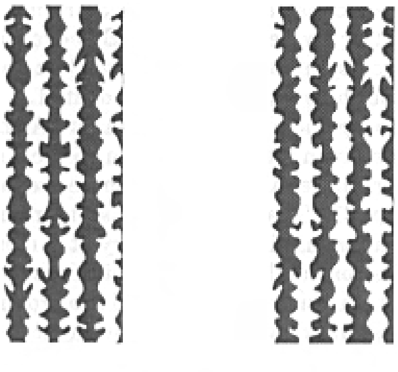

Figure 36.19. The Bahnsen columns. Although the black columns and the white columns have been equated for area and convexity, the symmetrical columns, black in (a) and white in (b), tend to be seen as figure. This is the most powerful demonstration that exists for the role of symmetry in perceptual organization. (From P. Bahnsen, Eine Untersuchung über Symmetrie und Asymmetrie bei visuellen Wahrnehmungen. *Zeitschrift fur Psychologie*, 1928, *108*. Reprinted with permission of Johann Ambrosius Barth.)

is so easily overcome indicates that it is less important an organizational principle than has been believed.

In Figure 36.22a, from Rock and Leaman (1963), symmetry fails to make its presence known as a factor in perceptual organization. This shape is often perceived as asymmetric even though it is symmetrical about two oblique axes. Figure 36.22b and c, modified from Attneave (1982), shows a related example: (b) tends to be seen as a nearly random distribution of eight dots; (c) shows symmetry *could* be perceived if the dots were interpreted as lying on an oblique plane in three-dimensional space. If the rules of organization placed much weight on discovering an organization that maximizes symmetry, then the symmetric interpretations of (a) and (b) might be easier to achieve than they are. If the perceptual system *is* searching for symmetry, it gives up rather quickly.

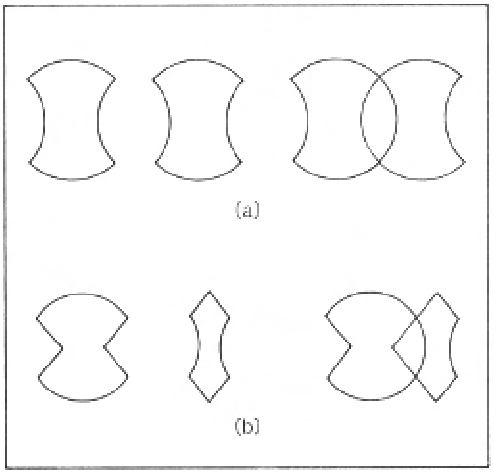

Figure 36.20. A further demonstration of the weakness of the Gestalt law of symmetry. In both (a) and (b) two symmetric figures are superimposed. In each case the composite is organized into two overlapping, asymmetric figures rather than as two abutted, symmetric ones, presumably because of the factors of good continuation and convexity. (From G. Kanizsa, *Organization in vision*. Copyright 1979 by Gaetano Kanizsa. Reprinted with permission of Praeger Publishers.)

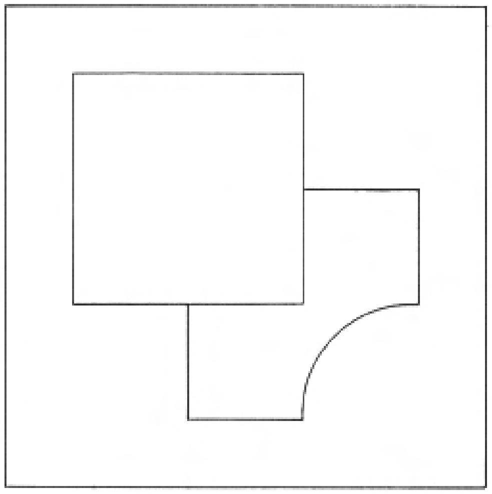

Figure 36.21. Kanizsa's amodal completion phenomenon. This stimulus is usually perceived as a square that is partially occluding a second shape. In addition, observers generally have biases about the shape of the occluded part of the figure; that is, the second figure is amodally completed. With this stimulus, subjects do not complete the second figure in a symmetrical fashion, although this would be possible and is predicted by the prägnanz principle. (From G. Kanizsa, *Organization in vision*. Copyright 1979 by Gaetano Kanizsa. Reprinted with permission of Praeger Publishers.)

Before we interpret symmetry in the light of prägnanz and likelihood, let us note that despite the demonstrations of Kanizsa and Attneave, symmetry is undeniably a salient property of forms and textures (Julesz, 1971; Palmer, 1982; Perkins, 1982; Royer, 1981; see also the discussion of pattern goodness in Section 7.4). Indeed, Palmer, Julesz, and others have formulated models to account for the ease with which symmetry is often detected. On the other hand, Uttal (1975) has demonstrated that symmetry has no role in the detection of dot patterns embedded in dynamic, random dot noise. In any case, the fact that perceivers are adept at detecting symmetry does not imply that symmetry has a major (or even any) role in determining perceptual organization. The same holds true of other properties like coplanarity, rectangularity, and familiarity. This is a subtle but important point that has not been adequately recognized in the literature, so let us elaborate.

Flicker is an extremely salient property of visual stimuli, so much that it is frequently used as an attention-capturing device in visual information displays. Also, flicker can affect the appearance of spatial patterns (see Watson, Chapter 6). Yet to our knowledge, flicker has never been proposed as an organizational factor in vision. It may well be an important organizing factor; but we have chosen it because it has not been proposed as such despite its great perceptual salience. Perhaps flickering elements tend to be grouped into perceptual wholes, but that cannot be taken for granted without empirical evidence demonstrating that it is true and delineating any additional constraints that must be met for such grouping to occur (e.g., common flicker frequencies, phase relationships, or duty cycles). This argument parallels one made by the Gestalt psychologists (and others) about the role of learning in perception. For example, Kanizsa & Gerbino (1982) note that, "the importance of past experience in the interpretation of perceptual data does not legitimize its use as an explicative principle regarding the processes at the basis of the formation of the data themselves" (p. 187).

