6

Object Perception

Mary A. Peterson

What Is Object Perception?

Visual perception in general, and the visual perception of objects in particular, seems so immediate and effortless that it is difficult to comprehend its complexity. Consider, for example, the non-trivial question of what constitutes an object. Both philosophers and psychologists have occupied themselves with trying to find the necessary and sufficient properties of objects (Hirsch, 1982; Wiggins, 1980). Based on infant research, Elizabeth Spelke and her colleagues have defined objects as solid entities that (a) exhibit spatio-temporal continuity, (b) cohere within their boundaries when they move, and (c) move only when contacted by another object (Spelke, 1990; Spelke, Guthiel, & Van de Walle, 1995). On Spelke's definition, animals and immaterial entities are excluded from the object category, and well they should be, at least for common usage of the term "object." Bloom (1996) correctly excludes other entities, including puddles, shadows, holes, illusory objects, and parts of objects (e.g., fingers and cup handles). Ittelson (1996) excludes pictures of objects because they are two-dimensional (2-D) rather than three-dimensional (3-D), as real objects are.

Distinctions between those entities that count as real objects and those that do not are critical if one is concerned with classifying those entities we *judge or know* to be real objects. However, most investigators of visual perception use the term "object perception" both more broadly and more narrowly than it is used by the authors discussed above. The term *object perception* is used more broadly by perception psychologists because it encompasses processes that

- integrate within and segregate between elements in the visual input;
- assign shape and 3-D structure to some of those elements;
- permit recognition of previously-seen shaped entities; and
- determine the manner in which attention is focused on the shaped entities.

Hence, investigators of visual perception typically use the term *object perception* to apply to both animate and inanimate objects, to pictured (2-D) as well as real (3-D) objects, and even to illusory objects.

This chapter will cover research and theory on both shape and object perception. The first section of this chapter will cover the processes involved in segmenting the visual field into contours and grouped regions; the second section will cover shape assignment. Object recognition theories will be summarized in the third section, and different visual architectures that specify the relationships among segmentation, shape assignment and recognition will be presented in the fourth section. Finally, the relationship between attention and object perception will be covered in the fifth section.

Before continuing, I should note the ways in which vision scientists use the term *object perception* more narrowly than it is used by philosophers and theorists concerned with defining what constitutes an object. The critical difference is that many of the conceptual or judgmental processes necessary to distinguish real objects from other entities are not included in the term *object perception,* as typically used by perception psychologists. Although visual perception is affected by some types of knowledge embodied in previous experience, it seems immune to influences from other types of knowledge (e.g., Peterson, Harvey, & Weidenbacher, 1991; Peterson, Nadel, Bloom, & Garrett, 1996). Consider, for example, the classic demonstration shown in Figure 6.1a (Hochberg, 1978). Figure 6.1b shows that the number 4 is embedded in Figure 6.1a. However, this familiar shape is not perceived unless the viewer is informed that the display contains the number 4, and unless time sufficient for careful inspection is provided. Göttschaldt (1926) created displays like Figure 6.1a to demonstrate that familiar shape does not affect segmentation, a position that has subsequently been shown to be incorrect (Peterson, 1994a). What Göttschaldt's (1926) demonstrations actually show is that objects cannot be perceived effortlessly unless the critical features defining those objects can be readily extracted from the display. The line terminator features of the number 4 are obscured by the continuous contour in Figure 6.1a (Hochberg, 1971; M. G. Moore, 1930; Woodworth, 1938). On subsequent encounters with Figure 6.1a, the initial percept of a closed loop might be supplemented quickly by knowledge of past experience, which might initiate a search for the number 4. However, the search processes employed under such conditions are secondary to the initial, or primary, perception.

Research designed to distinguish between primary versus secondary perceptual processes and between object knowledge based upon those different processes would certainly be worthwhile, and might be useful in bridging the terminological gaps between philosophers and psychologists, and between investigators of infant and adult perception. It might even allow bridges between the study of object perception and object categorization. Because object perception research is mostly concerned with initial perception, however, the distinction between primary and secondary perceptual processes will not be considered further in this chapter.

Segmentation

This section will summarize research and theory concerning processes by which the visual field is segmented, or differentiated, into contours, regions, and groups. We start with contour segregation because it is fundamental for object perception. Grouping processes and region-detecting processes are considered next.

Contour Perception

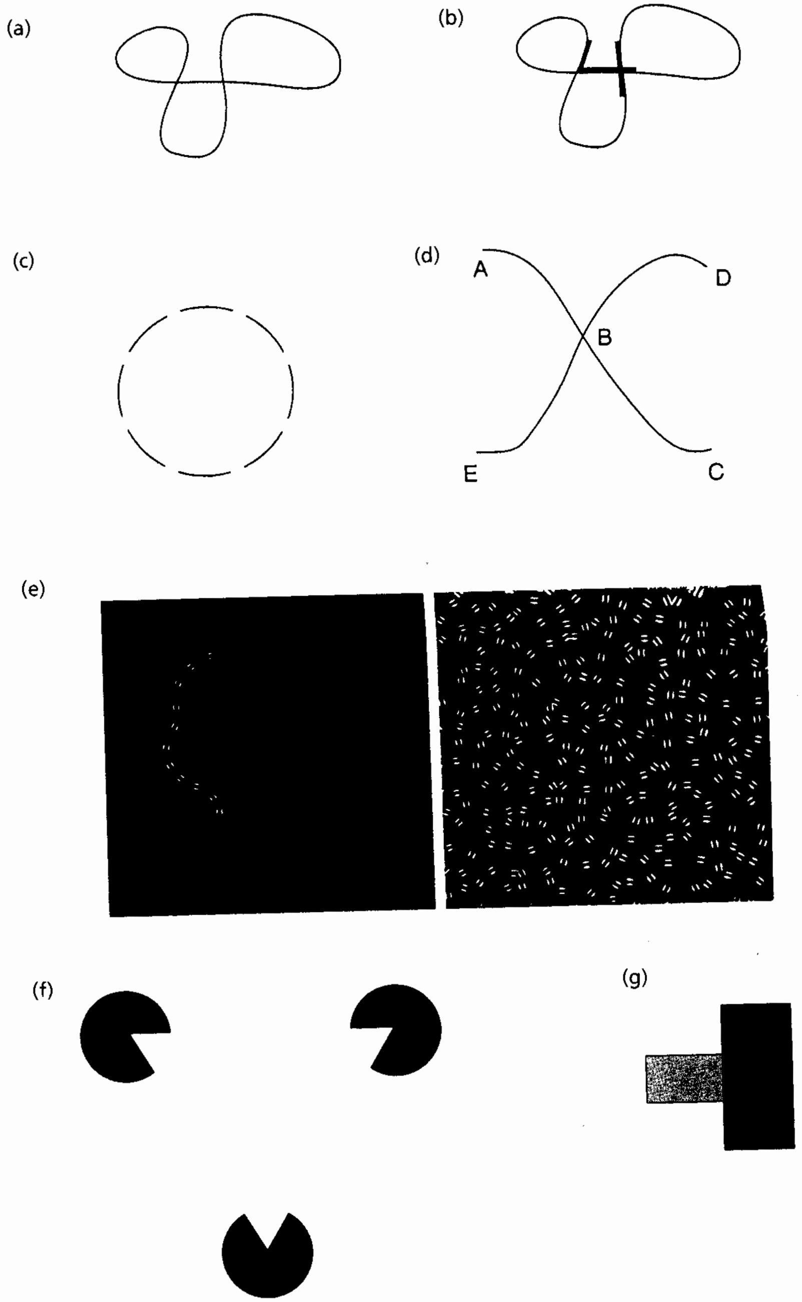
The General Case

Objects are bounded by contours. Although the boundaries of physical objects ar tinuous, the contours extracted by visual processes are likely to be discontinuous, o mented. Therefore, as part of the process of segregating contours from other ele some process must integrate the contour segments (Grossberg & Mingolla, 1985; U 1990). Ullman proposed that contour segregation occurs more readily for more "s contours. According to Ullman, contour salience increases as the orientation sim between neighboring contour segments increases.

Ullman's salience computation implemented the Gestalt psychologists' propos the visual system has an inherent tendency to group segments into contours along entailing the smallest change in curvature. This tendency, called "good continua operates to group both segments of fragmented contours, as in Figure 6.1c, and seg of continuous contours where they intersect other contours, as in Figure 6.1d. The view was that, by virtue of integration by good continuation, fragmented contours (t "virtual contours" by Kanizsa, 1987) were as real as real contours. (For recent beha evidence consistent with this hypothesis, see Rensink & Enns, 1995; Han, Humphr Chen, 1999a).

The Gestalt psychologists supposed that good continuation operates very early course of perceptual organization. After Hubel and Wiesel (1968) showed that cells first layer of visual cortex (V1) were differentially sensitive to stimulus bars of dif orientations, it was thought that V1 might be the neural substrate for contour integ and segregation mechanisms. Recent psychophysical, computational, and neurophysiol work has elucidated contour integration mechanisms, and has confirmed a role f cells. (For an excellent brief history, see Westheimer, 1999.)

Field, Hayes, and Hess (1993) used Gabor patches as both contour elements and ground elements (see Figure 6.1e), and examined the conditions under which the co could be segregated from the background. Field et al. found that, as long as the elen principal axes were misaligned by less than 60°, observers could accurately segregate tours from backgrounds even when (a) the elements differed in phase, and (b) the element distance was up to seven times the element width. (See also Beck, Rosenfe Ivry, 1990.) These effects at a distance demonstrated by Field et al. (1993) were inco ent with the classic understanding of the receptive field properties of V1 cells. Hov Kapadia, Ito, Gilbert, and Westheimer (1995) later showed that V1 cell response stimulus bar were enhanced substantially when a bar with the same orientation was lo nearby, yet outside their receptive field. Strikingly, the patterning of the effects de strated in V1 cells was very similar to the pattern obtained in psychophysical studies Field et al., 1993; Kapadia et al., 1995). The degree to which V1 cell activity was enh by nearby bars decreased as the principal axes of the bars became increasingly misali and as the distance between the two bars increased. Together, these psychophysica physiological results support the hypothesis that V1 cells do indeed play a role in co integration and segregation.



Modal and Amodal Contour Completion

So far, the discussion has focused on real and fragmented, or “virtual” contours. In cc trast, *modal* completion occurs where no explicit contour is present, yet an illusory, subjective, contour is perceived. An example is shown in Figure 6.1f, which appears to a white triangle resting on three black circles. The bounding contour of the white trian is a modal contour, in that it can be perceived. Yet it is a subjective contour becau despite appearances, there are simply no white contours in the display. The black pacm shapes with two straight edges serve as the inducing elements for the subjective conto When the subjective triangle is seen in Figure 6.1f, the black shapes appear to be circ completed behind the subjective triangle, due to *amodal* completion (see below).

Both physiological and behavioral evidence suggest that subjective contours are gen ated by early visual processes. Physiological investigations have identified cells in V1 a V2 that respond to both real and subjective contours shortly after stimulus onset (Gros Shapley, & Hawken, 1993; Peterhans & von der Heydt, 1989; von der Heydt & Peterha 1989). Behavioral evidence indicates that the time required to find subjective contc targets does not increase as the number of locations to be searched increases. Such resu suggest that subjective contours are generated in parallel across the visual field; focal atte tion is unnecessary (Davis & Driver, 1994; Gurnsey, Humphrey, & Kapitan, 1992). addition, psychophysical investigations demonstrate that, similar to real and virtual cc tours, modal contours are perceived between contrast reversed elements (Prazdny, 198 Thus, although the perceived outcomes are very different, there are clear similarities in t early processes that produce real and subjective contours.

The only straight contours present in Figure 6.1f are those of the three black induci elements. Nevertheless, the straight edges are not perceived as belonging to the black shap Rather, the black shapes are completed as circles lying behind the subjective triangle. T is a case of *amodal* contour completion. Amodal contour completion occurs when r lines (or edges) are perceived to connect behind occluding surfaces. This implicit contc completion is considered amodal because a connecting edge is not seen – in contrast modal completion, where contours that are not present in the physical display are nor theless perceived. (For review, see Kanizsa, 1987.) Psychophysical investigations indic: that amodal contours (and the amodal surfaces bounded by those contours) are complet sufficiently early in processing that observers cannot ignore them even when doing would improve their performance on experimental tasks (He & Nakayama, 1992, 199

←

Figure 6.1. (a & b) The number 4, visible in the drawing in (b), is hidden in the drawing in (a). is reprinted from *Perception*, 2/e by Hochberg, J., © 1964. Reprinted by permission of Prenti Hall, Inc., Upper Saddle Riber, NJ. (c) Fragmented contours grouped by good continuation. Intersecting contours grouped by good continuation into continuous segments ABC and EBD. The left and right fields show a sample target and comparison display used by Field et al. (199 reproduced with permission from Elsevier Science. (f) A subjective contour triangle. (g) A g rectangle occluded by a black rectangle. According to Kellman and Shipley's (1991) relatability rı the edges of the gray rectangle do not complete amodally.

