3 On Figures, Grounds, and Varieties of Surface Completion

Mary A. Peterson
University of Arizona

FIGURE-GROUND SEGREGATION: UNANSWERED QUESTIONS

Two regions can be defined for each edge in the visual field, one lying on each side. Typically, the edge appears to shape only one of these two regions, and figure-ground segregation is said to occur. Figures are regions of the visual field that appear to have a definite shape, a shape bestowed in part by their bounding contour (i.e., their edge). If figures are familiar objects, they can be identified (barring brain damage or the imposition of external noise). The region adjacent to the figure is locally shapeless near the edge it shares with the figure; this shapeless region is called the ground (short for background) because it often appears to complete amodally behind the figure. Although figure-ground segregation is a venerable topic in visual perception (cf. Rubin, 1915/1958), a number of questions about figures and grounds remain unanswered.

One critical longstanding question is, Why can figures be identified if they are familiar, whereas grounds cannot? A traditional answer to this question is that the determination of figure versus ground precedes access to object memories and that figures, not grounds, are matched to object memories. On this view, grounds cannot be identified because they are not matched to object memories. The observation
that identifiability is coupled to figural status has been taken as support for this figure-ground first assumption.

Consider Fig. 3.1. When the black region appears to be the figure, it can be identified as a vase; the white regions appear shapeless near the borders of the vase and seem to continue behind the vase. Alternatively, when the white regions appear to be the figures, they can be identified as face profiles, whereas the black region appears to be shapeless near the borders of the faces, apparently continuing behind them. The coupling between identifiability and figural status does not constitute support for the figure-ground first assumption, however, because couplings cannot support inferences of causality (Peterson, 1999a). Many theorists adopt the causal approach for functional reasons, arguing that it would be computationally inefficient to match against object memories those regions that will ultimately appear shapeless. However, when arguments regarding computational inefficiency are raised, it is worth pointing out that the design features of vision are not necessarily those that are deemed efficient a priori.

A second question concerning figure-ground segregation, related to the first question, is, Why is the region adjacent to the figure shapeless? Some have answered this question by claiming that one-sided contour assignment is obligatory

(Driver & Baylis, 1996) and that a region lacking a contour (as the ground does) is necessarily shapeless. But one-sided contour assignment is not obligatory. It has long been known that certain contours shape both adjacent regions and that others shape neither region (Kennedy, 1974, p. 96). Moreover, it is generally agreed that the two regions on either side of a contour (or edge) are assessed for the Gestalt configural cues of symmetry, relative convexity, relative area, closure, and so on.

FIG. 3.2. Sample displays used in our research. A and B were used by Peterson et al (1991). Gestalt configural cues and monocular depth cues favor the interpretation that the black center region is the figure. The white surround regions portrait two standing women (in A) and two profile faces (in B). C is a schematic of a stimulus used by Peterson and Gibson (1994a). "Must shape recognition follow figure-ground organization? An assumption in peril." Psychological Science, Blackwell Publishing Co. Reprinted with permission. The white region is symmetric around a vertical axis, whereas the black asymmetric region portrays a portion of a seahorse. Thus, for purposes of figure assignment at the edge shared by the black and white regions, the Gestalt configural cue of symmetry competes with an object memory cue. D and E were used by Peterson and Gibson (1993). In these displays, the black and white regions are approximately equated for area and convexity, but the black region is high in denotivity, portraying a portion of a table lamp in D and a portion of a guitar in E.
before figure assignment is complete. Given that a substantial amount of configural processing has taken place on both sides of the contour, it remains to be explained why the region adjacent to the figure is perceived to be shapeless when one-sided contour assignment does occur. The question regarding apparent shapelessness applies regardless of whether the apparently shapeless region would be a familiar or a novel shape if it were seen as the figure (e.g., compare the white regions of Fig. 3.2, parts A and E).

A third unanswered question concerning figure-ground segregation is whether regions adjacent to figures are necessarily completed behind the figure. Can they be completed in front of the figure? Must they be completed at all? The answer will ultimately depend on the extent to which the factors that influence figure-ground segregation determine depth segregation.

A fourth question, related to the third, is whether or not two-dimensional displays are good surrogates for three-dimensional displays. The terms figure and ground, introduced by Rubin and used by the Gestalt psychologists, were originally used to describe the perception of two-dimensional displays. The Gestalt psychologists explicitly assumed that two-dimensional displays were good surrogates for three-dimensional displays (e.g., Rubin, 1915/1958); more recently, others have argued that they are not (e.g., Marr, 1982). In fact, two-dimensional displays may be good surrogates for three-dimensional displays for investigating some questions but not others.

A fifth question concerning figure-ground segregation is whether or not it should be considered a stage of processing. Many investigators consider figure-ground segregation to be a stage of processing that serves to separate the wheat of candidate objects from the chaff of shapeless grounds. Following figure-ground segregation, the former can be matched to object memories, and can serve as targets for action, whereas the latter can be safely eliminated from further processing. It is easy to conceptualize figure-ground segregation as a stage of processing if one accepts the assumption that figure-ground assignment precedes access to object memories. Hence, the answer to the fifth question is likely to be related to the answer to the first question.

CHAPTER OUTLINE

In the remainder of this chapter, I will address the five questions raised previously. In the second section, I review evidence indicating that, contrary to traditional assumptions, memories of object structure (at least) are accessed before figural status is determined (Question 1). Next, I discuss a model that accounts for these effects and provides a mechanism for the apparent shapelessness of grounds (Question 2). Recent evidence supporting the model is reviewed. At the end of the third section, I begin to consider whether regions adjacent to a figure are necessarily amodally completed behind the figure (Question 3). Then, I show that in some ways two-dimensional displays are good surrogates for three-dimensional displays, whereas in other ways they are not (Question 4). Consideration of three-dimensional displays has led me to conclude that figure-ground segregation is not a stage of processing (Question 5) and to propose that it is simply one possible outcome of the interactions among cues to shape and depth. Other outcomes are possible, some of which will be discussed in this chapter. Conceiving of figure-ground segregation as an outcome rather than as a stage opens up new questions about how various visual properties interact to produce the perception of shape and distance. Some of these questions are raised at the end of the chapter.