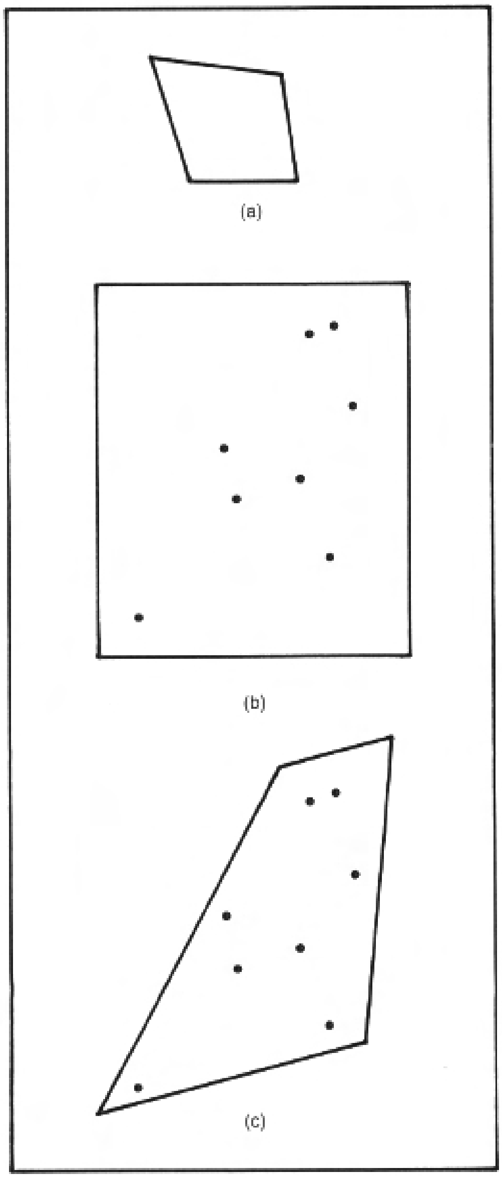

Figure 36.22. Symmetry and orientation. (a) A shape whose symmetry often goes undetected. (b) Dots, usually perceived as a nearly random scattering lying on a surface oriented perpendicular to the line of sight. The scattering could, however, be interpreted as symmetric (about one axis) but viewed at a slant, as indicated in (c) with the help of a reference frame. The latter organization is difficult to achieve in (b), indicating that percepts are not organized spatially to maximize symmetry. (From F. Attneave, Prägnanz and soap bubble systems: A theoretical exploration. In J. Beck (Ed.), *Organization and representation in perception*. Lawrence Erlbaum Associates, 1982. Reprinted with permission.)

The demonstrations of Kanizsa and of Attneave appear to indicate that symmetry will be a detectable property of visual configurations only if symmetry exists in the perceptual units carved out on the basis of other factors besides symmetry. Thus symmetry (like familiarity) may become important only *after* the perceptual organization of the stimulus has been determined.

It is important to note how different this view of symmetry is from other interpretations, such as the soap-bubble concept that symmetry is actively imposed on perceptual organizations, or the notion that the perceptual system actively hunts for a veridical description of the distal stimulus pattern that includes symmetry. The demonstrations just cited are not favorable to either of the latter two explanations. But even if a bias did exist that favored symmetric over asymmetric descriptions of the stimulus, that bias could be accommodated at least as well by the likelihood principle as by prägnanz. Attneave (1982) noted that, "for a given retinal pattern, a possible distal source that contains certain regularities . . . is more probable than one that does not" (p. 21). That is, if one possible interpretation of a proximal stimulus manifests symmetry, it is most likely this interpretation is correct, because it is highly improbable that an asymmetric distal stimulus would lead to a proximal stimulus that could be organized symmetrically without distortion.

Kanizsa (1979) has expressed concern that the effects of symmetry (and of geometric regularity in general) are so easily overridden by other factors in determining perceptual grouping (see Figures 36.20, 36.21, and 36.22). Indeed, the very demonstrations most often summoned to illustrate the law of symmetry are frequently marred by an inadvertent confounding of the effects of convexity or area with those of symmetry; the result has been an exaggeration of symmetry's apparent role in organization. This bias led Kanizsa to conclude, "Although it may be doubtful that there is a tendency to maximum regularity in our perceptual system, it is probable that a 'tendency toward the selection of a regular figure' exists on the part of gestalt psychologists when a perceptual law is being illustrated" (p. 110).

Kanizsa's skepticism is justified. The demonstrable vulnerability of symmetry to domination by other factors forces us to question whether symmetry is as important to organization as traditionally has been believed. However, it does not force us to dismiss symmetry altogether: the Bahnsen columns of Figure 36.19 show symmetry determining perceptual organization when other confounding factors (such as area or convexity) are eliminated. Perhaps symmetry tips the scales of organization only when other, more potent factors are completely balanced and so favor none of the competing organizations.

In any event, the Bahnsen columns *do* show an effect of symmetry in its own right, and so the prägnanz versus likelihood explanations of this phenomenon must be contrasted. The prägnanz explanation is quite straightforward because symmetry is the quintessence of simplicity: having specified one component of a symmetric figure, the remaining components need not be reiterated. The asymmetric columns in Bahnsen's figures need not be represented in perception. The symmetric figure occludes the asymmetric ground, and, as Rubin argued, the ground has no contours or shape of its own but merely continues behind the figure.