Kellman and Shipley (1991) articulated a relatability rule that predicts when amodal completion will occur. The relatability rule states that amodal contour completion will occur only when smoothly curving extensions of interrupted contours meet at an angle less than 90°. Hence, the black inducing elements complete amodally as circles in Figure 6.1f, because the smoothly curving extensions of the outer contours of the inducing elements meet each other. The edges of the gray shape in Figure 6.1g would not complete amodally, however, because the smoothly curving extensions of the inducing element do not meet. The relatability rule captures local constraints on contour connectivity.

I end this section by raising the possibility that both modal and amodal completions are generated by the same processes that integrate real and virtual lines; hence, neither may be special cases after all. Consistent with this possibility, Kellman, Yin, and Shipley (1998) showed that those amodal contours that satisfy the relatability rule have some of the same properties as modal contours. Dresp and Bonnet (1993) showed that the properties of real and modal contours overlap. Moreover, real and subjective contours function similarly as substrates for certain higher-level processes (Peterson & Gibson, 1994b).

Looking Beyond V1 to Explain Contour Segregation, Integration, and Completion

A number of investigators, including Kanizsa (1987), Rock (1987), and Wallach and Slaughter (1988), showed that familiarity affects modal completion. C. Moore and Cavanagh (1998) demonstrated familiarity effects on virtual contour completion. Furthermore, Hochberg and Peterson (1993) and Zemel, Behrmann, Mozer, and Bavelier (under review) demonstrated that familiar shapes are more likely than unfamiliar shapes to be completed amodally. And, Sekuler (1994) showed that in addition to local processes, more global processes, such as the symmetry of the completed figure, play a role in early amodal completion processes. These results suggest that one must look beyond V1 to gain a full understanding of contour integration and segregation processes.

Grouping

Grouping Factors

In addition to good continuation, the Gestalt psychologists identified a number of factors that increase the likelihood that a set of entities will be grouped together and segregated from other entities. For instance, elements that are *similar* are likely to be grouped together. *Similarity* can be determined over any number of dimensions, such as shape, color, or size. An example of grouping by similarity can be seen in Figure 6.2a. In addition, elements that are *close to* one another are likely to group together. The display in Figure 6.2b is likely to be grouped into columns because of the factor of *proximity*. Proximity appears to be determined by the perceived distance separating the elements rather than by the physical distance, when the two differ (Rock & Brosgole, 1964). As well, elements that *move together* are likely to be grouped together. If the elements in columns 1, 3, and 5 of Figure 6.2c were to move upward while the elements in columns 2 and 4 remained stationary, the moving elements would group together by virtue of sharing a *common fate* and would segregate from the stationary elements. Although common fate was traditionally defined for moving versus stationary elements, or for elements moving in opposite tions, Leonards, Singer, and Fahle (1996) recently found that temporal modulati brightness operates to segregate the visual field as well.

Level at Which Grouping Operates

Evidence obtained from a variety of sources suggests that grouping processes are visual processes, as the Gestalt psychologists proposed. Supporting evidence was obt in a task in which observers are asked to categorize a target letter appearing at fixati one of two letters. The target letter is surrounded by a number of distractor letters lyi its right and left sides (B. A. Eriksen & C. W. Eriksen, 1974). Distractors located at a distance from the target are more likely to interfere with the target response when group with the target (by virtue of similarity or common fate) than when they d

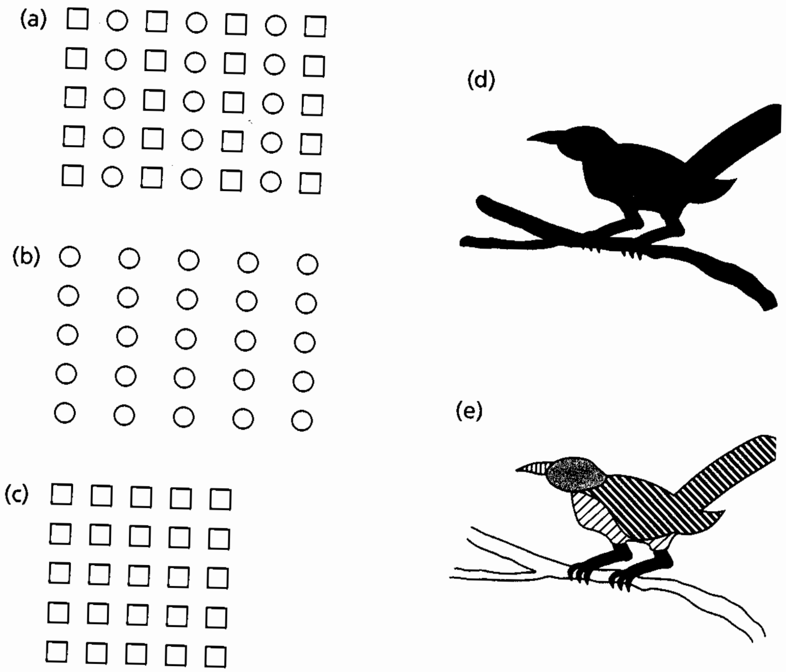


Figure 6.2. (a) An example of grouping by similarity. (b) An example of grouping by proximity Grouping by common fate would occur if the elements in columns 1, 3, and 5 were to move upw while the elements in columns 2 and 4 remained stationary. (d) According to Palmer and R (1994), this display would first be perceived as a unified connected region and later segregated two objects, a bird and a branch. (e) Because of the different textures in this display, it woul treated as up to eight regions at the entry level. Later processes would integrate across the unif connected regions to yield two objects, a bird and a branch.

group with the target (Baylis & Driver, 1992; Driver & Baylis, 1989; Fox, 1998; Harms & Bundeson, 1983; Humphreys, 1981; Kramer & Jacobson, 1991). These results can be taken as evidence that the Gestalt grouping laws are early automatic processes that are not overridden by task-dependent attentional allocation. It remains difficult to pinpoint how early or late the relevant grouping is accomplished, especially because physiological investigations have not shed light on where the grouping factors operate (but see Tononi, Sporns, & Edelman, 1991). Nevertheless, these behavioral results suggest that attention spreads across the entities defined by grouping factors, even when task performance would be improved by focusing attention on the target.

The grouping factors do not all follow the same time course, however. For example, Ben-Av and Sagi (1995) found that proximity grouping is perceived faster than similarity grouping and dominates performance under brief exposure conditions; whereas similarity grouping is perceived somewhat later in rime and dominates performance under long exposure conditions (see also Han et al., 1999a; Han, Humphreys, & Chen, 1999b). Thus, the grouping factors should not be considered a homogeneous set.

Region Formation

Integration and segregation processes are required for regions of homogeneous stimulation as well as for contours and groups of elements (Koffka, 1935). The detection of closed contours might be involved in integrating and segregating homogeneous regions from the outside-in. Homogeneous regions can also be formed from the inside-out by a "region-growing" type of integration processes, analogous to contour integration processes, whereby neighboring image locations are linked by virtue of sharing the same property (e.g., Mumford, Kosslyn, Hillger, & Herrnstein, 1987).

Uniform Connectedness

Recently, Palmer and Rock (1994) outlined a theory of perceptual organization in which regions of homogeneous or uniform visual properties ("uniform connected regions," UCRs) serve as "entry level units" – that is, as the units forming the substrate for other segregation and integration processes. Palmer and Rock (1994) proposed that once UCRs have been isolated in the visual array, subsequent processes can operate either to create divisions within UCRs, as in Figure 6.2d, where a homogeneous black region is seen as two objects – a bird and a branch; or to integrate across UCRs, as in Figure 6.2e, where the regions of different luminance and rexture are integrated into a single object – a bird. According to Palmer and Rock, the principle of "uniform connectedness" (UC) has the privileged position of defining the fundamental units for later segregation and grouping processes.

A Privileged Cue?

Uniform connectedness is surely one of the early integration/segmentation factors employed by the visual system; but the claim that it is the fundamental factor is controversial. Two issues of continued relevance to object perception underlie the debate. A first issue is

whether the fundamental units for object perception are global, bounded regions, or wheth they are smaller units (see Boselie, 1994; Boselie & Leeuwenberg, 1986; Hochberg, 196 1980; Kimchi, 1998; Peterson & Hochberg, 1983, 1989). A second issue is whether an one factor constitutes the fundamental, or dominant, segmentation factor, or whether U and the Gestalt grouping and configural factors constitute a subset of a larger set of facto that cooperate to organize the visual field (Peterson, 1994b, 1999).

Consistent with Peterson's view that UC operates as one cue among many, Han et a (1999) found that grouping by a cue known to operate quickly – proximity – was accom plished as fast as grouping by UC and was not enhanced when combined with UC. How ever, they found that grouping by a cue known to operate more slowly – similarity – wa accomplished more slowly than grouping by UC and was enhanced when combined wit UC. Furthermore, developmental research suggests that UC is not a dominant factor i infants' organization of the visual world (Spelke, 1988). However, consistent with Palme and Rock's view that UC defines the entry level elements for perception, Watson and Krame (1999) found that, in adults, other things being equal, attention may select regions define by UC, even when the selection of larger units would speed task performance (Watson & Kramer, 1999). Additional research is required to determine whether UC has the privilege position of defining the first fundamental units for perceptual organization or whether it i simply one of many cues, each of which has different strengths and time courses.

Shape Assignment

The integration and segregation of contours, groups, and regions is not sufficient for shap perception because not all regions in the visual field are perceived to have shape; some ar perceived as shapeless backgrounds. Contours can be described as shared by two regions one lying on each side. Whenever two regions share a contour, two perceptual outcome are possible. One outcome is that the contour is assigned to one region only; whereas th adjacent region is left contour-less. In this case, the region to which the contour is assigne is the "figure"; the adjacent region is the "ground." By virtue of contour ownership, th figure appears to have a definite shape, whereas the adjacent ground does not, at least nea the contour it shares with the figure. When this outcome, termed *figure-ground segregation* is perceived the shared contour is seen as an occluding contour, in that it appears to occlude parts of the ground (i.e., the ground appears to continue behind the figure). A example is shown in Figure 6.3a.

A second outcome that can be perceived when two adjacent regions share a contour is tha the shared contour can be assigned to both regions rather than to just one region (Kennedy 1973, 1974). When this outcome, called *figure-figure segregation*, is perceived, the share contour signifies the meeting of two surfaces or objects, both of which appear to be shape by the contour. The two surfaces can appear to lie on the same depth plane, as in a til pattern (Figure 6.3b), or to slant in depth, as in the two surfaces of a cube that meet at common edge (Figure 6.3c). Examples such as Figures 6.3b & 6.3c demonstrate that one-sided contour assignment is not "obligatory," as some have claimed (Baylis & Driver, 1995)

In some situations, such as the one depicted in Figure 6.3c, figure-figure segregation i

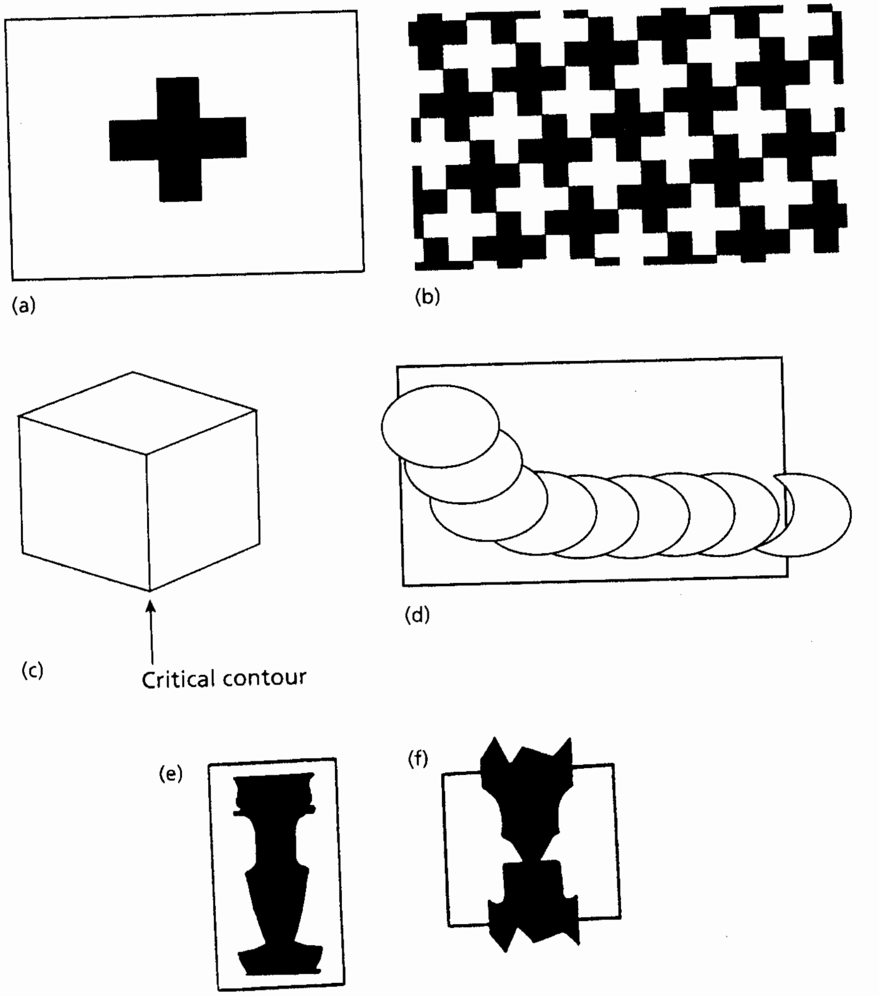


Figure 6.3. (a) An example of figure-ground organization in which the contour shared by the black and white regions bounded by the rectangle is assigned to the black region. The black region is seen as the shaped figure (a cross), whereas the white region appears to be a shapeless ground, continuing behind the cross. (b) A tile pattern in which the contours shared by the black and white regions are assigned to both regions. Both regions appear to be shaped and to lie on the same depth plane. (c) The critical contour signifies the meeting of two faces of a cube. (d) Eight ovals and one crescent. The contours in this display are assigned to one side only. All regions, except the outer two, appear to be figures along one portion of their bounding contour and grounds along another portion of their bounding contours. (From Hochberg (1980), copyright © 1980 by Academic Press, reproduced by permission of the publisher.) (e, f) Displays used by Peterson et al. (1991). The white high-denotative surrounds are more likely to be seen as figures when the displays are upright rather than inverted. (Reprinted with permission from the American Psychological Association.)