TWO-DIMENSIONAL DISPLAYS

Object Memory Effects on Figural Status

My colleagues and I found that, contrary to the figure-ground first assumption, memories of the structure of known objects are accessed sufficiently early in perceptual processing to affect figure assignment (Peterson & Gibson, 1994a, 1994b; Peterson, Harvey, & Weidenbacher, 1991). These results were obtained in experiments using both two-dimensional and three-dimensional displays. The research conducted with two-dimensional displays will be discussed in this section; research using three-dimensional displays will be discussed in the third section.

In our stimuli, the edge between two regions sketches an identifiable portion of a known object along one side only. The side of the edge on which the portion of the known object is sketched is called the high-denotative side. The term high-denotative reflects the fact that a large percentage of pilot observers viewing the high-denotative regions agreed on which shape was portrayed along the high-denotative side; in other words, the high-denotative side of the border reliably elicited the same basic level shape perception in pilot observers. The opposite side of the border, called the low-denotative side, did not elicit high agreement about which object, if any, it portrayed. We found it impossible to create articulated regions that were meaningless to all observers, although we could create regions that elicited agreement from less than 25% of our pilot observers. We interpret the denotivity of a stimulus as an index of its goodness of fit to an object memory, with higher levels of denotivity indicating better matches.

Sample two-dimensional stimuli are shown in Fig. 3.2, parts A-E, where the high-denotative regions portray portions of a standing woman in white on the left and right sides of the black central shape (Fig. 3.2, part A), profiles of two identical faces in white (Fig. 3.2, part B), a portion of a seahorse in black (Fig. 3.2, part C), a portion of a table lamp in black (Fig. 3.2, part D), and a portion of a guitar in black (Fig. 3.2, part E). In many of the displays we used (e.g., Fig. 3.2, parts A-C), Gestalt configural cues (e.g., symmetry, smallness of relative area, enclosure)
or the monocular depth cue of interposition, or both, favored the interpretation that the low-denotative region was the figure. In other two-dimensional displays, the high- and low-denotative regions were equated for the Gestalt configurational cues (and the monocular depth cues were absent; e.g., Fig. 3.2, parts D and E). In this second type of display (but not the first), the only known cue to figural status that distinguished between the two regions was the object memory cue.

We presented our two-dimensional displays in both upright and inverted orientations and sometimes in orientations in between (e.g., Peterson & Gibson, 1994a, 1994b; Peterson et al., 1991). (The term upright refers to the orientation of the display in which the known object was portrayed in its typical upright orientation. Inverted displays were rotated 180° around the z axis.) We manipulated stimulus orientation because the object recognition literature had shown that it takes less time to identify objects when they are seen in their typical upright orientation than when they are rotated away from their upright (e.g., Jolicœur, 1988; Tarr & Pinker, 1989). These identification results could be taken to indicate that some critical threshold necessary for identification is reached earlier in time for upright displays than for displays misoriented from upright. (See Perrett, Oram, & Ashbridge, 1998, for a recent articulation of this view.)

We related these orientation-dependent identification effects to figure assignment as follows. It is generally accepted that figure assignment occurs quite early in the course of perceptual organization. Therefore, it seemed reasonable to argue that only those variables assessed within some critical time frame can affect figure assignment. Thus, if memories of object structure do affect figure assignment, their influence might be evident when displays with a high-denotative region are shown in an upright orientation but not when they are shown in an inverted orientation. For inverted displays, figural status might be determined before memories of object structure reach some critical threshold necessary to affect figure assignment. Changing the orientation of our displays from upright to inverted left unchanged the other cues that were present and known to affect figural status (e.g., overlap, symmetry, enclosure, smallness of relative area, convexity). Therefore, any increased likelihood of seeing the high-denotative regions of our displays as figures when the stimuli were upright versus inverted could be attributed to object memory cues that were present in the upright condition and diminished or absent in the inverted condition.

Consistent with the idea that memories of object structure can affect figure assignment (provided they are accessed quickly), we found that high-denotative regions were more likely to be seen as figures when the displays were upright rather than inverted. The same pattern of results was found both under brief masked exposure conditions in which observers reported which region first appeared to be the figure (Gibson & Peterson, 1994; Peterson & Gibson, 1994a) and under long exposure conditions in which observers reported perceived reversals of figure and ground (Peterson et al., 1991; Peterson & Gibson, 1994b). These results led us to propose that memories of object structure can be accessed sufficiently early in the course of perceptual organization to influence figure assignment. Inverting the displays slowed access to the critical object memories, thereby diminishing, or removing, their contributions to perceived organization.

On the basis of these experiments, we proposed that edges, detected early in the course of perceptual processing (and not just edges assigned to regions already partially or wholly determined to be figures) are the substrate for accessing memories of object structure. In our view, edge-based access to memories of object structure occurs in parallel with assessments of the configural cues and the depth cues, rather than afterwards (the parallel hypothesis). Our results can also be understood within an hierarchical interactive model in which regions (rather than edges) are matched to object memories following initial assessments of the configural cues (Vecera & O’Reilly, 1998, 2000). Presently, it is difficult to distinguish the hierarchical interactive model from our parallel model (Peterson, 1999b; Vecera & O’Reilly, 2000, but see Peterson, 2003). Both models account for our data indicating that object memories are accessed before figure assignment is complete and for converging evidence provided by others (e.g., Vecera & Farah, 1997). The critical difference is whether initial access to object memories is considered to be edge based or region based.

### Cue Competition

Our experiments conducted with two-dimensional displays indicated that the object memory cue does not necessarily dominate putatively lower level cues to figure status, such as interposition, symmetry, and smallness of relative area. Nor do low-level cues necessarily dominate the putatively higher level object memory cue. We investigated this question directly in experiments using the displays shown in Fig. 3.3 (Peterson & Gibson, 1994a). In each display, a high-denotative region and a low-denotative region shared a central edge. The horizontal symmetry (i.e., reflectional symmetry around a vertical axis) of the high- and low-denotative regions was manipulated orthogonally. The stimuli were exposed briefly (14 ms to 100 ms) and masked immediately. Observers reported whether the region lying on the left or the right side of the central edge appeared to have a definite shape by pressing a key to their left or their right. They were given an option to press neither key if both regions appeared to be shaped by the central border or if neither region did.