Although symmetrical figures abound in the natural environment (as with snowflakes, leaves, faces, etc.), so do asymmetric ones (uneven terrain, jagged rocks, winding streams, etc.), and many natural scenes lack any symmetry whatsoever. An ecological survey would help to resolve the likelihood question, but it is probable that asymmetric objects are more frequently encountered in the environment than symmetric ones. In that case, the prägnanz principle might appear to have prevailed over likelihood in best accounting for symmetry. However as we have noted, if a symmetrical interpretation can be found for a stimulus, it is likely to be a correct interpretation, and so in this fashion the likelihood principle can accommodate the effects of symmetry on perceptual organization (which, as we have seen, are not particularly strong) as well as can prägnanz.

We return to the effects of symmetry in this chapter in the context of information theory and coding theory, which attempt

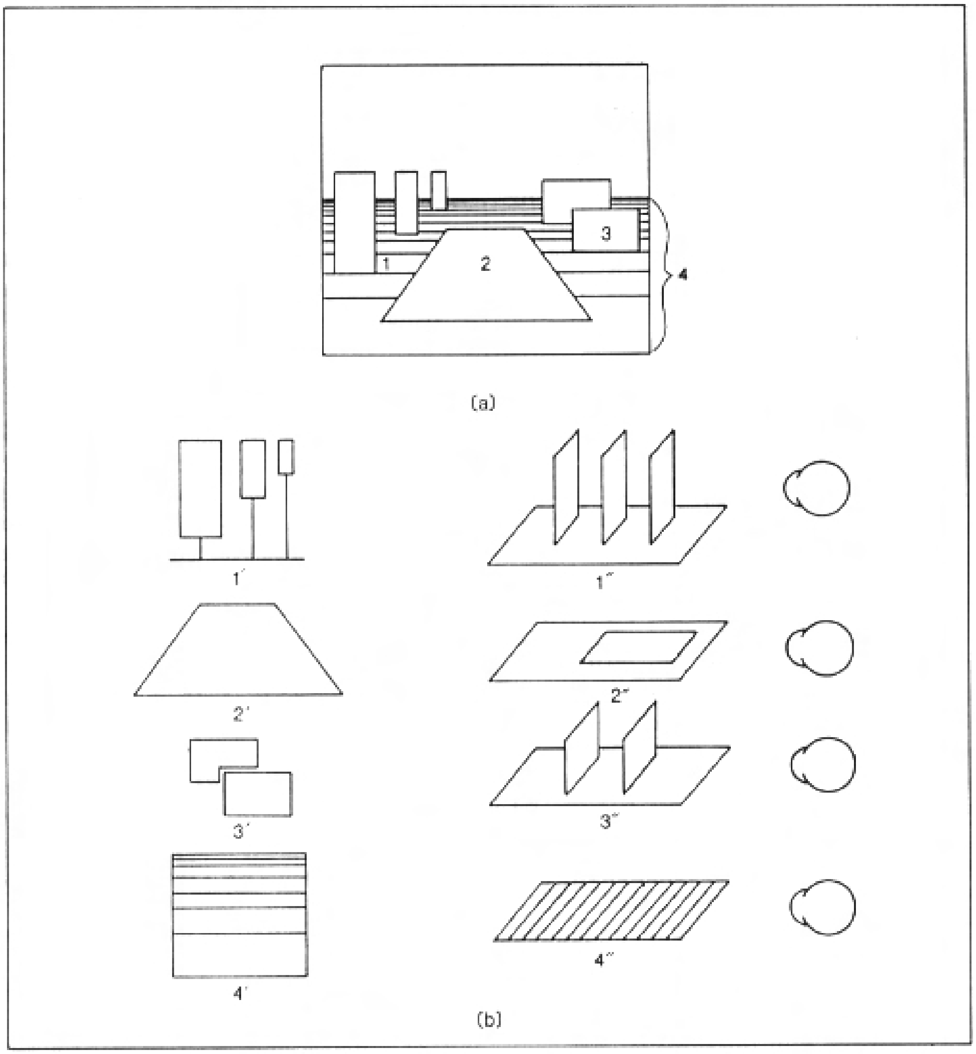

Figure 36.23. Depth cues and prägnanz. (a) A scene containing four monocular depth cues: relative size, linear perspective, interposition (occlusion), and texture-density gradients. (b) Isolation of the four cues, in the left column as viewed in (a) and in the right column as viewed from the side. As the text explains, it is not clear that the depthful interpretation of the scene in any way simplifies the scene. (From J. Hochberg, *Perception* (2nd Ed.). Prentice-Hall, Inc., 1978. Reprinted with permission.)

to deal with symmetries and other forms of regularity (or redundancy) in stimuli in a more formal fashion.

4.7. Assessment

The preceding analyses have focused on the Gestalt principles of area, proximity, similarity, closure, good continuation, convexity, and symmetry. Other principles of some importance not discussed include orientation (Koffka, 1935/1963, p. 190), enclosure or surroundedness (Rubin, 1921), color and contrast (Rock, 1975, p. 261), and common fate (Wertheimer, 1923/1950). Although these and other laws could be analyzed in detail, such an exercise would add little to the picture that has emerged. A rough balance sheet or scorecard (cf. that of Gregory, 1974) shows the following: the likelihood principle provides the better explanation for the effects of closure, area, proximity, and similarity. Both symmetry and good continuation seem to be handled equally well by either. The same is probably true for Rubin's principle regarding the one-sided function of contour, according to which any given contour (a line or edge) is perceived as belonging only to the figure, not to the ground. The law explains why it is difficult if not impossible to perceive the faces and the vase simultaneously in Rubin's famous figure-ground demonstration.

5. ADDITIONAL PERCEPTUAL PHENOMENA TO BE EXPLAINED

Having assessed the various Gestalt laws of organization, let us turn to a series of specific perceptual problems and phenomena that have occupied center stage in theoretical analyses of perceptual organization for many years. These include the perception of space and depth, the Necker cube, impossible figures, subjective contours, and apparent motion.