clearly the preferred organization. That may be because the Y- and arrow junctions ir figure are themselves early cues to 3-D structure (Enns & Rensink, 1991; Humm Biederman, 1992). In other situations, such as the one depicted in Figure 6.3a, how figure-ground segregation seems to be preferred. The likelihood of seeing figure-figure sus figure-ground segregation can be attributed to the cross-region balance of (a) config factors identified by the Gestalt psychologists and others, (b) contour recognition proce and (c) monocular and binocular depth cues. These factors are discussed next.

Gestalt Configural Factors

The Gestalt psychologists elucidated a number of factors that affect the likelihood th region will be attributed figural status while its adjacent region will be attributed gro status; these factors are called the Gestalt configural factors. Regions that are (a) small area than their surrounds, (b) symmetric (especially around a vertical axis), (c) con and/or (d) enclosed are likely to be seen as figures; whereas their adjacent regions are li to be seen as grounds. Demonstrations devised by Gestalt psychologists in the first ha the twentieth century supported these claims (for reviews, see Hochberg, 1971; Pomer & Kubovy, 1986).

Recent research replicated and extended the demonstrations of the Gestalt psych gists. For instance, Kanizsa & Gerbino (1976) tested the importance of global conve and found that it is a stronger cue to figural status than symmetry. The importanc closure as a Gestalt configural cue was recently confirmed by Kovács & Julesz (1994) used a detection task rather than the phenomenological reports favored by the Ge psychologists. Kovács and Julesz (1994) obtained lower detection thresholds for tar presented near the center of a region bounded by a closed curve than for targets preser at an equivalent distance from the contour outside the bounded region. Given that cl regions tend to be seen as figures, these results replicate and extend research conducte Wong and Weisstein (1983) who reported that detection of high spatial frequency tar is superior when targets fall on the figure rather than the ground. Similarly, using a c tour-matching paradigm as an indirect measure of perceived organization, to avoid sc of the demand character of phenomenological report, Driver, Baylis and Rafal (19 recently confirmed the importance of smallness of relative area as a segregation cue.

Modern research has revealed new factors that can be added to the list of config cues. Brown and Weisstein (1988) showed that when different spatial frequency patte cover two adjacent regions, the region covered with the higher spatial frequency is likel be seen as the figure. O'Shea, Blackburn and Ono (1994) showed that the region t contrasts more with the background is likely to be seen as the figure. And, Hoffman Singh (1997) showed that regions with distinctive parts (defined as large area, con excursions into adjacent regions) are more likely to be seen as figures than grounds.

Does "Configural" Imply "Global"?

Although configural cues are often considered global, holistic cues, recent research in cates that configural factors can be computed locally. For instance, Hoffman and Sing

(1997) part distinctiveness is measured locally. Further, Stevens and Brooks (1988) showed that convexity can operate locally. In addition, Han et al. (1999b) and Kimchi (1994) have shown that configural factors are not necessarily mediated by global, object-wide mechanisms. These findings are important because they are consistent with the evidence indicating that figural status need not be assigned to an entire bounded region. A region can be figure along one portion of its contour and ground along another portion (Hochberg, 1980, 1998; Hoffman & Singh, 1997; Peterson & Hector, 1996), as illustrated in Figure 6.3d.

Level at Which Configural Cues Operate

Recent research confirms the Gestalt claim that configural factors are computed early in processing. Peterson and Gibson (1994a) showed that symmetry can determine figure assignment in masked exposures as short as 28 ms (but not in 14-ms masked exposures). An approach used recently to partition perception into early and late-acting processes has been to test brain-damaged individuals who distribute attention preferentially toward the side of space or the side of an object contralateral to the brain damage (contralesional spaces or contralesional sides of objects; Heilman & Van Den, 1980; Kinsbourne, 1970; see Chapter 7 for more details.) Processing that occurs in unattended contralesional space can be considered "preattentive" – that is, can be considered to occur before the intentional allocation of attention. Driver et al. (1992) showed that the configural cue of smallness of relative area operates effectively to determine figure-ground segregation in contralesional space. Similarly, Driver et al. (1992) found that the configural cue of symmetry affected figure-ground segregation normally in a patient who was unable to consciously attend to the contralesional sides of the figures he saw, and hence, unable to judge accurately whether or not the figures he perceived were symmetric.

None of the factors described above determined which of two adjacent regions will appear to be shaped (i.e., will be seen as the figure) 100% of the time, especially when other competing configural cues are present. Furthermore, the likelihood of assigning shape to one region or the other is affected by contour recognition cues and depth cues, as well as by configural cues. The perceived segregation depends upon the balance of cues across regions competing for figural status (Peterson, 1994a, 1999).

Quick Access to Memories of Object Structure

It was traditionally assumed that access to memories of objects occurs only after grouping and segregation processes have produced the figures or objects in the visual array (e.g., Köhler, 1929; Neisser, 1967; Biederman, 1987). Contrary to this assumption, Peterson and her colleagues found evidence that object memories activated by contours can serve as one more shape-assignment cue. Their results were obtained using stimuli like those shown in Figures 6.3e and 6.3f in which adjacent regions sharing a contour differed in the degree to which they resembled known objects. One, "high denotative," region was a good depiction of an upright known object when it was seen as figure (e.g., the white regions in Figures 6.3e and 6.3f), whereas the other, "low denotative," region was not (i.e., the black regions in Figures 6.3e and 6.3f). (The denotivity of each region was determined by between-subjects agree-

ment in a pre-test in which observers listed all the known objects each region res
when it was seen as figure.) Other relevant cues, such as configural cues and the mo
and binocular depth cues, were sometimes present in these displays, and when th
present, these cues favored the interpretation that the low-denotative region was the
figure (e.g., Peterson, Harvey & Weidenbacher, 1991; Peterson & Gibson, 1993, 1

Peterson et al. (1991) compared the likelihood that the high-denotative region w
as the shaped figure when the displays were upright (as shown in Figures 6.3e an
versus inverted (i.e., as seen when you turn the book upside down). Such rotation
picture plane does not change the configural or depth cues present in the displays sh
Figures 6.3e and 6.3f. However, rotation in the picture plane does slow down ac
object memories (Jolicoeur, 1985, 1988; Tarr & Pinker, 1989). Peterson et al.
reasoned that the delay induced by inversion in the picture plane might be suffic
remove, or to reduce, any influence from object memories that normally affects the
gation outcome. Therefore, Peterson and her colleagues argued that any increase
dency to see high-denotative regions as figures in upright compared to inverted d
would constitute evidence that object memories, activated early in the course of perc
processing, can affect the segregation outcome. Consistent with this prediction, Pe
et al. (1991; Gibson & Peterson, 1994; Peterson & Gibson, 1994a, 1994b) foun
high-denotative regions were more likely to be seen as figures when the displays we
right than inverted. They attributed these effects to a set of quick recognition pr
operating simultaneously on both sides of the contour shared by two adjacent regi

The objects portrayed by inverted high-denotative regions are not necessarily unrec
able; they can be recognized once they are seen as figure. Nevertheless, influences from
memories on shape assignment are diminished or absent for inverted displays. This
indicates that only those object memories that are accessed quickly affect shape assign
object memories accessed later in time (as when misoriented displays are used) do not
ence shape assignment. Consistent with this conclusion, neither priming nor knowle
what the inverted high-denotative region portrays alters the orientation effects (Gib
Peterson, 1994; Peterson et al., 1991). Effects of object memories on shape assignme
evident only when a match between the stimulus and a memory representation codi
structure of the object can be made quickly. Additional evidence implicating access to
ries of object structure is that no effects of object memories on shape assignment are ob
when the parts of the known object portrayed by the high-denotative region are retaine
their spatial interrelationships are rearranged, or scrambled. Furthermore, effects of
memories on shape assignment have been found only when the contours that serve
substrate for access to object memories are detected early in processing. For example,
nance contours and subjective contours support contour recognition effects, whereas bi
lar disparity contours, available later in processing, do not (Peterson & Gibson, 1993, 1

Neuropsychological investigations are consistent with the proposal that the pro
subserving quick access to object memories should be considered early visual processe
instance, a visual agnosic individual whose object identification was severely impaire
ertheless showed normal influences from contour recognition processes on figure-gr
responses (Peterson, de Gelder, Rapcsak, Gerhardstein, & Bachoud-Lévi, 2000).
results indicate that contour recognition processes operate outside of conscious awar
and are a subset of the processes required for conscious object recognition/identific

In addition, Peterson, Gerhardstein, Mennemeier, and Rapcsak (1998) tested individuals with unilateral brain damage whose attenrion was biased away from the contralesional contours of the regions of the experimental displays. Nevertheless, contour recognition processes seemed to operate normally on the unattended contralesional contours, suggesting that contour recognition processes proceed without the benefit of focused attention.

Physiological evidence is consistent with the claim that shape assignment is accomplished early in visual processing. Zipser, Lamme, and Schiller (1996) measured a response in V2 cells 80–100 ms after stimulus onset that was evident when near, shaped, figures, but not grounds, fell on the cells' receptive fields. If this differential activation indeed indicates that shape assignment or figure-ground segregation has been accomplished, these data confirm the view that those processes occur early in visual processing. Because the inferior temporal cortex, located downstream from V2, and important for object recognition, can be activated 60 ms after stimulus onset, however, these data are also consistent with the proposal that object memories can affect figure-ground segregation.

Just as for the other segregation-relevant cues, the likelihood that the region providing a good fit to object memories will be seen as figure depends on the balance of other cues. In other words, the cue originating in quick access to object memories does not always dominate the configural and depth cues (Peterson & Gibson, 1993, 1994a). Indeed, just as none of the other configural cues or depth cues is a necessary component of the segregation process, neither is a good fit to an object memory. Therefore, segregation can proceed without substantial contributions from object memories for novel objects, as it can proceed without contributions from the configural cue of symmetry for asymmetric objects (or without convexity for concave objects). However, when known objects are present, the segregation process can benefit from prior experience, as it can benefit from convexity when convex objects are present.

Depth Cues

Many depth cues, including binocular disparity (stereo), contour, motion parallax, texture, and shading, affect the likelihood that shape will be assigned to the region lying on one or the other side of a contour. Some of these cues, such as shape from shading, are determined early in processing (Braun, 1993; Ramachandran, 1988; Sun & Perona, 1997), whereas others, such as shape-from-stereo, may unfold over a longer time course (Julesz, 1971; Sun & Perona, 1997).

Cue combination studies address the question of whether these cues to 3-D shape interact early in processing or whether they are computed independently until each pathway produces an estimate of depth. Some evidence indicates that the 3-D cues combine linearly to determine shape, a finding that is consistent with the latter possibility. But departures from linear combination have also been obtained, and these departures are consistent with the possibility that the cues to 3-D shape interact early in processing (Bülthoff, 1991; Parker, Cumming, Johnston, & Hurlburt, 1995). Neuropsychological case studies of patients who are impaired at seeing shape from shading, but relatively intact at seeing shape from edge cues (and vice versa) provide evidence consistent with the hypothesis of separate pathways (Battelli, Casco, & Sartori, 1997; Humphrey, Symona, Herbert, & Goodale, 1996), but do not necessarily speak to the independence of those pathways.