The data from an unpublished experiment (N = 24 participants; 100-ms stimulus exposure) are shown in Table 3.1. The results clearly indicate that both memories of object structure and the configural cues of symmetry affect perceived figural status. Averaging over symmetry conditions, high-denotative regions were more likely to be seen as figures when the displays were upright (74.5%) rather than inverted (58%). And, considering only inverted displays with one symmetric and one asymmetric region (regardless of denotivity), symmetric regions (72%) were
Low Denotative

High Denotative

Asymm

\[ \text{HD} = A \quad \text{LD} = A \]

Symm

\[ \text{HD} = S \quad \text{LD} = S \]

FIG. 3.3 The four types of stimuli used by Peterson and Gibson (1994a). "Most shape recognition follow figure-ground organization? An assumption in peril." Psychological Science, Blackwell Publishing Company. Reprinted with permission. In all stimuli, a high-denotative and a low-denotative region shared a central edge. In this figure, high-denotative regions are shown on the left in black, and low-denotative regions are shown on the right in white. In the experiment, the high-denotative regions appeared equally often on the left and the right and in black and in white. Also in this figure, the white regions are outlined in black on the top, bottom, and outer sides. This black outline is necessary to delimit the white regions of the stimuli from the white background. In the experiment, the black and white stimuli were shown on a gray background; no black outline surrounded the white regions. Symm = Symmetric around a vertical axis drawn through the center of the region. Asymm = Asymmetric.

more likely than adjacent asymmetric regions to be seen as figures. This percentage is consistent with recent estimates of the strength of the symmetry cue reported by others (e.g., Driver, Baylis, & Rafal, 1992).

The condition that permits an assessment of how the Gestalt cue of symmetry and the object memory cue fare when they are placed in conflict is the condition in which the low-denotative region was symmetric (LD\(_S\)) and the high-denotative region was asymmetric (HD\(_A\)). For inverted versions of these LD\(_S\)/HD\(_A\) displays, the high-denotative region was seen as the figure in only 29% of the trials (i.e., the symmetric low-denotative region was seen as the figure in 71% of the trials). This percentage reflects the strength of the symmetry cue when the object memory cue is absent or diminished by virtue of stimulus inversion. When competition from memories of object structure was present in the upright condition, however, the high-denotative regions were twice as likely to be seen as figures as they were in the inverted condition (59% vs. 29%). In other words, the percentage of figure reports consistent with the symmetry cue decreased markedly, to 41%. The comparison of performance obtained with upright and inverted displays revealed that memories of object structure do affect figural status when they are placed in competition with the Gestalt configural cue of symmetry.

Despite the substantial and significant increase in the likelihood of seeing HD\(_A\) regions as figures in upright compared with inverted LD\(_S\)/HD\(_A\) displays, the high-denotative regions were not seen as figures in the upright displays much more often than half the time. This finding suggests that the object memory cue does not necessarily dominate the putatively lower level cue of symmetry; each cue seemed to determine figure assignment approximately half the time. Thus, it is not the case that the high-level cue necessarily dominates the low-level cues, as many others have assumed (e.g., Götzschaldt, 1926/1938; Köhler, 1929/1947), nor is it the case that the low-level cues necessarily dominate the high-level cue, as others have claimed (e.g., Pylyshyn, 1999). The data in Table 1, along with the previous observations published by Peterson and Gibson (1994a), suggest that the object memory cue is simply one more cue to figural status; it is not given substantially more weight than the configural cue of symmetry.

In the experiment just described and others like it, observers were given the option to report that both regions appeared shaped by the central border or that neither region did. Observers rarely used this third response option. This finding implies that one-sided contour (or edge) assignment occurred in these experiments. In other words, when a cue did not determine figural status, the region carrying that cue appeared to be shapeless near the central edge. Consider LD\(_S\)/HD\(_A\) displays: When symmetry did not determine figure assignment, the symmetric low-denotative region appeared shapeless near the edge it shared with the high-denotative region. (Hence, its symmetry could not be seen.) Likewise, when memories of object structure did not determine figural status, the high-denotative region appeared to be shapeless near the border it shared with the low-denotative figure. (Hence, it could not be identified.)

\[ \begin{array}{|c|c|c|}
\hline
\text{High Denotative} & \text{Low Denotative} \\
\hline
\text{ASYMM} & \text{ASYMM} & \text{SYMM} \\
\text{Upright} & 72\% & 59\% \\
\text{Inverted} & 64\% & 29\% \\
\hline
\text{SYMM} & \text{Upright} & 87\% & 80\% \\
\text{Inverted} & 72\% & 67\% \\
\hline
\end{array} \]

\textbf{Note.} ASYMM = asymmetric; SYMM = symmetric.
The local apparent shapelessness of a high-denotative region adjacent to a low-denotative figure can easily be accounted for by a theory in which object memories are accessed for figures and not for grounds. On a theory in which memories of object structure are accessed before figural status is determined, how can one account for the apparent shapelessness of a high-denotative region seen as a ground? And on either theory, how can one account for the apparent shapelessness of a symmetric region seen as a ground? The parallel interactive model of configural analyses (PIMOCA), described in the next section, explains apparent shapelessness by positing competition between configural cues on opposite sides of an edge. PIMOCA explains the apparent shapelessness of the relatively weakly cued side of an edge, regardless of its denotivity.

Accounting for the Apparent Shapelessness of Grounds in Two-Dimensional Displays

PIMOCA (Peterson, 2000; Peterson, de Gelder, Rapcsak, Gerhardstein, & Bachoud-Lévi, 2000) integrates the parallel hypothesis (Peterson, 1994b, 1999b; Peterson & Gibson, 1994a) with inhibitory cross-border connections, a feature commonly employed in interactive hierarchical models (see Kienker, Sejnowski, Hinton, & Schumacher, 1986; Vecera & O’Reilly, 1998). According to PIMOCA, illustrated in Fig. 3.4, configural cues on the same side of an edge cooperate, whereas configural cues on opposite sides of an edge compete.

In PIMOCA, memories of object structure are considered to be configural cues, as are the traditional Gestalt configural cues. This is because, for effects of object memories on figure assignment to be observed, the parts of known objects must be in the proper configuration, and the known object must be portrayed in its typical orientation. Effects of object memories on figure assignment were not found when high-denotative regions were reconfigured by rearranging the parts, a change that rendered them low in denotivity (Gibson & Peterson, 1994; Peterson, et al., 2000; Peterson, Gerhardstein, Mennemeier, & Rapcsak, 1998; Peterson et al., 1991). Showing observers the known object from which the reconfigured version was created and pointing out the correspondence between the parts did not restore the object memory effects on figure assignment (Peterson et al., 1991). Thus, it is the configuration itself that is important for the object memory effects, not higher level knowledge regarding the object. Furthermore, the memories of object structure relevant to figure assignment are not holistic (Peterson, 2002). Instead, the relevant remembered configurations appear to be spatially limited (i.e., smaller than the whole object).