5.1. Depth Cues

Let us start this analysis with Figure 36.23, drawn from Hochberg (1978, pp. 141–143). Figure 36.23a shows a scene that most observers see as a landscape containing one rectangle lying on the ground and five rectangles standing upright, perpendicular to the receding ground plane. At least four monocular depth cues are contained in this figure: (1) relative size, wherein the proximally larger rectangles are perceived as nearer and not necessarily as larger; (2) linear perspective, wherein distally parallel lines are depicted as converging toward a vanishing point beyond the visible horizon; (3) interposition, wherein nearer surfaces can occlude our view of more distant ones; and (4) texture-density gradients, wherein the parallel, horizontal lines depicting the ground become more dense at greater distances. (Additional cues are contained in this picture, some of which indicate depth, such as height in picture plane, and some of which contradict depth, such as absence of binocular disparity or motion parallax, but we can ignore these factors for the moment.) The figures in the left-hand column of Figure 36.23b depict frontal views of portions of this scene that illustrate these cues; those in the right-hand column show side views depicting the usual three-dimensional interpretation of the scene.

Hochberg (1978) has outlined the Gestalt interpretation of these depth cues as further instances of the prägnanz principle. Considering the topmost pair of figures in Figure 36.23b, is it not simpler to perceive three identical rectangles at three different distances (right column) than three different-sized rectangles at the same distance? Considering the linear perspective cue with the second row of figures, is it not simpler to perceive a square lying flat on the ground than a trapezoid standing upright (cf. the Ames window)? In the third row, is it not simpler to see one complete rectangle partially occluded by another than to see a rectangle abutted in the same depth plane with an irregular, L-shaped figure? Finally, in the bottom row, is not a regularly spaced set of lines receding into the distance simpler than an irregularly spaced set that does not recede? Although these questions are phrased rhetorically, their answers are not at all obvious, for each implies an unstated cost-benefit analysis which must demonstrate that the three-dimensional interpretation is in fact the simpler one (Pomerantz & Kubovy, 1981). Considering the relative size cue as an example, *why* are three identical figures at different depths simpler than three different figures at the same depth? Such a claim entails additional assumptions about the processes and structures used to encode stimuli; the claim does not follow directly from the idea of prägnanz.

The likelihood explanation for these depth cues follows a now-predictable pattern. Concerning relative size, for example, rarely in the natural environment do objects line up in exactly the same depth plane; a typical scene contains numerous objects, most of which are at different depths. However, it is common to see arrays of nearly identically sized objects, as with a field of dandelions or a herd of cattle. In the absence of contradictory information, the perceptual system might assume that several objects of similar shape have the same size. Similar arguments can be made for the other depth cues, but in each case, an ecological survey would be required to settle the matter.

To clarify further the contrast between the two principles, recall the Ames room, shown in Figure 36.24. The room's distal shape is trapezoidal, but when seen monocularly from the proper observation point, it is perceived as rectilinear. When objects of known, familiar size (such as people) are placed within the room, the room is still perceived as rectilinear, and the objects' apparent sizes are therefore distorted. To explain this outcome, prägnanz would hold that a rectangular room is simpler than a trapezoidal one, whereas likelihood would hold that the former is more likely than the latter. Concerning size distortions, prägnanz would claim that altering the sizes of objects does not affect their simplicity; the patterns remain the same even though the elements change size. Some loss in simplicity is incurred by making the people's sizes so variable, but that might be offset partially by making their perceived distances uniform (cf. Figure 36.23) and is certainly compensated for by the perceived regularity of the room. The likelihood principle, on the other hand, must rationalize the size disparity by appealing to the great size variations that exist among human beings, ranging from giants to dwarfs. In sum, each approach does a presentable, albeit not totally convincing, job of handling the Ames room. The tacks taken by the two are similar in that both explain an illusion by appealing to processes that normally foster veridical perception. But no obvious method suggests itself for distinguishing the two empirically.

Hochberg (1978, p. 139) has shown that at least two depth cues cannot be handled by prägnanz. These are *illumination direction*, discussed previously, and *familiar size*. Illumination from above is no simpler than from below; it is only more likely. Similarly, familiar size is by definition dependent on learning, and, although it is in a sense simpler to see objects at their proper sizes, this is not the sense intended by prägnanz. Further, Gregory (1972, p. 183) points to the cue of aerial perspective, in which distant objects appear more blue and less distinct than nearer ones. This effect is due to atmospheric haze, which scatters light and absorbs its constituent wavelengths differentially. It would be farfetched to link this cue with simplicity; more likely this cue is acquired (either phylogenetically or ontogenetically) through experience with the environment.

Returning to our balance sheet, we see again that neither prägnanz nor likelihood has won a decisive victory. Let us next probe some domains in which differences between the two principles may be most apparent, beginning with the perception of the Necker cube and then examining impossible figures, subjective contours, and apparent motion.

5.2. The Necker Cube

This stimulus, discussed earlier and shown in Figure 36.2a, has been at the center of debates about perceptual organization for decades. Why is this figure normally seen as a cube in depth rather than as a flat design? Why does it tend to reverse spontaneously in depth as we observe it, and why is one of the two three-dimensional orientations (where the cube is seen from above) normally preferred to the other? As usual, prägnanz holds that a three-dimensional organization is simpler, whereas likelihood holds that it is more likely. (The relevant arguments have all been presented before.) Multistability is explained (at least in part) by both principles by claiming that both three-dimensional interpretations are viable and so neither should dominate the other. This notion, when combined with neurophysiological account of multistability (in terms of Köhler's satiation processes or Attneave's multistable electronic circuits), has considerable appeal (but see Rock, Chapter 33); and, although the details of such processes need much additional elaboration, neither principle is favored here. The preference for one perceived orientation over the other seems more readily explained by likelihood than by prägnanz; solid objects are viewed from above more often than from below, but it is not clear why the former would be any simpler than the latter.