Investigations of how depth cues interact with configural cues and activated object mer ries are, unfortunately, rare. Hence, it is not possible at this point in time to draw conclusions about how these different cues to shape are combined. In conducting resea to address this question, it will be important to bear in mind a consideration recently rai by Tittle, Norman, Perotti, and Phillips (1997) that different depth cues may be imp tant for different properties of 3-D structure. For instance, Tittle et al. (1997) showed t binocular disparity is relatively more important for the perception of scale-independ aspects of shape (e.g., whether a shape is spherical versus cylindrical) than for scale-pendent aspects of shape (e.g., magnitude of surface curvature).

As for the configural cues, none of the depth cues is 100% predictive of perceived de especially when other, contradictory depth cues are present. Instead, it seems that ceived depth corresponds to the depth signaled by the ensemble of cues in any partic scene (Landy, Maloney, & Young, 1990), although different depth cues may have dif ent strengths (Cutting & Vishton, 1995), as different configural cues do.

A review the object perception processes must address the question of what object me ries are like, and how object recognition occurs. Accordingly, a brief review of theorie object recognition is given next.

Theories of Object Recognition

An adequate theory of object recognition must account for

- the accuracy of object recognition over changes in object size, location, and orienta (and it would be preferable if this account did not posit a different memory record each view of every object ever seen);
- the means by which the spatial relationships between the parts or features of an ob are represented; and
- the attributes of both basic-level and subordinate-level recognition (e.g., recogni of a finch as both a bird and as a specific kind of bird).

Current competing object recognition theories differ in their approach to each of t attributes (see Biederman, 1987, 1995; Tarr, 1995; Tarr & Bülthoff, 1998).

Recognition by Components Theory

According to the Recognition by Components (RBC) theory, proposed by Biederman (1 1987), objects are parsed into parts at concave portions of their bounding contours, and parts are represented in memory by a set of abstract 3-D components, called "geons." Be RBC was proposed, other theorists had stressed the importance for recognition of both cave regions of the bounding contour of objects (e.g., Hoffman & Richards, 1985; M 1982; Marr & Nishihara, 1978) and 3-D representational components (i.e., cylinders, Bin

1981; Marr, 1982; Marr & Nishihara, 1978). Biederman expanded the set of components from cylinders to generalized cones (i.e., cross-sections swept out in depth along an axis). Further, Biederman (1995) showed that a finite set of 3-D geons (N = 24) can be defined by combining a small set of binary or trinary contrasts that can easily be extracted from two-dimensional images. Thus, in RBC, a representation of an object's 3-D structure was derived from contrasts extracted from a single 2-D view. The contrasts specify the shape of the cross-section of a geon, the shape of the axis of the geon, and changes in the size of the cross-section as it is swept along the axis. (Contrasts include the following: for edges, whether they are straight or curved, parallel or non-parallel, converging or diverging; and for cross-sections, whether they shrink, expand, or remain constant in size as they move along the geon axis.) Sample geons and some objects constructed from them are shown in Figure 6.4a.

The contrasts from which the geons are constructed are viewpoint invariant properties or "non-accidental properties," in that they are unlikely to occur in the image as an accident of viewing position (Lowe, 1985, 1987; Witkin & Tenenbaum, 1983). For instance, under most viewing conditions, except for accidental views, curved lines do not look straight, nor do converging lines appear parallel. Biederman and his colleagues (Biederman, 1987, 1995; Biederman & Gerhardstein, 1993, 1995) argued that geon extraction is viewpoint invariant because the contrasts that specify the geons are viewpoint invariant. The prediction that object recognition should be viewpoint invariant followed, provided that the same geons (and geon relations, see below) could be extracted from the image in different views. Thus, according to RBC theory, only a small number of views of each object need to be represented in memory.

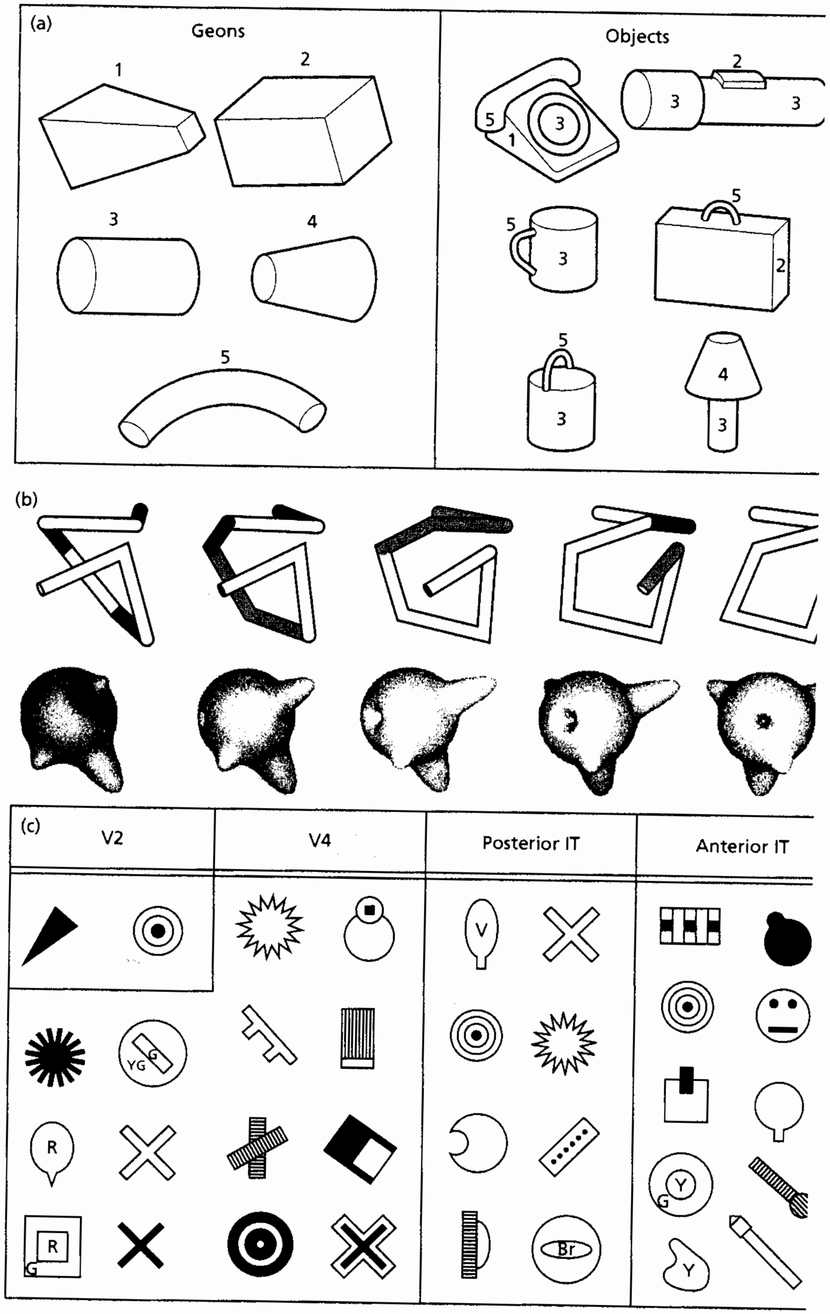
RBC specified that the spatial relations between the geons comprising an object are specified in terms of categorical relationships such as "top-of," below, or "next-to," rather than in metric terms (Biederman, 1987; Hummel & Biederman, 1992). It is known that object recognition fails when the parts are rearranged (Cave & Kosslyn, 1993; Peterson et al., 1991). Nevertheless, prior to RBC, little consideration had been given to the question of how the spatial relationships between the parts of an object are coded.

Evidence: Pro and Con

As all good theories should be, RBC makes testable predictions and consequently, is falsifiable. After the publication of Biedermans' (1987) article, research and theorizing on object recognition flourished, and continues to flourish today. Research investigated questions such as whether or not (a) bounding contours and concave cusps are as important as claimed by RBC, (b) object recognition is viewpoint invariant, (c) RBC can account for both subordinate and basic level recognition, and (d) RBC's proposals concerning the coding of spatial relationships. The fact that many objects can be recognized from their bounding contour alone indicates that bounding contours are highly important for object

→

Figure 6.4. (a) Sample geons and objects constructed from them. Reprinted from Biederman (1995), with permission of MIT Press. (b) Paperclip objects (top row) and spheroid objects (bottom row) used in tests of Multiple Views Theory (reprinted from Logothetis et al. (1994), with permission from Elsevier Science). (c) Examples of the shapes that activated cells at various levels in the ventral processing stream.



recognition, as specified by RBC (Hayward, 1998; Peterson, 1994a), although bounding contours are clearly not the whole story (Riddoch & Humphreys, 1987). Consistent with part-based theories such as RBC, evidence suggests that the concave portions of bounding contours are more important than other contour segments (Baylis & Driver, 1994; Biederman, 1987; Hoffman & Singh, 1997; Hoffman & Richards, 1985; Braunstein, Hoffman, & Saidpour, 1989). Furthermore, recent research by Saiki and Hummel (1998) indicates that the spatial relationships between parts of objects are represented differently than spatial relationships between different objects. Although this last finding does not directly support RBC, it does suggest that a complete theory of object recognition must account for the coding of spatial relationships between object parts.

Overall, the research suggests that certain elements of RBC theory must be retained, but other elements should probably be abandoned. In particular, evidence that neither geon extraction nor object recognition is viewpoint invariant (Brown, Weisstein & May, 1992; Tarr & Pinker, 1989) led to the formulation of a competing theory, discussed next.

Multiple Views: Evidence and Theory

Psychophysical evidence suggests that object recognition is not viewpoint invariant as proposed by RBC (e.g., Bülthoff & Edelman, 1993; Tarr & Pinker, 1989). Furthermore, physiological evidence indicates that cells may code for individual views of objects (Logothetis, Pauls, & Poggio, 1995) and faces (Perrett et al., 1985). Accordingly, Bülthoff, Edelman, Tarr and their colleagues (Bülthoff & Edelman, 1993; Edelman & Weinshall, 1991; Tarr, 1995; Tarr & Bülthoff, 1995) adopted a different theoretical approach to object recognition, proposing that multiple two-dimensional views of objects are represented rather than just a few 3-D views. According to Multiple Views Theory, object recognition is view-dependent (rather then view-independent) in that objects seen in new views must undergo some time-consuming process before they are matched to stored views and recognized.

In addition, Tarr and Bülthoff (1995) argue that the geon-based representations of RBC theory fail to account for either basic or subordinate level recognition (see also Kurbat, 1994). They criticize the RBC geons for their coarseness, arguing that geon-based representations could not distinguish between members of different basic level categories such as a horse and a cow, and could not represent the differences between two horses, two cows, or two dogs. Yet humans can easily make these sorts of distinctions. In contrast to RBC, exemplar representations, such as those in Multiple Views Theory, can readily represent the differences between objects by representing their different salient features. Similarities between objects can be made explicit through multidimensional feature interpolation (Poggio & Edelman, 1990).

Criticisms of Multiple Views Theory

Multiple Views Theory has not yet specified the exact form of the representation used for common objects. Much of the research supporting RBC has been conducted with open "paperclip" or spheroid objects such as those in Figure 6.4b, where all of the parts were identical save for length, and the bends in the paperclips were the salient features used for recognition. But such objects may not be representative of the objects humans recognize. Another criticism is that the object representations in Multiple Views Theory are too much like two-dimensional templates. It is well known that template-like representations leave perception susceptible to disruption by slight changes in any object features (Neisser, 1967), regardless of whether they lie on, or internal to, the object's bounding contour. Yet human perception is notoriously robust to such changes. It was just this robustness of object perception that led early theorists to propose that object memories were view-independent, size-independent, location-independent, etc. (i.e., in Marr's (1982) terminology, they were "object-centered" representations). It is feared that Multiple Views Theory might require an unreasonably large number of representations for each object. Moreover, it is not clear how different views of objects are determined to be the same object, rather than similar but different objects. Another criticism is that in Multiple Views Theory there is no provision, save for that implied in template matching, for representing the spatial relations between parts of objects. This is a drawback, given the importance of the spatial relations between features, and the behavioral distinctions between the spatial relations between objects per se, and between parts of objects (see above).

These criticisms point to research that must be done to elaborate Multiple Views Theory. Current attempts to resolve these problems include (a) using the bounding contour of the object to integrate different views, and (b) exploring the feasibility of categorical coding for spatial relations between features (for a summary of recent research see Tarr and Bülthoff, 1998). Recall that both bounding contours and categorically coded spatial relations play important roles in RBC. It may turn out that a complete theory of object recognition must incorporate principles of both RBC and Multiple Views Theory (Suzuki, Peterson, Moscovitch, & Behrmann, under review; Tarr & Bülthoff, 1998). There also remains the possibility that, in addition to orientation-dependent representations (such as those identified by Logothetis et al., 1995), there exist object-centered representations (i.e., representations that permit orientation-independent object recognition) (Corballis, 1988; Solms, Turnbull, Kaplan-Solms, & Miller, P., 1998; Turnbull & Mccarthy, 1996).