According to PIMOCA, the perception of shape in two-dimensional displays consisting only of configural cues is only the perception of shape attributes, such as symmetry versus asymmetry, convexity versus concavity, closure, area, part configuration, familiarity, and so on. When these configural attributes are suppressed in such two-dimensional displays, shape cannot be seen. Thus, the cross-edge competition in PIMOCA accounts for a weakly cued region adjacent to a relatively strongly cued region being perceived as locally shapeless, regardless of whether it is high or low in denotivity. Processing of configural properties continues on the strongly cued side of the shared edge, and a shape is ultimately perceived there.

Tests of PIMOCA’s Predictions

We recently tested PIMOCA’s predictions regarding cross-edge inhibition using the stimuli shown in Fig. 3.5–3.7 (Peterson & Kim, 2001). On each trial, two shapes were shown sequentially—a black silhouette followed by a line drawing. All of the black silhouettes were novel shapes; hence, their edges were low in denotivity on the black side. As well, they were all symmetric, enclosed, and substantially smaller in area than the white screen on which they were displayed. The black silhouettes appeared centered on fixation, and a new one appeared on every trial. Thus, for all silhouettes a number of configural factors cued the black regions to be figures, as did the cues of expectation and fixation location (Peterson & Gibson, 1994b). Most (75%) of the silhouettes were low in denotivity (LD) along the white (W) side of their left and right borders as well as along the black (B) side. Samples of these $B_{LD}W_{LD}$ silhouettes are shown in Fig. 3.5.

The remaining 25% of the silhouettes were high in denotivity (HD) on the white side of their left and right borders. Samples of these $B_{LD}W_{HD}$ silhouettes are shown in Fig. 3.6. For the $B_{LD}W_{HD}$ silhouettes, one configural factor—memory for object structure—cued the white side of the black-white border as the figure. Because the black side of the border was strongly cued to be the figure by many configural factors (see previous listing), we expected observers to perceive the black regions as the shaped figures and the white regions as shapeless grounds in these $B_{LD}W_{HD}$.

FIG. 3.6. Four sample $B_{LD} W_{HD}$ silhouettes used by Peterson and Kim (2001). The white sides of the left and right borders of the black regions are high in denotivity. The object sketched in part on the white side is a hand in A, a bell in B, a leaf in C, and a faucet in D. B is from Visual Cognition, 8 (3/4/5), Peterson & Kim, pp. 329–348. Copyright 2001, reprinted with permission of Psychology Press, Ltd., Hove, UK.

FIG. 3.7. Sample line drawings used by Peterson and Kim (2001). A–D are line drawings of known objects; E–H are line drawings of novel objects. Line drawings of known objects were always preceded by $B_{LD} W_{LD}$ silhouettes like those in Fig. 3.5. Half of the line drawings of known objects were also preceded by $B_{LD} W_{LD}$ silhouettes like those in Fig. 5; these were control trials. The other half of the line drawings of known objects were preceded by $B_{LD} W_{HD}$ silhouettes in which the RD region, perceived as ground, sketched a portion of the same basic level object. Examples of these experimental trials would be Fig. 3.7, parts A and B, preceded by Fig. 3.6, parts A and B, respectively. B and F are from Visual Cognition, 8 (3/4/5), Peterson & Kim, pp. 329–348. Copyright 2001, reprinted with permission of Psychology Press, Ltd., Hove, UK. E–H are from Journal of Verbal Learning and Verbal Behavior, 23, Kroll & Potter, pp. 39–66. Copyright 1984. reprinted with permission from Academic Press.

silhouettes as well as in the $B_{LD} W_{LD}$ silhouettes. Indeed, that is what observers typically reported seeing when they looked at these displays within the context of the experiment (and for the brief durations used in the experiment). According to PIMOCA, the apparent shapelessness of the white regions near the border of the black figure is mediated in part by inhibition of object memories that were accessed on the relatively weakly cued side of the edge.

We tested for the proposed inhibition using a priming paradigm. Silhouettes like those in Fig. 3.5 or 3.6 were presented for 50 ms as primes before line drawings like those in Fig. 3.7. Participants categorized the line drawings as known or as novel objects by pressing one of two buttons as quickly and as accurately as they could. Line drawings were exposed until a response was made.

A different silhouette was shown before each of the line drawings. $B_{LD} W_{LD}$ silhouettes (Fig. 3.5) were shown before all of the line drawings of novel objects.
The novel line drawings and their paired silhouette primes were included only so the observers had to make a decision before responding to the real line drawings: none of the novel line drawings was preceded by a silhouette with a matching shape in either the figure or the ground. The responses to the novel line drawings will not be considered further in this chapter. B_{LD}W_{HD} silhouettes were also shown before half of the line drawings of known objects (these were control prime trials). B_{LD}W_{HD} silhouettes were shown before the remaining half of the line drawings of known objects. On these experimental prime trials, a portion of the same basic level object portrayed by the line drawing was sketched along the white side of the borders of the silhouette prime (cf. Fig. 6.2, parts 4 and B with Fig. 3.7, parts A and B). The borders of the silhouette and the line drawing were different (i.e., a different version of the same basic level object was portrayed by each) to increase the likelihood that any priming we observed would be mediated by memories of object structure rather than by edge descriptions alone.

We were primarily interested in the participants' latency to correctly categorize the known line drawings, depending on the nature of the preceding prime silhouette.

![Control-Experimental RTs](image)

**FIG. 3.8.** Data from Peterson and Kim (2001, Experiment 2). SOA = stimulus onset asynchrony between the silhouette and the line drawing. In this experiment, the silhouette was exposed for the entire SOA. At the shortest SOA, observers were significantly slower to correctly classify known objects in the experimental condition than in the control condition. This difference may reflect the inhibition of object memories accessed for regions adjacent to strongly cued figures (i.e., the black silhouettes).