Hochberg and McAlister (1953) contrasted the two Kopfermann (1930) patterns shown in Figure 36.25 (taken from

Hochberg, 1978, p. 139) and discovered that the traditional Necker cube in Figure 36.25*a* is more likely to be perceived in depth than is the pattern in (b), even though each is an equally correct depiction of a wire cube. The effect was explained by Hochberg and McAlister as follows (see also Hochberg & Brooks, 1960, discussed further later): Part (a) is seen in depth and part (b) as flat because each is simpler when interpreted that way. They devised an objective metric for simplicity based on weighted counts of the number of interior angles, the number of different angles, and the number of continuous lines. Basically, (a) is seen as a cube in depth because all the angles of a cube are right angles; if it were seen as flat, the figure would contain a number of odd angles. On the other hand, (b) is seen as flat because when seen that way all of its angles are identical (60°); further, seeing (b) in depth would entail splitting all three of the internal, continuous line segments into two parts, thereby breaking good continuation and making the figure more complicated than it would be if it were left flat.

The preceding analysis is clearly Gestalt-like, but it goes beyond vague appeals to prägnanz by explaining how simplicity might be measured, at least for drawings of cubes. A likelihood account for the effects in figure 36.25 would note that for the figure in (b) to be a cube it would have to be viewed from exactly the one angle that would make objectively separate contours become continuous in the proximal stimulus. It is more likely, then, that the figure is a flat design. This line of reasoning is not as successful in explaining why the pattern in (a) is seen

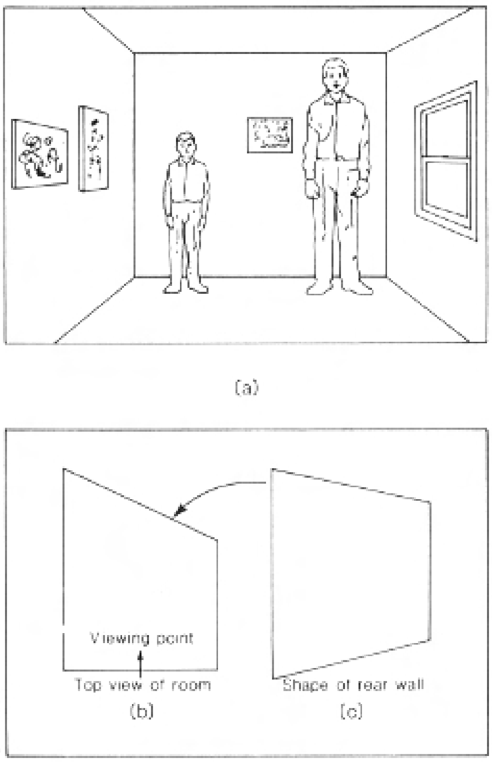

Figure 36.24. The Ames distorted room. Adelbert Ames designed this room which, although it contains no rectangular surfaces, appears normal and "square" to an observer peering in from a properly placed observation hole. (a) What the eye (or camera) would see from this station point. (b and c) Clarification of the the actual layout of the room. Observers perceive the room as square and the two people as stationed at about the same distance but differing greatly in size; in fact, the two people are about the same size but are stationed at different distances. This effect can be explained by either the prägnanz or the likelihood principle. [(a) From I. Rock, *An introduction to perception*. Macmillan Publishing Co., 1975. Reprinted with permission.]

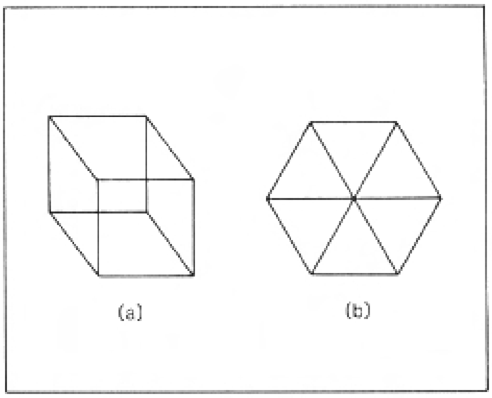

Figure 36.25. Perspective drawings of two Kopfermann cubes. Both of these patterns are equally good perspective drawings of a wire cube, but the design (a Necker cube) in (a) is most likely to be perceived as a cube, whereas that in (b) is most likely to be perceived as a flat design. Hochberg & McAlister (1953) provided a quantitative model to explain why.

in depth, but many of the arguments given before (e.g., that trapezoids, or in this case parallelograms, are likely to represent rectangles) could be applied.

Kanizsa (1979, p. 106) has provided a companion demonstration to the Figure 36.25, which is shown in Figure 36.26 and which places important contraints on either explanation. As with Figure 36.25, (a) tends to be seen in depth whereas (b) is seen as flat. But seeing (a) in depth does not make it simpler than seeing it flat; it is an irregular, asymmetric figure in either case. Similarly, it is not clear that (b) is simpler when seen as flat than if it were seen in depth. [Note that (b) can also be seen in depth as an irregular pyramid viewed from straight above, but it seems neither simpler nor more complicated this way.] Kanizsa's point is that the apparent improvement in regularity seen with the cubes in Figure 36.25 may be only an accidental by-product of other principles that govern perceptual organization. In this case most of the results in both Figures 36.25 and 36.26 may be due to good continuation, which (as we have seen) is in turn explained equally well by likelihood as by prägnanz.