Open Issues

In addition to the issues discussed at the end of the preceding section, two other issues must be considered in order to understand object recognition. The first concerns the role of local features in object memories. For the most part, theorists assume that object recognition is a global or holistic process. However, both behavioral and computational evidence (Mozer, Zemel, Behrmann, & Williams, 1992; Peterson & Hector, 1996; Ullman, 1998) suggests that object recognition can be mediated by local cues. Those local cues that are necessary and sufficient for object recognition have yet to be determined. Furthermore, mounting evidence suggests that the local components of representation are affected by experience (Lin & Murphy, 1997; Mozer et al., 1992; Schyns, Goldstone, & Thilbaut, 1998; Zemel et al., under review). Future research exploring the nature of the local cues mediating object recognition, the degree to which they are learned, and the interactions between local and global cues will be important for object recognition theory.

The second open issue concerns the nature of the representational primitives. In both Multiple Views Theory and RBC Theory, there is a clear resemblance between the object and the representational components. Indeed, it is easier to think about the components of object representations as being similar to the nameable or visible parts of objects than it is to think about them as abstractions bearing little or no resemblance to the consciously perceived object parts. However, alternative conceptions exist. One possibility is that objects are represented by their Fourier components (e.g., Graham, 1989, 1992). Another possibility is that objects are represented by complex shape components such as those suggested by research by Tanaka (1993; Kobatake & Tanaka, 1994) (see Figure 6.4c). Tanaka and his colleagues discovered that cells in monkey temporal cortex are selective for complex shape components, many of which bear little resemblance to either whole objects or parts of objects. Furthermore, these investigators uncovered a columnar organization in the temporal cortex, where cells within a column share a similar selectivity. Many questions about these components await further investigation: Can ensembles of these components be used to represent the entire set of objects the monkeys can recognize? Is the selectiviry changed by experience? Are some components best described as coding global features and others as coding local features? The answers to these questions will constrain future theories of object recognition.

Models of the Relationship Between Segmentation, Shape Assignment, and Object Recognition

In order to understand how object perception proceeds, it is important to understand how the component processes of integration and segmentation, shape assignment, and object recognition are ordered. Which precede the others? Which serve as substrates for others? In what follows, I first discuss traditional hierarchical models. Next, I summarize a parallel model my colleagues and I have proposed. Finally, I point out the open questions that must be addressed to adjudicate between these models.

Hierarchical Models

The Gestalt psychologists proposed that segmentation, shape assignment and recognition were ordered serially and hierarchically, with grouping and segmentation completed first and forming the substrate for shape assignment, and shaped regions in turn providing the substrate for, and necessarily being determined prior to, access to object memories. (For some evidence consistent with the proposal that segmentation is completed before shape assignment see Sekuler & Palmer, 1992; for contradictory evidence see Bruno, Bertamini, & Domini, 1997; Peterson et al., 1991; Kellman et al., 1998.)

An influential model of vision proposed by David Marr (1982) was also serial and hierarchical. Unlike the Gestalt psychologists, Marr concentrated on the traditional depth cues at the expense of the configural cues, arguing that the sphere of influence of the latter was restricted to 2-D displays, which represent only a small subset of the conditions under which the visual system operates. (The current isolationism between those who study the perception of shape based upon configural cues versus depth cues can be traced to Marr's position.)

According to Marr, visual input proceeds through a number of stages, illustrated in Figure 6.5a. The first stage of processing is the primal sketch, in which edges are ma explicit. The second stage entails the construction of the 2½-D sketch, in which surfac and viewer-relative orientations emerge. The third stage is the construction of the 3- model, and as a final step, the 3-D model is matched to 3-D object models stored memory. In Marr's theory, there is a clear sequence from edge extraction through 3- shape assignment before object memories are accessed. (Marr's theory was proposed befo either the RBC Theory or the Multiple Views Theory of object recognition. Indeed, t RBC Theory owes much to Marr and Nishihara's (1978) work.)

More recent interactive hierarchical models of the relationship among segmentatio shape assignment, and object recognition allow feedback from higher levels to influen processing at lower levels. However, these models maintain a hierarchical structure in th lower-level processes must at least be initiated before higher-level processes are initiated, illustrated in Figure 6.5b (McClelland, 1979, 1985; McCelland & Rumelhart, 198 Rumelhart & McClelland, 1982, 1986; Vecera & O'Reilly, 1998).

In hierarchical views of perceptual organization, configural and depth cues are consi ered lower-level, or bottom-up, cues – cues that do not require access to higher-level memo representations, and shape assignment based upon these cues is considered a lower-le process than object recognition. Consequently, according to these accounts, object mem ries cannot be accessed before shape assignment and perception is at least partially accor plished. On the basis of the evidence that contour recognition processes influence sha assignment, my colleagues and I proposed a parallel model, discussed next.

A Parallel Model

Recall that investigations with figure-ground displays indicated that object memories access quickly in the course of processing affect the shape assignment. The cues arising from the activated object memories did not dominate the other configural cues or depth cues; n did the configural and depth cues constrain access to object memories. Rather, activat object memories seem to serve as one more cue among the many cues that contribute the likelihood that a region will be seen as a shaped figure rather than a shapeless groun Critically, object memories affected shape assignment only when they were accessed quickl Influence from object memories could be removed either by inverting the stimuli (ar thereby delaying the access to object memories), or by using contours detected later rath than earlier in processing (e.g., random-dot stereo edges versus luminance edges).

On the basis of this evidence, my colleagues and I proposed that, as soon as contours a segmented in the visual input, quick access to object memories via contours is initiate The model is a parallel model because object memories are accessed via contour-bas mechanisms at the same time that other processes assess the Gestalt configural cues and t depth cues, and all of these processes interact to affect shape assignment (see Figure 6.5 We do not suppose that time course of all of these processes is the same. We suppose on that shape assignment based upon configural cues and/or depth cues does not prece access to object memories, either partially or wholly, as would be assumed on hierarchic models (Peterson, 1994a, 1999; Peterson & Gibson, 1994a, 1994b).

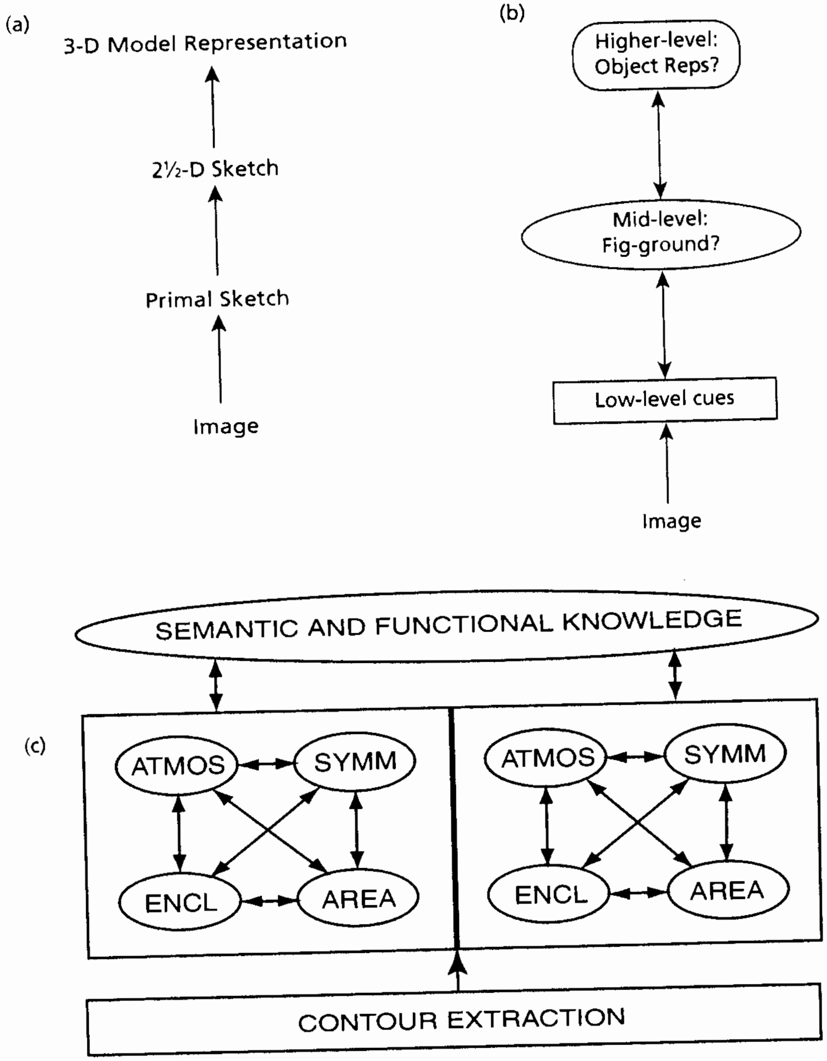


Figure 6.5. (a) A sketch of Marr's serial hierarchical theory. (b) An interactive hierarchy. (c) The parallel interactive model proposed by Peterson et al. (2000). A selection of shape processes are shown operating on both sides of a contour extracted early in processing, including ATMOS (Access to Memories of Object Structure), SYMM (Symmetry), ENCL (Enclosure), and AREA. Facilitatory connections exist between shape processes operating on the same side of the contour (indicated by double-headed arrows). Inhibitory connections exist between shape processes operating on opposite sides of the contour (indicated by T-endings). Feed-forward and feedback connections to and from Semantic and Functional Knowledge are also indicated by double-headed arrows. (Originally published in Peterson et al. (2000, Figure 9); reprinted with permission from Pergamon Press.)

A Continuing Debate

Hierarchical interactive models have been adapted to account for the evidence indicatin that shape assignment is affected by access to object memories (e.g., Vecera & O'Reill 1998, 2000). In Vecera and O'Reilly's (2000) model, lower-level processes do not constrain the operation of higher-level recognition processes and effects of recognition processes on the processing at lower levels are evident at the earliest time slices. It will b difficult to distinguish this version of an interactive hierarchical model from a paralle model (e.g., Peterson, 1999; Peterson et al., 2000).

Many theorists prefer hierarchical models (serial or interactive) to parallel models because they believe that hierarchical models are better able than parallel models to accoun for the perception of novel or unrecognized objects (e.g., Marr, 1982; Warrington, 1982) However, this belief is based on the incorrect assumptions that (a) in parallel accounts inputs from higher-level object memories are necessary for shape assignment and perception, and (b) high-level influences must dominate low-level factors. Neither of these assumptions is necessary and neither is held in the parallel models proposed by Peterso (1999; Peterson et al., 2000; also see above).

A hierarchical model implicitly underlies the notion of object files proposed by Kahnema and Treisman (1984) to account for perceived object continuity over changes in perceive identity, shape, color, or location. They proposed that temporary representations of objects are created at an intermediate hierarchical level, before object identity is established These "object files" mediate object continuity over changes in object features such as location, color, and shape, provided that the changes are not too extreme. Priming experiments support the existence of object files (e.g., Kahneman, Treisman, & Gibbs, 1992 Treisman, Kahneman, & Burkell, 1983), but experimental evidence suggests that objec files may code some aspects of object identity as well (Gordon & Irwin, 1996; Henderso & Anes, 1994). Thus, object files must be understood within a hierarchical model only i one assumes that spatio-temporal continuity is less likely to be maintained over changes i object identity than over changes in other object features. Although this prediction migh be generated by a serial hierarchical model, it is not necessary on a parallel account.

Another reason underlying a preference for both serial and interactive hierarchical approaches to perceptual organization is that brain structures are thought to be arranged hierarchically. For instance, occipital cortex is initially activated via cortical connections earlier in time than temporal and parietal cortices, which in turn, are initially activated before frontal cortex. (Indeed, Vecera and O'Reilly (2000) argue that their hierarchical interactive model of figure-ground assignment is an architectural model rather than a processing model.) It is tempting to associate functional stages such as those proposed by Mar (1982) with these sequentially activated brain regions. However, it must be remembered that there are massive feedback connections between brain regions as well as feed-forward connections (Felleman & Van Essen, 1991; Zeki, 1993). These feedback connections from brain regions activated later in time via cortical connections can alter the activity in brain regions activated earlier in time. When these feedback connections are taken into consideration, it becomes very difficult to pinpoint the stage at which various aspects of perceptual organization are accomplished (Braddick, 1992; Peterson, 1999).

Consider the V2 cells that respond differentially to figures than to grounds (Zipser et al., 1996). This "figure" response in V2 occurs after some cells in temporal cortex have responded, so feedback may be involved (Zipser et al., 1996). This is not to say that activity did not occur in V2 prior to the measured "figure" response. But it is simply not clear whether or not the prior activity should be taken as constituting emergent figure-ground segregation, as might be predicted on any interactive hierarchical view. Alternatively, the prior activity represents some other perceptual function, such as border detection or grouping (see Zipser et al., 1996), neither of which can properly be said to produce emergent figure-ground segregation, as discussed previously.