3. Figures, Grounds, and Completion

... (control vs. experimental). If the memories of object structure matching the ground side of B_{LD}W_{HD} silhouettes were inhibited, as predicted by PIMOCA, then the latency to correctly classify line drawings of objects from the same basic level category should be longer on experimental prime trials than on control prime trials. Peterson and Kim (2001) obtained evidence for the predicted inhibition. As shown in Fig. 3.8, observers' mean latency to accurately categorize line drawings as known was longer on experimental than on control prime trials. This pattern was obtained when the stimulus onset asynchrony (SOA) between the silhouette prime and the line drawing target was short—50 ms (Exp. 2) or 83 ms (Exp. 1).

We took these results to reflect the inhibition of object memories matching the white side of the border of the B_{LD}W_{HD} experimental prime silhouettes. We investigated the longevity of the inhibition by testing for priming at longer SOA conditions. Our results suggested that the object memories were inhibited for a brief time only. These results have been replicated with masked primes and a different control condition (e.g., using control prime silhouettes that portrayed portions of objects from a different superordinate class than their paired line drawings; Skow Grant, Lampignano. Kim, & Peterson, 2002). These results are consistent with the proposal that memories of object structure matched by the relatively weakly cued side of a border are inhibited.

According to PIMOCA, cooperation between configural cues on the same side of an edge and competition between configural cues on opposite sides of an edge produces the perception of definite shape on one side and shapelessness on the other side—one-sided edge assignment. PIMOCA's predictions regarding inhibition extend to all configural cues on the relatively weakly cued side of an edge. Therefore, PIMOCA can account for the perceived shapelessness of symmetric regions, and convex regions, seen as grounds. Tests of whether these inhibitory effects extend to the other configural cues are currently underway.

Possible Neural Substrates of PIMOCA

... suppression arising from competitive mechanisms has been observed in ventral area V4 of the monkey (e.g., Jagadeesh, Chelazzi, Mishkin, & Desimone, 2001; Reynolds, Chelazzi, & Desimone, 1999). Consistent evidence in humans was obtained using functional magnetic resonance imaging (fMRI; Kastner, De Weerd, Desimone, & Ungerleider, 1998). The suppression in those experiments arose from competition between two objects present in the same receptive field rather than from competition between configural cues on opposite sides of an edge. However, other evidence indicates that area V4 might be critical for figure-ground perception. For instance, Pasupathy and Connor (1999) found that V4 neurons respond to local convexity, one of the Gestalt configural cues. Peterson (2003) proposed that the spatially limited configurations mediating the object memory effects on figure assignment might be coded in V4 as well. In addition, Kastner, de Weerd, and Ungerleider (2000) observed neural correlates of texture segregation in V4 and TEO. Hence, it would no be unreasonable to investigate whether evidence of
the competitive interactions between configural cues on opposite sides of a border can be observed in V4.

Apparent Shapelessness and Completion

In two-dimensional displays, where no other depth cues or transparency cues are present, the perception of shapelessness on one side of a border may just be relative depth perception. It is thought that the ground, lacking a border near the figure, appears to complete amodally behind the figure. Must shapeless regions adjacent to figures appear to continue behind the figure? The illustration in Fig. 3.9 shows that when transparency cues are added, portions of the region adjacent to the figure can separate into two surfaces—one completes in front of the figure (the transparent gray surface); the other completes amodally behind the figure (the white surface). Fig. 3.9 demonstrates that the two figural properties of being shaped by a border and being in front of adjacent regions at that border can be uncoupled; the figure appears to lie behind the gray transparency (although it does appear to lie in front of the white ground).

Another demonstration that the figural properties of shape and relative nearness can be uncoupled is discussed in the next section. In addition, question 4 (whether two-dimensional displays are good surrogates for three-dimensional displays) and question 5 (whether figure-ground segregation should be considered a stage of processing) will be considered in the next section.

THREE-DIMENSIONAL DISPLAYS

To investigate whether or not object memories affect the assignment of figural status in three-dimensional displays, Peterson and Gibson (1993) created black and white stereograms by adding binocular disparity to the two vertical edges of either the high-denotative or the low-denotative region of stimuli like those shown in Fig. 3.2, parts D and E. (Fig. 3.10 shows sample stereograms.)

These stereograms were presented on a large random dot background. Each display was shown in two conditions. In one condition, binocular disparity specified that the high-denotative region was in front of the projection plane and that the low-denotative region lay on the projection plane. This condition was called the cooperation condition because the object memory cue and the binocular disparity cue cooperate to specify that the high-denotative region is the figure. In the other condition, binocular disparity specified that the low-denotative region was in front of the projection plane and that the high-denotative region lay on the projection plane. This condition was called the competition condition because the object memory cue and the binocular disparity cue specify that the figure lies on opposite sides of the central edge.\footnote{Peterson and Gibson (1993) used the terms \textit{cooperative} and \textit{competitive}. Strictly, these terms apply only on the assumption that figural regions (i.e., shaped regions) are necessarily perceived nearer to the viewer than adjacent regions.}

The critical condition was the competition condition. We predicted that if the object memory cue affects the perceived organization of three-dimensional displays, then observers should be less likely to see the disparity-consistent organization in competitive stereograms than in cooperative stereograms. That is, observers should be less likely to see the low-denotative region as the shaped entity lying in front of the high-denotative region in competitive stereograms than they are to see the high-denotative region as the shaped entity lying in front of the low-denotative region in cooperative stereograms. If observers were equally likely to see the disparity-consistent interpretation in both cooperative and competitive stereograms, that would indicate that memories of object structure do not affect the perceived organization of three-dimensional displays.

We showed these displays to observers for long durations (30–40 s) and asked them to report whether the region lying on the right or the left of the central edge appeared to have a definite shape and to lie in front of the adjacent region at the central border. Observers indicated what they saw throughout the viewing period by pressing one of two horizontally oriented keys. As in previous experiments, they were given the option to press neither key if both regions appeared to have a definite shape, or if neither region appeared to be in front of the central edge.