5.3. Impossible Figures

Impossible figures have been treated as interesting curiosities of perception, but their implications for perceptual organization may be profound. Figure 36.27 presents two well-known examples, the three-stick clevis and the impossible triangle. In each instance, our perception tends to organize the figure into a three-dimensional object that could not be realized physically. We tend not to see them as flat designs drawn on paper, although this interpretation would eliminate the problem of impossibility. Figure 36.28 reduces the paradox to perhaps its simplest case. In Figure 36.28a most often we see two rectangles, one of which occludes our view of the other. In (b) we see a multistable organization of two overlapping rectangles containing conflicting depth cues. At times, the left rectangle appears as a nearer figure that occludes our view of the right rectangle, while at other times the converse organization prevails. Following a few such alternations of perceived depth, the conflict becomes apparent to the observer. But despite the conscious awareness of the difficulty of interpreting the figure in depth, the stimulus resists a flat interpretation. Note that this stimulus need not be seen as an impossible figure because the rectangles can be interpreted as possessing horizontal slits that allow them to be

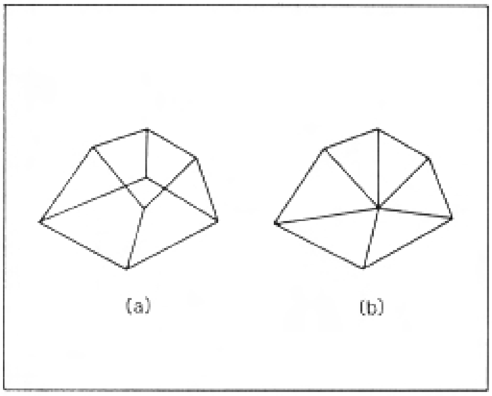

Figure 36.26. Kanizsa's distorted cubes. The two patterns presented here parallel the two in Figure 36.25, except that the object depicted is a distorted cube. Nevertheless, the perceptual effects are the same as in Figure 36.25, with (a) being perceived in depth and (b) as flat. This demonstration indicates that depthful interpretations are not necessarily imposed on stimuli to increase their geometric regularity. (From G. Kanizsa, *Organization in vision*. Copyright 1979 by Gaetano Kanizsa. Reprinted with permission of Praeger Publishers.)

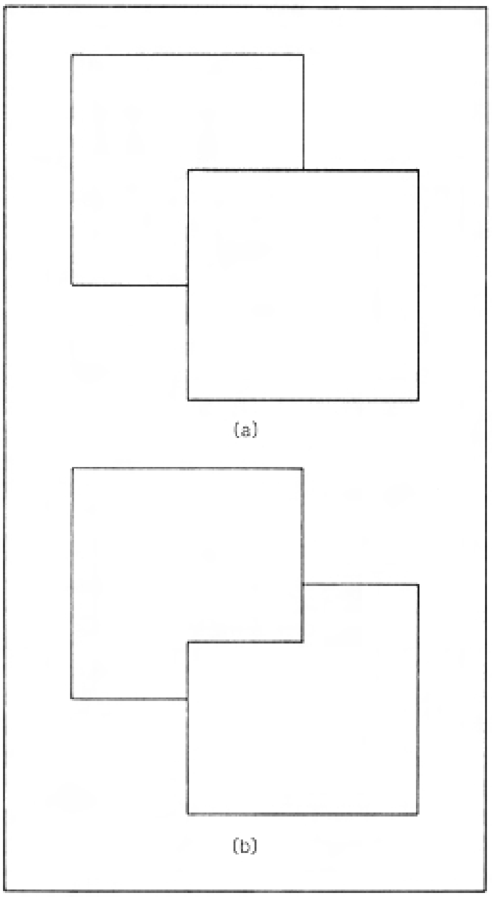

Figure 36.28. Interposition and perceived depth. The stimulus in (a) is most often perceived as one square in front of another. The same is true in (b), but which square is in front of the other is difficult to determine, and the whole figure may appear impossible. For some observers, the display spontaneously alternates in depth like a Necker cube, even though the impossibility would be eliminated if it were perceived as a flat design or as two sheets with slits that have been slipped through one another.

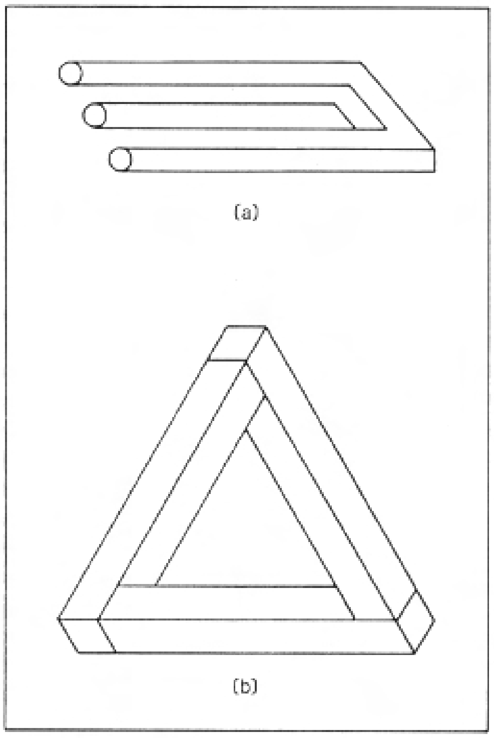

Figure 36.27. The three-stick clevis and the impossible triangle. These are perhaps the best-known of the "impossible figures." Rather than interpreting these patterns as flat designs, the visual system attempts to structure them in depth, which turns out to be an impossible task. These figures present a challenge to all theories of perceptual organization. [(a) From D. H. Schuster, A new ambiguous figure: A three stick clevis. *American Journal of Psychology, 77*. Copyright 1964 by University of Illinois Press. Reprinted with permission. (b) From L. Penrose & R. Penrose, Impossible objects: A special type of visual illusion. *British Journal of Psychology*, 1958, *49*. Reprinted with permission.]