At this time, we simply cannot distinguish between an interactive hierarchical account and a parallel account of how object (or contour) recognition processes interact with other early visual processes. Attempts to elucidate what aspects of object memories (i.e., structure, function, semantics) are accessed at various processing stages will be critical in resolving this debate. (For research relevant to this issue see Gordon & Irwin, 1996; Henderson & Anes, 1994; Kahneman & Treisman, 1984; Kahneman, Treisman, & Gibbs, 1992; Peterson et al., 1991; Peterson et al., 1996; Treisman, 1988; Treisman, Kahneman, & Burkell, 1983.)

Attention and Object Perception

In this section, we consider research concerning the relationship between object perception and attention. We begin by considering whether attention makes something an object. Imagine attending to entities in order to count them, for example. Wolfe and Bennett (1997) define an object as ". . . a numerable thing as distinct from a collection of numerable things and as distinct from unenumerable 'stuff'." We can certainly count objects, but we can also count other things, like the spaces between the words on this line of text, for example. Enumerability does not make something an object. Similarly, one can attend to spaces as well as to objects, but attending does not necessarily make a space an object (Rubin, 1915/1958; Peterson & Gerhardstein, under review; Peterson, Gerhardstein, Mennemeier, & Rapcsak, 1998; Peterson & Gibson, 1994b). Thus, neither attention nor enumerability is sufficient for object perception. The related question of whether attention, or intention, on the viewer's part is necessary for object perception is currently being explored, as a consequence of pioneering work by Mack and Rock (1998). This topic is covered next. Then, in the following section, we consider the evidence indicating that there exists an object-based form of attention that is distinct from sparial attention.

Is Attention Necessary for Object Perception?

Inattentional Blindness

Mack and Rock (1998; Mack, Tang, Tuma, Kahn, & Rock, 1992; Rock, Linnett, Grant, & Mack, 1992) found that a large percentage of observers were effectively blind to the unexpected onset of an object when they were performing a difficult discrimination task. The discrimination task entailed judging which of the two arms of a cross was longer, when the difference between the arms was quite small. The cross was exposed briefly (e.g., 200 ms) and followed by a masking stimulus. On the third trial on which observers performed this task, a simple geometric object was presented in one of the four quadrants sketched by the cross at the same time that the cross was presented. When questioned shortly after this critical trial, many observers reported that they had not seen anything unusual. Some observers did report that something unusual had happened on the critical trial, but they were unable to report the simple geometric shape of the object that had been shown (e.g., a square or a triangle). Mack and Rock called this phenomenon "inattentional blindness." They argued that if the observer's attention or intention is not directed to perceiving an object, then object perception does not occur. Note that inattentional blindness is necessarily inferred from performance on a memory task rather than from performance on an online perception task. (Presumably, if observers knew they would have to occasionally detect objects, their perceptual intentions would change to accommodate this object detection task.)

The phenomenon of inattentional blindness raises the possibility that one may need attention (or intention) to perceive objects consciously. This in turn raises a question about terminology: Should the term "perception" (and the term "object perception" in particular) be reserved only for conditions in which observers can report being consciously aware of what they perceived? I argue that it should not. Research summarized in this chapter indicates that many of the component processes involved in object perception can be computed outside the observer's attentional focus, outside of awareness. Other evidence comes from work by C. M. Moore and Egeth (1997), who adapted the Mack and Rock paradigm and presented convincing evidence that grouping occurs without intention or attention. Regardless of whether or not the term "object perception" is ultimately reserved for conscious object perception, the important question of how attention or intention contributes to *conscious* object perception remains.

Stimulus Selection

A related debate in the search literature concerns whether or not it is possible for a stimulus to draw attention automatically if an observer is not intending to search for that stimulus in the first place. Pop-out effects have often been taken as evidence that unusual stimulus features or abrupt stimulus onsets can attract attention (e.g., Treisman, 1988; Yantis, 1993, 1996). "Pop-out" occurs when a single target stimulus differing on some basic feature from other "distractor" stimuli is detected quickly, and target detection latency does not increase as the number of distractors increases (e.g., the time to detect a red dot amongst green dots does not increase appreciably as the number of green dots increases). The quick detection responses were originally attributed to "stimulus selection" – the automatic allocation of attention to the distinct stimulus feature in the display. But Mack, Rock, and others pointed out that, in experiments demonstrating pop-out effects, observers are typically given advance information about the identity of their target feature or stimulus. Therefore, pop-out effects cannot serve as evidence that targets automatically attract attention in virtue of being different from the other display items. In experiments in which attention and task set were carefully controlled, Folk, Remington, and Johnston (1992) and Gibson and Kelsey (1998) failed to find evidence for stimulus selection, consistent with the view that task set determines what observers perceive. These results are consistent with the hypothesis that attention/intention is necessary for conscious perception, although it must be remembered that no "consciousness standard" exists; surely no single task currently in the experimentalist's repertoire provides a universally accepted standard.

Binding

Attention may be required to bind together the various properties of an object, such as its color, form, and movement (Treisman, 1988; Treisman & Gelade, 1980) as well. Treisman and her colleagues argued that, without attention, such features can be combined incorrectly, and illusory conjunctions can occur (e.g., illusory conjunctions of color and form, or form and motion). Prinzmetal (1981, 1995) showed that, when grouped displays are not attended, illusory conjunctions are more likely to occur within grouped entities than across grouped entities. (Note that Prinzmetal's results, like Moore and Egeth's, suggest that grouping itself can occur without attention.) Wolfe and Bennet (1997) recently demonstrated that attention is necessary to conjoin the features of an object, at least for conscious report. These authors argue that, prior to the allocation of attention, objects are nothing more than loose collections of basic features organized on the basis of spatio-temporal properties (i.e., Kahneman & Treisman's object files). However, it is important to remember that inaccessibility to conscious report does not necessarily imply that perceptual organization has not occurred. There is some evidence that binding has occurred, even when it cannot be measured via conscious reports (Robertson, 1998; Wojciulik & Kanwisher, 1998).

Experiments such as these indicate that attention must be considered if we are to understand object perception, but so must the relationship between perception and conscious report. I turn next to consider the evidence indicating that a specialized form of object-based attention exists.

Object-Based Attention

It has long been known that attention can be allocated to locations in space that are different from the location where the eyes are directed (Posner, 1980). More recently, it has been shown that attention can be allocated to objects independent of the spaces they occupy (e.g., Duncan, 1984; Driver & Halligan, 1991; Gibson & Egeth, 1994; Treisman, Kahneman & Burkell, 1983). Evidence that attention can be "object-based" and not just "space-based" takes various forms. One form of evidence for "object-based" attention, is that it takes longer to move attention a given distance between two objects than the same distance within an object (Egly, Driver, & Rafal, 1994; Egly, Rafal, Driver, & Starrveveld, 1994). Another manifestation of object-based attention is that observers require less time to report about two features of a single object than about two features of different objects. This second effect is obtained even when the two objects overlap each other and occupy essentially the same location (Duncan, 1984; Goldsmith, 1998), and even when portions of the single object are occluded by another object (Behrmann, Zemel, & Mozer, 1998). A third demonstration of object-based attention entails moving objects. When a cued object moves to a new location, attention moves with the object, rather than (or in addition to)

---

staying in the cued location (Kahneman, Treisman, & Gibbs, 1992; Tipper, Brehau
Driver, 1990; Tipper, Driver, & Weaver, 1991). Thus, attention seems (a) to spread
readily within an attended object than between two objects, (b) to encompass the pe
tible features of an attended object, and (c) to move with an object.

In summary, the study of object perception and the study of attention are curr
intertwined. Objects may not be perceived consciously unless they are encompassed b
observer's intentions. Once objects are perceived consciously, they form a unique subs
for the spread of attention that is distinct from a purely spatial substrate. Questions re
ing the relationship between space and objects have been raised throughout this cha
and will continue to be raised in the future. Questions regarding the relationship bet
conscious perception and action are important as well (see Chapters 7 and 10) for elal
tion of these points.

Suggested Readings

Parasuraman, R. (Ed.) (1998). *The attentive brain.* Cambridge, MA: MIT Press.
Hochberg, J. (Ed.) (1998). *Perception and cognition at century's end.* NY: Academic Press.
Inui, T., & McClelland, J. L. (Eds.) (1996). *Attention and performance, XVI: Information integ*
in perception and communication. Cambridge, MA: MIT Press.
Tarr, M. J., & Bülthoff, H. (Eds.) (1999). *Object recognition in man, monkey, and machine.*
bridge, MA: MIT Press. (Printed from a special issue of *Cognition* (1998), 67 (1–2).)

Additional Topics

Colour and Surface Detail

To what extent do color and surface detail influence object recognition? (Biederman & Ju,
Price & Humphreys, 1989; J. W. Tanaka & Presnell, 1999)

Are Holes Objects?

For discussions of holes, see Cassati & Varzi (1995); Bloom (1996); Bloom & Giralt (in p
Hochberg, 1998.

Tactile and Auditory Object Perception

What general principles cut across different modalities in which object perception occurs (e.g.,
principles are shared by visual, tactile, and auditory object perception and what specific prin
are employed in each modality)? For work on auditory object perception see Darwin and
(1999). For work on tactile object perception see Klatzsky and Lederman (1995, 1999).

References

Battelli, L., Casco, C., & Sartori, G. (1997). Dissociation between contour-based and texture-
shape perception: A single case study. *Visual Cognition, 4,* 275–310.
Baylis, G. C., & Driver, J. (1992). Visual parsing and response competition: The effect of gro
factors. *Perception & Psychophysics, 51,* 141–162.

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

Baylis, G. C., & Driver, J. (1994). Parallel computation of symmetry, but not repetition in single visual objects. *Visual Cognition, 1,* 377–400.

Baylis, G. C., & Driver, J. (1995). Obligatory edge-assignment in vision: The role of figure- and part-segmentation in symmetry detection. *Journal of Experimental Psychology: Human Perception & Performance, 21,* 1323–1342.

Beck, J., Rosenfeld, A., & Ivry, R. (1990). Line segregation. *Spatial Vision, 4,* 75–101.

Behrmann, M., Zemel, R. S., & Mozer, M. C. (1998). Object-based attention and occlusion: Evidence from normal participants. *Journal of Experimental Psychology: Human Perception & Performance, 24,* 1011–1036.

Ben-Av, M. B., & Sagi, D. (1995). Perceptual grouping by similarity and proximity: Experimental results can be predicted by intensity autocorrelations. *Vision Research, 35,* 853–866.

Biederman, I. (1985). Human image understanding: Recent research and a theory. *Computer Vision, Graphics, and Image Processing, 32,* 29–73.

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

Biederman, I. (1995). Visual object recognition. In S. F. Kosslyn & D. N. Oshershon (Eds.), *An invitation to cognitive science* (2nd ed., pp. 121–165). Cambridge, MA: The MIT Press.

Biederman, I., & Gerhardstein, P.C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception & Performance, 19,* 1162–1182.

Biederman, I., & Gerhardstein, P.C. (1995). Viewpoint-dependent mechanisms in visual object recognition: A reply to Tarr and Bülthoff (1995). *Journal of Experimental Psychology: Human Perception & Performance, 21,* 1506–1514.

Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology, 20,* 38–64.

Binford, T. O. (1981). Inferring surfaces from images. *Artificial Intelligence, 17,* 205–244.

Bloom, P. (1996). Possible individuals in language and cognition. *Current Directions in Psychological Science, 5,* 90–94.

Bloom, P., & Girault, N. (in press). *Psychological Science.*

Boselie, F. (1994). Local and global factors in visual occlusion. *Perception, 23,* 517–528.

Boselie, F., & Leeuwenberg, E. (1986). A test of the minimum principle requires a perceptual coding system. *Perception, 15,* 331–354.

Braddick, O. (1992). Motion may be seen but not used. *Current Biology, 2,* 597–599.

Braun, J. (1993). Shape-from-shading is independenr of visual attention and may be a 'texton.' *Spatial Vision, 7,* 311–322.

Braunstein, M. L, Hoffman, D. D., & Saidpour, A. (1989). Parts of visual objects: An experimental test of the minima rule. *Perception, 18,* 817–826.

Brown, J. M., & Weisstein, N. (1988). A spatial frequency effect on perceived depth. *Perception & Psychophysics, 44,* 157–166.

Brown, J. M., Weisstein, N., & May, J. G. (1992). Visual search for simple volumetric shapes. *Perception & Psychophysics, 51,* 40–48.

Bruno, N., Bertamini, M., & Domini, F. (1997). Amodal completion of partly occluded surfaces: Is there a mosaic stage? *Journal of Experimental Psychology: Human Perception & Performance, 23,* 1412–1426.

Bülthoff, H. H. (1991). Shape from X: Psychophysics and computation. In M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing.* Cambridge, MA: The MIT Press.

Bülthoff, H. H., & Edelman, S. (1993). Evaluating object recognition theories by computer psychophysics. In T. A. Poggio & D. A. Glaser (Eds.), *Exploring brain functions: Models in neuroscience* (pp. 139–164). NY: Wiley.