Before each trial, observers were instructed to view the stereograms under one of two intentional sets (Peterson & Hochberg, 1983). On half of the trials, they were asked to try to perceive the region on the right of the central edge...
FIG. 3.10. Two sample stereograms used by Peterson and Gibson (1993). If fusion is achieved by diverging the eyes, the stereogram pair in the top of the figure is a cooperative pair, and the stereogram pair in the bottom of the figure is a competitive pair. In cooperative stereograms, binocular disparity and object memories both specify that the high-denotative region (shown in black here) is the figure (shaped and in front). In competitive stereograms, binocular disparity specifies that the low-denotative region (shown in white here) is in front, whereas object memories specify that the high-denotative region is the shaped entity; the shaped entity is typically perceived to lie in front of the adjacent region. Hence, the object memory cue competes with binocular disparity. The stereograms in Peterson and Gibson's (1993) experiments were shown on a random-dot background, only a portion of which is reproduced here.


Percent Disparity-Consistent Interpretation

FIG. 3.11. A schematic of the typical results obtained by Peterson and Gibson (1993) on trials on which observers tried to perceive the disparity-consistent interpretation of the two types of stereograms. Co-op = cooperation condition. Compet = competition condition. In the Co-op condition, the high-denotative region was the figure in the disparity-consistent interpretation. In the Compet condition, the low-denotative region was the figure in the disparity-consistent interpretation.

as satisfying the two criteria listed previously. On the other half of the trials, they were asked to try to perceive the region on the left of the central edge as satisfying those two criteria. (Across the four stimuli used in this experiment, the right-left location of the high- and low-denotative regions was balanced within and counterbalanced across observers.) The data were summarized as the mean durations that the region specified to be in front by binocular disparity (i.e., the disparity-consistent interpretation) was perceived as the shaped, near figure at the central edge on trials on which observers followed instructions to try to perceive that region as the figure (see Fig. 3.11).²

In cooperative stereograms, the high-denotative region was more likely than the low-denotative region to be perceived as the shaped occluding figure at the central

²Results were very similar when expressed as the mean duration of the first percept reported on
edge (i.e., the disparity-consistent interpretation). These results were expected regardless of whether disparity alone determined the perception of three-dimensional displays or whether object memories contributed as well. The contributions of these two cues cannot be separated unless one compares performance in the cooperation condition with performance in the competition condition.

Critically, in competitive stereograms, the low- and high-denotative regions were each perceived as the figure at the central edge approximately half of the time. In these stereograms, binocular disparity specified that the low-denotative region was the occcluding object at the central border, but memories of object structure specified that the high-denotative region was the shaped entity there. The results suggest that, in these particular three-dimensional displays, neither the binocular disparity cue nor the object memory cue was necessarily dominant in determining which of the two adjacent regions was the shaped, near region at the central edge.

In summary, these results showed that memories of object structure can affect figure assignment in three-dimensional displays, even when the depth cue of binocular disparity specifies that the edge shared by a high- and a low-denotative region should be assigned to the low-denotative region.

**Perceived Distance**

In another experiment using black and white stereograms, like those in Fig. 3.10, we asked observers to report the apparent distance between the projection plane and the black and white surfaces at each of their sides. That is, observers estimated the perceived distance between the white region and the projection plane near its outer edge and near the central edge; they also estimated the perceived distance between the black region and the projection plane near its outer edge and near the central edge. Magnitude estimations of perceived distance were elicited for each type of stereogram under each of the two intentional sets (i.e., "try to perceive the black region as the figure" and "try to perceive the white region as the figure"). These magnitude estimations can reveal whether or not the disparity cue is necessarily suppressed when observers see the disparity-inconsistent interpretation of competitive stereograms.

As can be seen in Fig. 3.12, part A, when observers saw the disparity-consistent interpretation of cooperative stereograms, they saw a sharp, deep depth discontinuity between the high- and low-denotative regions at the central edge. The high-denotative region appeared to be located at the same distance from the projection plane as a standard stereogram with the same disparity (magnitude = 100), whereas the low-denotative ground appeared to lie far behind the figure (and close to the projection plane) at both the central edge and at its outer edge.

When observers saw the disparity-inconsistent interpretation of the cooperative stereograms, however, they perceived much less depth in the displays (Fig. 3.12, part C). Both the high- and low-denotative regions appeared to lie close to the projection plane, and a much smaller difference in depth was seen at the central edge.
border. The magnitude estimations suggest that observers may have been suppressing the binocular disparity cue to follow the instructions to try to see the low-denotative region as the figure in the cooperative stereograms. If so, the small percentage of time they succeeded in seeing this interpretation (see Fig. 3.11) suggests that this strategy was not very successful. I next consider the distance estimations obtained for competitive stereograms to investigate whether disparity was always suppressed when observers saw the disparity-inconsistent interpretation of a three-dimensional display.

Consider first the magnitude estimations obtained when observers perceived the disparity-consistent interpretation of competitive stereograms (i.e., the low-denotative region appeared to be the shaped region and to lie in front of the shapeless high-denotative region). Observers' magnitude estimations indicated that, when the disparity consistent interpretation was perceived, the low-denotative region of competitive stereograms appeared to be approximately as far in front of the projection plane as did the high-denotative region of cooperative stereograms (cf. magnitude estimations for the perceived figures in Fig. 3.12, parts A and B). A shallower depth step was perceived between the high- and low-denotative regions near the central edge of the competitive stereograms, however. Indeed, as can be seen in Fig. 3.12B, the high-denotative ground of competitive stereograms appeared to slant in depth across its spatial extent. It seemed to lie farther from the projection plane near the central edge than near its outer edge, although it still appeared to lie behind the low-denotative region there. Reasons for the perception of this slant will be considered later. For now, I leave the issue of slant aside and continue to explore the question of whether binocular disparity is necessarily suppressed when the disparity-inconsistent interpretation of competitive stereograms was perceived.

The magnitude estimations obtained when the disparity-inconsistent interpretation of competitive stereograms was perceived are not consistent with the suppression hypothesis (Fig. 3.12, part D). When, contrary to the disparity cue, the high-denotative region was seen as the figure in competitive stereograms, the low-denotative region did not appear to lie on or near the projection plane, as would be expected were the disparity cue suppressed. Rather, the low-denotative region appeared to lie much farther in front of the projection plane than the high-denotative region did when it was perceived as the ground in competitive stereograms (cf. Fig. 12, parts C and D).