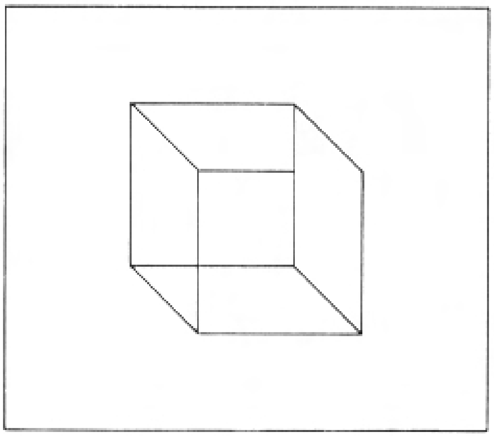

Figure 36.29. Hochberg's incomplete cube. This is a standard Necker cube except for one of the horizontal segments, which has been truncated at the point where it intersects a vertical line. Although this figure has a number of possible interpretations (such as a Necker cube with one opaque wall), many observers perceive it as an impossible figure. (From J. Hochberg, *Perception* (2nd ed.), Prentice-Hall, Inc., 1978. Reprinted with permission.)

slid through one another. But this "possible" interpretation eludes many observers, who experience only the alternating, impossible percepts.

Impossible figures are not easily accommodated by either prägnanz or likelihood because they seem neither simple nor likely. We should note that impossibility per se is unrelated to simplicity, because certain physically impossible structures (such as n-dimensional spaces and Klein bottles) are nonetheless fairly easy to describe mathematically. However, this observation fails to provide much redemption for prägnanz in explaining Figures 36.27 and 36.28; in any geometry these figures remain an enigma. Concerning likelihood, impossible figures clearly lie at the lower limit of the probability scale because they never occur as objects in the environment (Gregory, 1974; Pomerantz & Kubovy, 1981).

Impossible figures thus present a challenge for all existing theories of perceptual organization. One potential resolution of this dilemma would be to dismiss them as anomalies, special cases that could be handled by ad hoc mechanisms, so they would have no impact on general principles of perceptual organization. But a novel demonstration by Hochberg (1981a), shown in Figure 36.29, makes even this route untenable. The figure, which is a minor modification of the Necker cube, is a perfectly "possible" perspective drawing of a wire cube in which either (1) one of the wire edges has been clipped at its midpoint so it stops in midair, or (2) one of the surfaces of the cube is opaque, so the truncated edge is occluded from view. But despite these two "possible" interpretations, observers frequently see the figure as impossible, that is, as an object that could not exist in three-dimensional space. Were either a prägnanz or a likelihood principle operating, the "possible" interpretation should dominate. Yet many observers report spontaneous and in fact indifferent perceptual alternations between the possible and the impossible interpretations. Figure 36.30 (Pomerantz & Kubovy, 1981) shows a variation of this figure in which the "possible" interpretation is perhaps even more likely and plausible than the possible one of Figure 36.29. Even though this figure can easily be seen as an opaque triptych, most observers see it equally often as three perpendicular, rectangular surfaces twisted in an impossible manner.

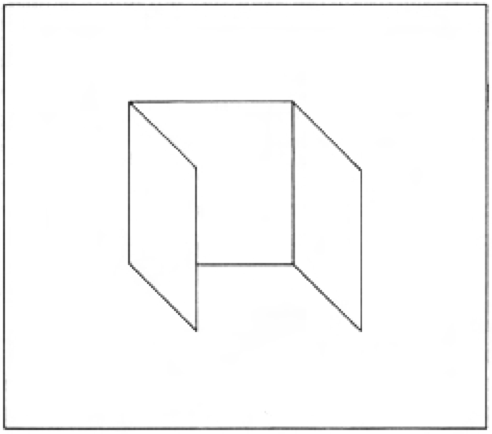

Figure 36.30. The ambiguous triptych. As with Figure 36.29, this modification of Hochberg's incomplete cube is often perceived as impossible even though it has a possible (and in fact quite likely) interpretation: a triptych, that is, a vertically oriented screen folded vertically into three sections. (From J. R. Pomerantz & M. Kubovy, Perceptual organization: An overview. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*. Lawrence Erlbaum Associates, 1981. Reprinted with permission.)

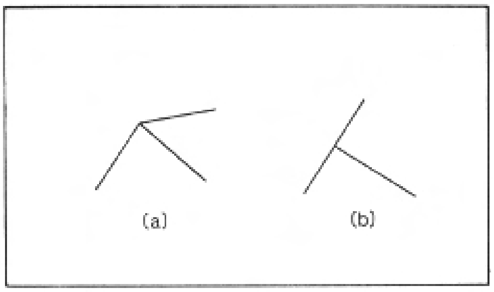

Figure 36.31. Vertices contain structural information. As Guzmán (1968) showed, information relevant to global structure can be gleaned from local regions such as vertices. The intersection in (a) can be interpreted as the corner of a cube, whereas that in (b) suggests depth by means of an occlusion cue (where one line terminates at the point of intersection with another). The principles of prägnanz and likelihood can be applied to local regions as well as to whole objects, although this shift in the level of analysis is not in the spirit of Gestalt theory.

How can these demonstrations be reconciled with either prägnanz or likelihood? As Hochberg (1981a) notes, excluding them as special cases would be gratuitous because both are derived from the same Necker cube that has figured so prominently in the perceptual organization literature. A more forthright reconciliation (alluded to in Section 3.3.3) requires shifting the *unit of analysis* upon which the prägnanz and likelihood principles operate. The literature on perceptual organization has neglected this crucial matter by assuming tacitly that simplicity or likelihood applies to entire stimuli (objects or perhaps entire scenes). But either principle could be applied to smaller units of analysis, such as to the local intersection of contours. Consider the intersections shown in Figure 36.31. In Figure 36.31a the intersection of three line segments can be interpreted as a flat design or as the corner of a cube; in Figure 36.31b the lines can be interpreted as a flat T-shaped configuration or as one contour passing behind an occluding edge (cf. Guzmán, 1968). Without reiterating all the arguments, either a prägnanz or a likelihood explanation could be formulated for perceiving these two intersections in depth. If the global organizations perceived for Figures 36.27–36.30 were determined at the level of local intersections, the dilemma of impossible figures might disappear: considered piecemeal, each narrowly circumscribed region of these figures is organized in a simple, likely fashion.

