macaque monkey, *Nature* **395**, 376–381.
Rock, I., 1975: *An Introduction to Perception*. New York: MacMillan.
Rubin, E., 1915/1958: Figure and ground, in D.C. Beardslee and M. Wertheimer (eds), *Readings in Perception*. Princeton, NJ: Van Nostrand, pp. 194–203. (Original work published 1915.)
Sagi, D. and Julesz, B., 1984: Detection versus discrimination of visual orientation, *Perception* **13**, 619–628.
Sandon, P.A., 1990: Simulating visual attention, *Journal of Cognitive Neuroscience* **2**, 213–231.
Schyns, P.G. and Murphy, G.L., 1994: The ontogeny of part representation in object concepts, in D.L. Medin (ed.), *The Psychology of Learning and Motivation*, Volume 31. San Diego: Academic Press, pp. 305–349.
Sejnowski, T.J. and Hinton, G.E., 1987: Separating figure from ground with a boltzmann machine, in M.A. Arbib and A.R. Hanson (eds), *Vision, Brain, and Cooperative Computation*. Cambridge, MA: MIT, pp. 703–724.
Shulman, G.L., 1992: Attentional modulation of a figural aftereffect, *Perception* **21**, 7–19.
Sperling, G. and Weichselgartner, E., 1995: Episodic theory of the dynamics of spatial attention, *Psychological Review* **102**, 503–532.
Tarr, M.J., 1995: Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three- dimensional objects, *Psychonomic Bulletin & Review* **2**, 55–82.
Treisman, A., 1988: Features and objects. The fourteenth Bartlett memorial lecture, *Quarterly Journal of Experimental Psychology* **40A**, 201–237.
Treisman, A. and Gelade, G.A., 1980: A feature-integration theory of attention, *Cognitive Psychology* **12**, 96–136.
Vecera, S.P., 1994: Grouped locations and object-based attention: Comment on Egly, Driver, & Rafal (1994), *Journal of Experimental Psychology: General* **123**, 316–320.
Vecera, S.P., Behrmann, M. and Filapek, J.C., in press: Attending to the parts of a single object: Part-based selection limitations, *Perception & Psychophysics*.
Vecera, S.P., Behrmann, M. and McGoldrick, J., 2000: Selective attention to the parts of an object, *Psychonomic Bulletin & Review* **7**, 301–308.
Vecera, S.P. and Farah, M.J., 1994: Does visual attention select objects or locations? *Journal of Experimental Psychology: General* **123**, 146–160.
Vecera, S.P. and Farah, M.J., 1997: Is visual image segmentation a bottom-up or an interactive process? *Perception & Psychophysics* **59**, 1280–1296.
Vecera, S.P. and Luck, S.J., 2000: Attention, in V.S. Ramachandran (ed.), *Encyclopedia of the Human Brain*. San Diego: Academic Press.
Vecera, S.P. and O'Reilly, R.C., 1998: Figure-ground organization and object recognition processes: An interactive account, *Journal of Experimental Psychology: Human Perception and Performance* **24**, 441–462.
Vecera, S.P. and O'Reilly, R.C., 2000: Graded effects in hierarchical figure-ground organization: Reply to Peterson (1999), *Journal of Experimental Psychology: Human Perception and Performance* **26**, 1221–1231.
Watson, S.E. and Kramer, A.F., 1999: Object-based visual selective attention and perceptual organization, *Perception & Psychophysics* **61**, 31–49.
Wolfe, J.M., 1994: Guided search 2.0: A revised model of visual search, *Psychonomic Bulletin & Review* **1**, 202–238.
Wolfe, J.M., 1998: Visual search, in H. Pashler (ed.), *Attention*. Hove, UK: Psychology Press, pp. 15–73.
Wong, E. and Weisstein, N., 1983: Sharp targets are detected better against a figure, and blurred targets are detected better against a background, *Journal of Experimental Psychology: Human Perception and Performance* **9**, 194–202.

Woodman, G.F., Vogel, E.K. and Luck, S.J., submitted: *Visual Search Remains Efficient when Working Memory is Full*. Manuscript submitted for publication.
Yantis, S., 1998: Control of visual attention, in H. Pashler (ed.), *Attention*. Hove, UK: Psychology Press, pp. 223–256.
Yantis, S. and Jonides, J., 1984: Abrupt visual onsets and selective attention: Evidence from visual search, *Journal of Experimental Psychology: Human Perception and Performance* **10**, 601–621.
Zipser, K., Lamme, V.A.F. and Schiller, P.H., 1996: Contextual modulation in primary visual cortex, *Journal of Neuroscience* **16**, 7376–7389.