Thus, perception of the disparity-inconsistent interpretation of competitive stereograms does not necessarily entail the suppression of the binocular disparity signal. In contrast, recall that when the configurual cue of object memory does not determine figure assignment, it is suppressed (Peterson & Kim, 2001; Skow Grant et al., 2002). It appears that the interactions between binocular disparity and the configurual cues are different from the interactions among the configurual cues themselves. This finding may not be surprising inasmuch as configurual cues are shape cues, whereas binocular disparity is predominantly a depth cue. The object memory cues and binocular disparity are processed in separate neural pathways (i.e., dorsal and ventral pathways, respectively; Ungerleider & Mishkin, 1982; Young, 1995). The interactions within the ventral pathway that determine perceived shape (e.g., the interactions modeled by PIMOSCA) might be quite different from the interactions between the ventral and the dorsal pathways.

I return now to the fact that the high-denotative regions of competitive stereograms appeared to slant upwards in depth from their outer border toward the central border both when they were seen as figures and when they were seen as grounds (Fig. 3.12, parts D and B, respectively). Note that a much smaller degree of slant was perceived in low-denotative regions of cooperative stereograms (Fig. 3.12, parts A and C). This originally puzzling result (see Peterson & Gibson, 1993) ultimately led me to understand that the slant of the region specified to lie behind in the black and white stereograms was ambiguous. The ambiguity was an unintentional consequence of the manner in which we made the stereograms, described next and illustrated in Fig. 3.13.

To portray, say, a black region on the right side of the central edge as lying farther in front of the projection plane than a white region on the left side of the central edge, we shifted the location of the entire black region to the right in the left eye's
view relative to its location in the right eye’s view. (This is the standard method for creating crossed disparity.) Shifting a region lying on the right of the central edge farther to the right necessarily left a space between the black and white regions in the left eye’s view. In random-dot stereograms, this space is typically filled with random dots uncorrelated with those already filling the two regions. In the black and white stereograms used by Peterson and Gibson (1993), this space was filled with the lightness of the unshifted region (white, in the sample case discussed). As a consequence, the region specified to be behind in the black and white stereograms was wider in the left eye’s view than in the right eye’s view. A display with these properties specifies either a flat surface lying on the projection plane and extending behind the front surface or a surface contacting the projection plane at its outer edge and slanting upwards in depth toward the central edge. (For discussion, see Howard & Rogers, 1995.) Thus, the slant of the region specified to lie behind by disparity was ambiguous in both the cooperative and the competitive black and white stereograms. It is noteworthy that the slanted interpretation was fitted to the back region predominantly in competitive stereograms (see Fig. 3.12, parts B and D); it was not perceived in the cooperative stereograms (see Fig. 3.12, parts A and C).

Might the fact that our black and white stereograms ambiguously specified a slanting surface account for our finding that memories of object structure affected the perceived organization of competitive stereograms? The slant of the ground can be rendered unambiguously in random-dot stereograms. Gibson and I (Peterson & Gibson, 1993, Experiment 1) failed to find object memory effects on figure assignment using random-dot stereograms. However, I did not then and I do not now take those results as evidence that object memory effects are present only when three-dimensional displays are ambiguous, for a number of reasons.

First, Gibson and I predicted that memories of object structure would not affect figure assignment in random-dot stereograms because it takes longer to detect the edge between the high- and low-denotive regions in random-dot stereograms than in black and white stereograms. (Luminance edges can be detected early in processing, whereas disparity edges can be detected only after the solution to the correspondence problem has begun to emerge.) As a consequence, edge-based access to object memories might not occur quickly enough in random-dot stereograms to affect figure assignment (Peterson & Gibson, 1993).

Second, even if the central edge is detected sufficiently quickly in random-dot stereograms, the dots comprising the stereograms may have added too much noise to the edge to support good edge-based matches to object memories. Evidence consistent with this latter hypothesis was obtained in an experiment in which observers viewed 300-ms masked exposures of stereograms in which binocular disparity indicated that a high-denotive region lay in front of the projection plane and an adjacent low-denotive region lay on the projection plane. Their task was to identify the object portrayed by the high-denotive region. For random-dot stereograms, observers identified only 29% of the objects portrayed by high-denotive regions. In contrast, for black and white stereograms, they identified 75% of the objects portrayed by high-denotive regions (Peterson & Gibson, 1991, 1993). To the extent that identification responses under these conditions can be taken as an index of the goodness of the edge-based match to object memories, these data suggest that matches to object memories mediated by edges in random-dot stereograms are poor compared with matches mediated by edges in black and white stereograms.

Furthermore, as discussed in the next section, that shaped apertures can be perceived demonstrates that object memories affect the perceived organization of unambiguous three-dimensional displays. Therefore, although the ambiguous slant of the background of the black and white stereograms used by Peterson and Gibson (1993) may have affected their perceived organization, it does not appear to be a necessary condition for the observation of object memory effects on figure assignment in three-dimensional displays.

**Shaped Apertures**

Consider a hand-shaped aperture cut into a surface mounted on a pole outdoors, as illustrated in Fig. 3.14. When this display is viewed from nearby, the binocular disparity cues are strong; each eye has different information about distant regions seen through the aperture. Other depth cues including linear perspective, relative size, texture density gradient, and interposition specify the layout of the scene beyond the surface into which the aperture has been cut. In addition, a few edges in the scene are good candidates for completion behind the mounted surface. These are all cues that the edge of the hand-shaped aperture belongs to the mounted surface. However, configural cues (including enclosure, convexity, and memories of object structure) specify that a shaped entity lies on the aperture side of the edge. What is perceived? The mounted surface with the aperture cut in it appears to be the near surface. The edge shared by the aperture and the near surface is seen as one of the occluding edges of the near surface. Nevertheless, the aperture appears to have a definite shape—that of a hand.

That a hand-shaped aperture can be perceived in the real three-dimensional scene illustrated in Fig. 3.14 indicates that object memory effects do extend to three-dimensional displays (as do the effects of the Gestalt configural cues). Therefore, for testing the role of configural cues in perception, two-dimensional displays are good surrogates for three-dimensional displays. In the three-dimensional version of Fig. 3.14, the distance to the objects seen through the aperture is specified unambiguously; nevertheless, object memory effects on perception are clearly evident. Hochberg (1998) and Palmer (1999) have also discussed cases in which shaped apertures can be seen. The case of a keyhole discussed by Hochberg is particularly telling because in that case we have a name for a shape that has primarily been seen in aperture form.