Cassati, R., & Varzi, A. C. (1995). *Holes and other superficialities.* Cambridge, MA: MIT Pres

Cave, C. B., & Kosslyn, S. M. (1993). The role of parts and spatial relations in object identifica *Perception,* 22, 229–248.

Corballis, M. C. (1988). Recognition of disoriented shapes. *Psychological Review, 95,* 115–12

Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances: The integra relative potency, and contextual use of different information about depth. In W. Epstein Rogers (Eds.), *Perception of space and motion* (pp. 71–118). San Diego: Academic Press.

Darwin, C. J., & Hukin, R. W. (1999). Auditory objects of attention: The role of interaural differences. *Journal of Experimental Psychology: Human Perception & Performance, 25,* 617–

Davis, G., & Driver, J. (1994). Parallel detection of Kanizsa subjective figures in the human system. *Nature, 371,* 791–793.

Dresp, B., & Bonnet, C. (1993). Psychophysical measures of illusory form perception: Fu evidence for local mechanisms. *Vision Research, 33,* 759–766.

Driver, J., & Baylis, G. C. (1989). Movement and visual attention: The spotlight metaphor b down. *Journal of Experimental Psychology: Human Perception & Performance, 15,* 448–456.

Driver, J., Baylis, G. C., & Rafal, R. D. (1992). Preserved figure-ground segregation and symm perception in visual neglect. *Nature* (London), *360,* 73–75.

Driver, J., & Halligan, P. W. (1991). Can visual neglect operate in object-centered coordinates affirmative single case study. *Cognitive Neuropsychology, 8,* 475–496.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Ex mental Psychology: General, 119,* 501–517.

Edelman, S., & Weinshall, D. (1991). A self-organizing multiple view representation of 3D obj *Biological Cybernetics, 64,* 209–219.

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locati Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: Gen 123,* 161–177.

Egly, R., Rafal, R., Driver, J., & Starrveveld, Y. (1994). Covert orienting in the split brain re hemispheric specialization for object-based attention. *Psychological Science, 5,* 380–383.

Enns, J. T., & Rensink, R. A. (1991). Preattentive recovery of three-dimensional orientation f line drawings. *Psychological Review, 98,* 335–351.

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

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in primate vi cortex. *Cerebral Cortex, 1,* 1–47.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual syst Evidence for a local "association field." *Vision Research, 33,* 175–193.

Folk, C.L., Remington, R.W., & Johnston, J. C. (1992). Involuntary covert orienting is contin on attentional control settings. *Journal of Experimental Psychology: Human Perception & Perf ance, 18,* 1030–1044.

Fox, E. (1998). Perceptual grouping and visual selective attention. *Perception & Psychophysics,* 1004–1021.

Gibson, B. S., & Egeth, H. (1994). Inhibition and disinhibition of return: Evidence from temp order judgments. *Perception & Psychophysics, 56,* 669–680.

Gibson, B. S., & Kelsey, E. M. (1998). Stimulus-driven attentional capture is contingent on attenti set for displaywide visual features. *Journal of Experimental Psychology: Human Perception & formance, 24,* 699–706.

Gibson, B. S., & Peterson, M. A. (1994). Does orientation-independent object recognition prec orientation-dependenr recognition? Evidence from a cueing paradigm. *Journal of Experime Psychology: Human Perception & Performance, 20,* 299–316.

Goldsmith, M. (1998). What's in a location? Comparing object-based and space-based model feature inregration in visual search. *Journal of Experimental Psychology: General, 27,* 189–219

Gordon, R. D., & Irwin, D. E. (1996). What's in an object file? Evidence from priming stud *Perception & Psychophysics, 58,* 1260–1277.

Göttschaldt, K. (1926/1938). Gestalt factors and repetition (continued). In W. D. Ellis (Ed.), *A sourcebook of Gestalt psychology*. London: Kegan Paul, Trech, Tubner, and Co., Ltd.

Graham, N. V. S. (1989). *Visual pattern analyzers*. New York: Oxford University Press.

Graham, N. V. S. (1992). Breaking the visual stimulus into parts. *Current Directions in Psychological Science, 1,* 55–60.

Grosof, D. H., Shapley, R. M., & Hawken, M. J. (1993). Macaque V1 neurons can signal illusory contours. *Nature, 365,* 550–552.

Grossberg, S., & Mingolla, E. (1985). Neural dynamics of perceptual grouping: Textures, boundaries and emergent segmentations. *Perception & Psychophysics, 38,* 141–171.

Gurnsey, R., Humphrey, G. K., & Kapitan, P. (1992). Parallel discrimination of subjective contours defined by offset gratings. *Perception & Psychophysics, 52,* 263–276.

Han, S., Humphreys, G., & Chen L. (1999a). Uniform connectedness and classical Gestalt principles of perceptual grouping. *Perception & Psychophysics, 61,* 661–674.

Han, S., Humphreys, G. W., & Chen L. (1999b). Parallel and competitive processes in hierarchical analysis: Perceptual grouping and encoding of closure. *Journal of Experimental Psychology: Human Perception & Performance, 25,* 1411–1432.

Harms, L., & Bundeson, C. (1983). Color segregation and selective attention in a nonsearch task. *Perception & Psychophysics, 33,* 11–19.

Hayward, W. G. (1998). Effects of outline shape in object recognition. *Journal of Experimental Psychology: Human Perception & Performance, 24,* 427–440.

He, Z. J., & Nakayama, K. (1992). Surfaces versus features in visual search. *Nature, 359,* 231–233.

He, Z. J., & Nakayama, K. (1994). Perceiving textures: Beyond filtering. *Vision Research, 34,* 151–162.

Heilman, K. M., & Van Den, A. (1980). Right hemisphere dominance for attention: The mechanism underlying hemispheric asymmetries of inattention (neglect). *Neurology, 30,* 327–330.

Henderson, J. M., & Anes, M. D. (1994). Roles of object-file review and type priming in visual identification within and across eye fixations. *Journal of Experimental Psychology: Human Perception & Performance, 20,* 826–839.

Hirsch, E. (1982). *The concept of identity*. New York: Oxford University Press.

Hochberg, J. (1968). In the mind's eye. In R. N. Haber (Ed.), *Contemporary theory and research in visual perception* (pp. 309–331). New York: Holt, Rinehart, & Winston.

Hochberg, J. (1971). Perception I: Color and shape. In J. W. Kling & L. A. Riggs (Eds.), *Woodworth and Schlossberg's experimental psychology,* (3rd ed., pp. 395–474). New York: Hold, Rinehart, & Winston.

Hochberg, J. (1978). *Perception* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall, Inc.

Hochberg, J. (1980). Pictorial functions and perceptual structures. In M. A. Hagen (Ed.), *The perception of pictures* (Vol. 2, pp. 47–93). New York: Academic Press.

Hochberg, J. (1998). Gestalt theory and its legacy. In J. Hochberg (Ed.), *Perception and cognition at century's end* (pp. 253–306). New York: Academic Press.

Hochberg, J., and Peterson, M. A. (1993). Mental representations of occluded objects: Sequential disclosure and intentional construal. *Giornale Italiano di Psicologia, 20,* 805–820. (Monograph edition published in English in honor of Gaetano Kanizsa.)

Hoffman, D. D., & Richards, W. A. (1985). Parts of recognition. In S. Pinker (Ed.), *Visual cognition*. Cambridge, MA: MIT Press.

Hoffman, D. D., & Singh, M. (1997). Saliende of visual parts. *Cognition, 63,* 29–78.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology (London), 166,* 106–154.

Hummel, J., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review, 99,* 480–517.

Humphrey, G. K., Symons, L. A., Herbert, A. M., & Goodale, M. A. (1996). A neurological dissociation between shape from shading and shape from edges. *Behavioral Brain Research, 76,* 117–125.

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

Ittelson, W. (1996). Visual perception of markings. *Psychonomic Bulletin & Review, 3,* 171–187.

Jolicoeur, P. (1985). The time to name disoriented objects. *Memory & Cognition, 13,* 289–303.

Jolicoeur, P. (1988). Mental rotation and the identification of disoriented objects. *Canadian Journa of Psychology, 42,* 461–478.

Julesz, B. (1971). *Foundations of Cyclopean perception*. Chicago: University of Chicago Press.

Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature (London), 290,* 91–97.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R Parasuraman (Ed.), *Varieties of attention* (pp. 29–61). New York: Academic Press.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object specifi integration of information. *Cognitive Psychology, 24,* 175–215.

Kanizsa, G. (1987). Quasi-perceptual margins in homogeneously stimulated fields. In S. Petry & G Meyer (Eds.), *The perception of illusory contours* (W. Gerbino, Trans.) (pp. 40–49). New York Springer-Verlag.

Kanizsa, G., & Gerbino, W. (1976). Convexity and symmetry in figure-ground organization. In M Henle (Ed.), *Vision and artifact*. New York: Springer Publishing Co.

Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys *Neuron, 15,* 843–856.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology, 23,* 141–221.

Kellman, P. J., Yin, C., & Shipley, T.F. (1998). A common mechanism for illusory and occluded object completion. *Journal of Experimental Psychology: Human Perception & Performance, 24,* 859–869.

Kennedy, J. M. (1973). Misunderstandings of figure and ground. *Scandinavian Journal of Psychology, 14,* 207–209.

Kennedy, J. M. (1974). *A psychology of picture perception*. San Francisco: Jossey-Bass Publishers.

Kimchi, R. (1994). The role of wholistic/configural properties versus global properties in visual form perception. *Perception, 23,* 489–504.

Kimchi, R. (1998). Uniform connectedness and grouping in the perceptual organization of hierarchical patterns. *Journal of Experimental Psychology: Human Perception & Performance.*

Kinsbourne, M., (1970). The cerebral basis of lateral asymmetries in attention. *Acta Psychologica, 33,* 193–201.

Klatzky, R. L., & Lederman, S. J. (1995). Identifying objects from a haptic glance. *Perception & Pychophysics, 57,* 1111–1123.

Klatzky, R. L., & Lederman, S. J. (1999). The haptic glance: A route to rapid object identification and manipulation. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII. Cognitive regulation of performance: Integration of theory and application*. Mahwah, NJ: Erlbaum.

Kobatake, E., & Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology, 71,* 856–867.

Koffka, K. (1935) *Principles of Gestalt psychology*. New York: Harcourt, Brace, & World, Inc.

Köhler, W. (1929/1947) *Gestalt psychology*. New York: New American Library.

Kovács, I., & Julesz, B. (1994). Perceptual sensitiviry maps within globally defined visual shapes. *Nature, 370,* 25 August, 644–646.

Kramer, A. F., & Jacobson, A. (1991). Perceptual organization and focused attention: The role of objects and proximiry in visual processing. *Perception & Psychophysics, 50,* 267–284.

Kurbat, M. (1994). Structural description theories: Is RBC/JIM a general purpose theory of human entry-level object recognition? *Perception, 23,* 1339–1368.

Landy, M. S., Maloney, L. T, & Young, M. J. (1990). Psychophysical estimation of the human depth combination rule. Sensor Fusion III: 3-D perception and recognition. *SPIE, 1383,* 247–254.

Leonards, U., Singer, W., & Fahle, M. (1996). The influence of temporal phase differences on texture segmentation. *Vision Research, 36,* 2689–2697.

Lin, E. L., & Murphy, G. L. (1997). Effects of background knowledge on object categorization and part detection. *Journal of Experimental Psychology: Human Perception & Performance, 23,* 1153–1169.

Logothetis, N. K., Pauls, J., Bueltoff, H. H., & Poggio, T. (1994). Viewpoint dependent object recognition in monkeys. *Current Biology, 4,* 401–414.

Logothetis, N. K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Current Biology, 5,* 552–563.

Lowe, D. (1985). *Perceptual organization and visual recognition.* Boston: Kluwer.

Lowe, D. (1987). Three-dimensional object recognition from single two-dimensional images. *Artificial intelligence, 31,* 355–395.

Mack, A., & Rock, I. (1998). *Inattentional blindness.* Cambridge, MA: MIT Press.

Mack, A., Tang, B., Tuma, R., Kahn, S., & Rock, I. (1992). Perceptual organization and attention. *Cognitive Psychology, 24,* 475–501.

Marr, D. (1982). *Vision.* San Francisco: W. H. Freeman.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London, B, 200,* 269–291.

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review, 86,* 287–330.

McClelland, J. L. (1985). Putting knowledge in its place: A scheme for programming parallel processing structures on the fly. *Cognitive Science, 9,* 113–146.

McClelland, J. L., & Rumelhart, D. E. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 2). Cambridge, MA: The MIT Press.

Moore, C., & Cavanagh, P. (1998). Recovery of 3D volume from 2-tone images of novel objects. *Cognition, 67,* 45–71.

Moore, C. M., & Egeth, H. (1997). Perception without attention: Evidence of grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception & Performance, 23,* 339–352.

Moore, M. G. (1930). Gestalt vs. experience. *American Journal of Psychology, 42,* 543–455.

Mozer, M. C., Zemel, R. S., Behrmann, M., & Williams, C. K. (1992). Learning to segment images using dynamic feature binding. *Neural Computation, 4,* 650–665.