Remarkably, in the shaped aperture example, the two properties of a figure’s border have been uncoupled (or at least, these properties have been assigned with differential strength to these two regions). The shaping property of the border has been assigned to the aperture, yet the occluding property of the border has been
assigned to the near surface. The demonstration that the two properties of a figure’s border can be uncoupled in three-dimensional displays lends further support to the idea that the interactions between depth cues and configural cues are different from the interactions among configural cues. The three-dimensional cues specify which of two adjacent regions is the near surface (and they may also specify shape), whereas the configural cues specify where a shaped entity lies relative to a border shared by two regions. It seems that these two properties of near and shaped are perceptually coupled only (1) when the two types of cues specify the same region as the shaped near surface or (2) when there is a null signal in one type of cue. Most two-dimensional displays used to investigate figure assignment meet Criterion (2): Configural cues are present, but depth cues tend to be absent. In such displays, the shaped entity appears to be the near entity. Accordingly, at first glance, two-dimensional displays may seem to be good surrogates for three-dimensional displays because, for three-dimensional displays, the near surface is likely to be the shaped surface. However, the shaped aperture case demonstrates that two-dimensional displays cannot stand in for three-dimensional displays for the purposes of investigating whether the shaping and occluding properties of a border are necessarily assigned to the same side. In other words, two-dimensional displays can sometimes but not always serve as surrogates for three-dimensional displays.

OLD AND NEW QUESTIONS

The five questions raised at the beginning of this chapter have been addressed, and partial answers have been provided to some of them. My colleagues and I have shown that memories of object structure affect figure assignment. Therefore, the traditional answer to the first question—Why can figures be identified if they are familiar, whereas grounds cannot?—is wrong. It is not the case that object memories are accessed for figures and not for grounds. Rather, the solution implemented by the visual system appears to be more complex. I discussed PIMOA, which

3Recall that in the black and white competitive stereograms tested by Peterson and Gibson (1993), the shaping and occluding properties of the figure border were perceptually coupled. Perhaps the ambiguous slant of the back surface allowed the coupling to be maintained. This hypothesis requires further investigation.

FIG. 3.14. A photograph of a real three-dimensional display in which a hand-shaped hole was cut into a surface and mounted on a pole outdoors. The shape of the hand is clearly perceived, despite (1) the hand appearing to be surface-free and (2) the edge it shares with the surface into which the hole was cut appearing to be an occluding border of that surface. Of course, this picture can only provide a twodimensional depiction of an actual three-dimensional display.
uses cross-edge inhibition to account in part for why regions adjacent to strongly cued figures cannot be identified and appear shapeless (Questions 1 and 2). PIMOCA is a model of the interactions between and among the configural cues, which constitute a subset of the cues that determine shape and occlusion. Additional research must investigate the form of the interactions among the configural cues and between the configural cues and the depth cues.

On the basis of examples discussed in this chapter, it is clear that regions adjacent to figures are not necessarily completed behind them (Question 3). When the region sharing an edge with a figure can be separated into two surfaces, a transparent overlay and an opaque ground, the transparent overlay completes in front of the figure, and the other surface appears to complete amodally behind the figure (Fig. 3.9). Under these conditions, scission has occurred for the region that is not shaped by the shared edge. The edge still appears to be an occluding contour of the figure. In the shaped aperture case, illustrated in Fig. 3.14, however, the edge shared by two regions appears to shape one region (the aperture) and to be the occluding contour of the region lying on the opposite side (and perhaps also to shape that region). The shaped aperture case demonstrates that two-dimensional displays can serve only a limited role as surrogates for three-dimensional displays.

The present chapter raises new questions, including the following: What is the ontological status of the shaped aperture? Is it a figure? It has a definite shape but it may not have a surface. It certainly does not appear to be a surface lying at the farthest distance visible through the aperture, nor does it appear to be a surface lying closer to the viewer than the surface into which the aperture has been cut. Physically, the shaped aperture has no surface, but can the visual system perceive a surfaceless shape? Or, in the absence of cues to transparency, does the visual system fit a transparent surface within the boundary of the shaped aperture? If so, then shaped apertures constitute a new case of amodal completion—a completed surface is present but is not perceived. Regardless of whether or not amodal completion occurs, the shaped aperture outcome is not a case of figure-ground segregation, which entails one-sided contour assignment. Instead, it is a special case of figure-ground segregation, the properties of which remain to be determined experimentally.

A second question raised by the shaped aperture case is whether the shaping property of figures has been assigned to one side only (the aperture side) or whether the near surface appears to be shaped by its occluding contours as well. The configural cues are not balanced in the shaped aperture example. Therefore, PIMOCA predicts that shape will not be seen on the side of the border across from the hand-shaped aperture. However, PIMOCA models the interactions among configural cues only. A critical aspect of the shaped aperture case is that depth cues are present, in addition to configural cues. Investigations of the extent to which the near surface appears to be shaped by the edge it shares with the shaped aperture will provide critical evidence regarding the interactions among configural cues and depth cues in determining both shape and relative depth.

In conclusion, on the basis of the research summarized here, it is clear that figure-ground segregation is not a stage of processing that simply provides the substrate for higher level visual processes. Rather it is one possible outcome of interactions among image segregation cues (including configural cues, depth cues, motion, texture, shading, etc.) cues that jointly determine perceived shape, perceived surface properties, and perceived relative depth. These attributes need not be coupled, although they often are, in particular when figure-ground segregation occurs.

REFERENCES


4

Visual Perceptual Organization: A Microgenetic Analysis

Ruth Kimchi
University of Haifa

INTRODUCTION

The visual world consciously perceived is very different from the retinal mosaic of intensities and colors that arises from external objects. We perceive an organized visual world consisting of discrete objects, such as people, houses, and trees, that are coherently arranged in space. Some internal processes of organization must be responsible for this achievement.

The Gestalt school of psychology was the first to study the problem of perceptual organization. According to the Gestaltists, organization is composed of grouping and segregation processes (Koffka, 1935; Kohler, 1929/1947). The well-known principles of grouping proposed by Max Wertheimer (1923/1955) identify certain stimulus factors that determine organization. These factors include proximity, similarity, good continuation, common fate, and closure.

Although the Gestalt work on perceptual organization has been widely accepted as identifying crucial phenomena of perception and their demonstrations of grouping appear in almost every textbook on perception, there has been, until the last decade or so, little theoretical and empirical emphasis on perceptual organization. This may be, in part, because perceptual organization is so deeply ingrained in visual experience that it is often hard to appreciate the difficulties involved in achieving it.