Mumford, D., Kosslyn, S. M., Hillger, L. A., & Herrnstein, R. J. (1987). Discriminating figure from ground: The role of edge detection and region-growing. *Proceedings of the National Academy of Sciences, 84,* 7354–7358.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton, Century, Crofts.

O'Shea, R. P., Blackburn, S. G., & Ono, H. (1994). Contrast as a depth cue. *Vision Research, 34,* 1595–1604.

Palmer, S., & Rock, I. (1994a). Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin & Review, 1,* 29–55.

Parker, A. J., Cumming, B. G., Johnston, E. B., & Hurlbert, A. C. (1995). Multiple cues for three-dimensional shape. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 351–364). Cambridge, MA: The MIT Press.

Perrett, D., Smith, P., Potter, D., Mistlin, A., Head, A., Milner, A., & Jeeves, M. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society, London,* [Biol], *223,* 293–317.

Peterhans, E., & von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *Journal of Neuroscience, 9,* 1749–1763.

Peterson, M. A. (1994a). Shape recognition can and does occur before figure-ground organization. *Current Directions in Psychological Science, 3,* 105–111.

Peterson, M. A. (1994b). The proper placement of uniform connectedness. *Psychonomic Bulletin and Review, 1,* 509–514.

Peterson, M. A. (1999). What's in a stage name? *Journal of Experimental Psychology: Human Perception & Performance, 25,* 276–286.

Peterson, M. A., de Gelder, B., Rapcsak, S. Z., Gerhardstein, P. C., & Bachoud-Lévi, A.-C. (Object memory effects on figure assignment: Conscious object recognition is not necess sufficient. *Vision Research, 40,* 1549–1567.

Peterson, M. A., & Gerhardstein, P. C. (under review). Effects of region-centered attentio object memory on figure assignment.

Peterson, M. A., Gerhardstein, P., Mennemeier, M., & Rapcsak, S. V. (1998). Object-ce attentional biases and object recognition contributions to scene segmentation in right hemis and left hemisphere-damaged patients. *Psychobiology, 26,* 557–570.

Peterson, M. A., & Gibson, B. S. (1993). Shape recognition contributions to figure-ground o zation in three-dimensional displays. *Cognitive Psychology, 25,* 383–429.

Peterson, M. A., & Gibson, B. S. (1994a). Must shape recognition follow figure-ground org tion? An assumption in peril. *Psychological Science, 5,* 253–259.

Peterson, M. A., & Gibson, B. S. (1994b). Object recognition contributions to figure-grou ganization: Operations on outlines and subjective contours. *Perception & Psychophysics, 56,* 564.

Peterson, M. A., Harvey, E. H., & Weidenbacher, H. L. (1991). Shape recognition inputs to f ground organization: Which route counts? *Journal of Experimental Psychology: Human Perc & Performance, 17,* 1075–1089.

Peterson, M. A., & Hector, J. E. (1996, November). *Evidence for the piecemeal nature of pre- object recognition processes.* Paper presented at the Annual Meeting of the Psychonomic Sc Chicago, IL.

Peterson, M. A., & Hochberg, J. (1983). Opposed-set measurement procedure: A quantitative ar of the role of local cues and intention in form perception. *Journal of Experimental Psych Human Perception & Performance, 9,* 183–193.

Peterson, M. A., & Hochberg, J. (1989). Necessary considerations for a theory of form perce A theoretical and empirical reply to Boselie and Leeuwenberg. *Perception, 18,* 105–119.

Peterson, M. A., Nadel, L., Bloom, P., & Garrett, M. F. (1996). Space and language. In P. Bl M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space* (pp. 553–577). bridge, MA: MIT Press.

Poggio, T., & Edelman, S. (1990). A network that learns to recognize three-dimensional ob *Nature, 343,* 263–266.

Pomerantz, J. R., & Kubovy, M. (1986). Theoretical approaches to perceptual organization: plicity and likelihood principles. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handb perception and performance, volume II: Cognitive processes and performance* (pp. 36:1–46). York: John Wiley & Sons.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32* 25.

Prazdny, K. (1983). Illusory contours are not caused by simultaneous brightness contrast. *Perce & Psychophysics, 34,* 403–404.

Price, C. J., & Humphreys, G. W. (1989). The effects of surface detail on object categorizatio naming. *Quarterly Journal of Experimental Psychology, 41A,* 797–828.

Prinzmetal, W. (1981). Principles of feature integration in visual perception. *Perception & Ps physics, 30,* 330–340.

Prinzmetal, W. (1995). Visual feature integration in a world of objects. *Current Directions in Ps logical Science, 4,* 90–94.

Ramachandran, V. S. (1988). Perception of shape from shading. *Nature* (London), *331,* 163–

Rensink, R., & Enns, J. T. (1995). Preemption effects in visual search: Evidence for low-level grou *Psychological Review, 102,* 101–130.

Riddoch, M. J., & Humphreys, G. W. (1987). A case of integrative visual agnosia. *Brain,* 1431–1462.

Robertson, L. C. (1998). Visuospatial attention and cognitive function: Their role in object pe tion. In R. Parasuraman (Ed.), *The attentive brain.* Cambridge, MA: MIT Press.

Rock, I. (1987). A problem-solving approach to illusory contours. In S. Petry & G. Meyer (

The perception of illusory contours (pp. 462–70). New York: Springer-Verlag.

Rock, I., & Brosgole, L. (1964). Grouping based on phenomenal proximity. *Journal of Experimental Psychology, 67*, 531–538.

Rock, I., Linnert, C. M., Grant, P., & Mack, A. (1992). Perception without attention: Results of a new method. *Cognitive Psychology, 24*, 502–534.

Rubin, E. (1958). Figure and ground. In D. Beardslee & M. Wertheimer (Ed. and Trans.), *Readings in perception* (pp. 35–101). Princeton, NJ: Van Nostrand. (Original work published 1915.)

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests of and extensions of the model. *Psychological Review, 89*, 60–94.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1). Cambridge, MA: The MIT Press.

Saiki, J., & Hummel, J. (1998). Connectedness and the integration of parts with relations in shape perception. *Journal of Experimental Psychology: Human Perception & Performance, 24*, 227–251.

Schyns, P., Goldstone, R. L., & Thilbaut, J-P. (1998). The development of features in object concepts. *Behavioral & Brain Sciences, 21*, 1–54.

Sekuler, A. B. (1994). Local and global minima in visual completion: Effects of symmetry and orientation. *Perception, 23*, 529–545.

Sekuler, A. B., & Palmer, S. E. (1992). Perception of partly occluded objects: A microgenetic analysis. *Journal of Experimental Psychology: General, 121*, 95–111.

Solms, M., Turnbull, O. H., Kaplan-Solms, K., & Miller, P. (1998). Rotated drawing: The range of performance and anatomical correlates in a series of 16 patients. *Brain & Cognition, 38*, 358–368.

Spelke, E. S. (1988). Where perceiving ends and thinking begins: The apprehension of objects in infancy. In A. Yonas (Ed.), *Perceptual development in infancy: The Minnesota symposium on child psychology* (Vol. 20, pp. 197–234). Hillsdale. NJ: Lawrence Erlbaum Associates.

Spelke, E. S. (1990). Principles of object perception. *Cognitive Science, 14*, 29–56.

Spelke, E. S., Gutheil, G., & Van de Walle, G. (1995). The development of object perception. In S. M. Kosslyn & D. N. Osherson (Eds), *Visual cognition: An invitation to cognitive science* (Vol. 2, 2nd ed.). Cambridge, MA: MIT Press.

Stevens, K. A., & Brooks, A. (1988). The concave cusp as determiner of figure-ground. *Perception, 17*, 35–42.

Sun, J., & Perona, P. (1997). Shading and stereo in early perception of shape and reflectance. *Perception, 26*, 519–529.

Suzuki, S., Peterson, M. A., Moscovitch, M., & Berhmann, M. (under review). Identification of one-part and two-part volumetric objects: Selective deficits in coding spatial arrangement of part in visual object agnosia.

Tanaka, J. W., & Presnell, L. M. (1999). Color diagnosticity in object recognition. *Perception & Psychophysics, 61*, 1140–1153.

Tanaka, K. (1993). Neuronal mechanisms of object recognition. *Science, 262*, 685–688.

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.

Tarr, M. J., & Bülthoff, H. H. (1998). Image-based recognition in man, monkey, and machine. *Cognition, 67*, 1–20.

Tarr, M. J., & Bülthoff, H. H. (1995). Is human object recognition better described by geon structural descriptions or by multiple views? Comment on Biederman and Gerhardstein (1993). *Journal of Experimental Psychology: Human Perception & Performance, 21*, 1494–1505.

Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology, 21*, 233–282.

Tipper, S. P., Brehaut, J., & Driver, J. (1990). Selection of moving and static objects for the control of spatially-directed attention. *Journal of Experimental Psychology: Human Perception & Performance, 16*, 492–504.

Tipper, S. P., Driver, J., & Weaver, B. (1991). Object-centered inhibition of return of visual atten-

tion. *Quarterly Journal of Experimental Psychology, 43A*, 289–298.

Tittle, J. S., Norman, J. F., Perotti, V. J., & Phillips, F. (1997). The perception of scale-depen and scale-independent surface structure from binocular disparity, texture, and shading. *Pe tion, 26*, 147–166.

Tononi, G., Sporns, O., & Edelman, G. M. (1991). Modeling perceptual grouping and fig ground segregation by means of active reentrant connections. *Proceedings of the National Aca of Sciences, 88*, 129–133.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psycho 12*, 97–136.

Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *Quar Journal of Experimental Psychology, 40A*, 201–237.

Treisman, A., Kahneman, D., & Burkell, J. (1983) Perceptual objects and the costs of filter *Perception & Psychophysics, 33*, 527–532.

Turnbull, O. H., & McCarthy, R. A. (1996). When is a view unusual? A single case stud orientation-dependent visual agnosia. *Brain Research Bulletin, 40*, 497–503.

Ullman, S. (1990). Three-dimensional object recognition. *Cold Spring Harbor Symposium on Qu titative Biology, 50*, 1243–1258.

Vecera, S. P., & O'Reilly, R. C. (1998). Figure-ground organization and object recognition p esses: An interactive account. *Journal of Experimental Psychology: Human Perception & Perfo ance, 24*, 441–462.

Vecera, S. P., & O'Reilly, R. C. (2000). A reply to Peterson. *Journal of Experimental Psychol Human Perception & Performance.*

von der Heydt, R., & Peterhans, E. (1989). Mechanisms of contour perception in monkey vis cortex. I. Lines of pattern discontinuity. *Journal of Neuroscience, 9*, 1731–1748.

Wallach, H., & Slaughter, V. (1988). The role of memory in perceiving subjective contours. *Per tion & Psychophysics, 43*, 101–106.

Warrington, E. K. (1982). Neuoropsychological studies of object recognition. *Philosophical Tra actions of the Royal Society of London, B, 298*, 15–33.

Watson, S. E., & Kramer, A. F. (1999). Object-based visual selective attention and perceptual ganization. *Perception & Psychophysics, 61*, 31–49.

Westheimer, G. (1999). Gestalt theory reconfigured: Max Wertheimer's anticipation of recent velopments in visual neuroscience. *Perception, 18*, 5–15.

Wiggins, D. (1980). *Sameness and substance.* Oxford, UK: Basil Blackwell.

Witkin, A. P., & Tenenbaum, J. M. (1983). On the role of structure in vision. In J. Beck, B. Ho & A. Rosenfeld (Eds.), *Human and machine vision.* NY: Academic Press.

Wojciulik, E., & Kanwisher, N. (1998). Implicit but not explicit feature binding in a Balint's tient. *Visual Cognition, 5*, 157–181.

Wolfe, J. M., & Bennett, S. C. (1997). Preattentive object files: Shapeless bundles of basic featur *Vision Research, 37*, 25–43.

Wong, E., & Weisstein, N. (1983). Sharp targets are detected better against a figure, and blurr targets are detected better against a background. *Journal of Experimental Psychology: Human P ception & Performance, 9*, 194–202.

Woodworth, E. G. (1938). *Experimental psychology.* New York: Henry Holt & Company.

Yantis, S. (1993). Stimulus-driven attentional capture. *Current Directions in Psychological Science,* 156–171.

Yantis, S. (1996). Attentional capture in vision. In A. F. Kramer, M. G. H. Coles, & G.D. Log (Eds.), *Converging operations in the study of visual selective attention* (pp. 45–76). Washingto DC: American Psychological Association.

Zeki, S. (1993). *A vision of the brain.* Oxford: Blackwell.

Zemel, R. S., Behrmann, M., Mozer, M. C., & Bavelier, D. (under review). Experience-depende perceptual grouping and object-based attention.

Zipser, K., Lamme, V. A. F., & Schiller, P. H. (1996). Contextual modulation in primary visu cortex. *The Journal of Neuroscience, 16*, 7376–7389.