Thus we can argue that impossible figures are explained satisfactorily by shifting the focus of organizational principles to local regions. This shift might be justified by constraints in the neural underpinnings of perceptual organization, in which interactions are limited to local regions (cf. the soap bubble analogy), or by arguing that organization occurs only within a limited region surrounding our momentary focus of attention (usually the fovea; cf. Hochberg, 1968, 1978). In either case there would be no process that oversees global organization to ensure that the local organizations are compatible with one another. The fact that percepts typically appear to be globally organized and free of internal contradictions, and that impossible figures are only rare curiosities, would then imply that when the perceptual system takes care of all the small details, the "big picture" will take care of itself. Perhaps a highly selective set of rules has evolved to achieve *local* organization that (except in rare, artificially contrived cases such as impossible figures) also gives us *global* organization.

5.4. Subjective Contours

This illusion, first demonstrated by Schumann (1904), involves the perception of contours in regions of homogeneous luminance. Several examples are shown in Figure 36.32. Figure 36.32a reproduces Schumann's original figure; (b) shows what is probably the most common demonstration of the effect, from Kanizsa (1955). In the left half of (b) the perception of a complete triangle that is whiter than the figure's background is especially compelling. Figure 36.32c (Kanizsa, 1955) shows a compelling subjective edge that has a lustrous appearance; in Figure 36.32d (Coren, 1972) most observers perceive the word "FEET," printed in block letters.

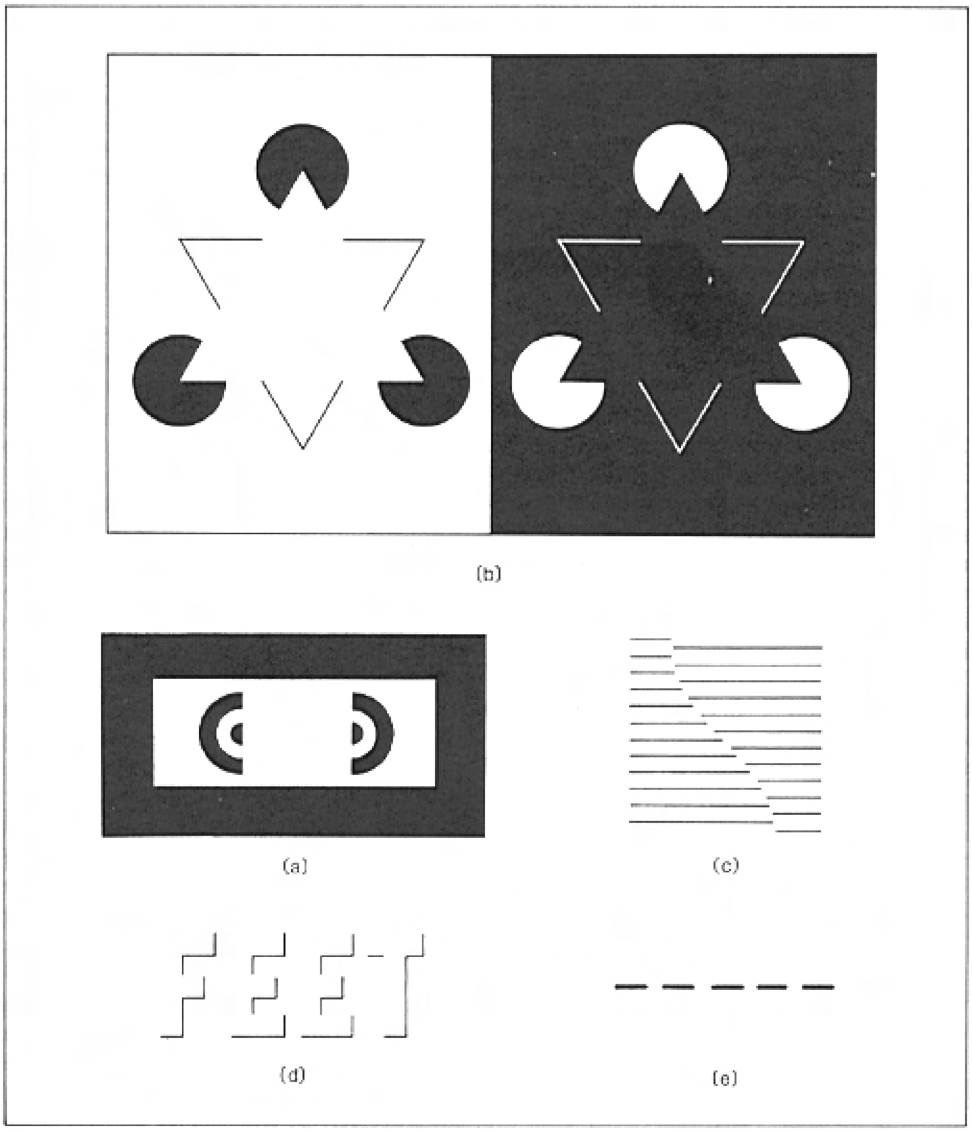

Figure 36.32. Subjective contours. (a) Schumann's (1904) original demonstration of subjective contours; a central square, slightly whiter than the background, can be seen, especially if the stimulus is not scrutinized too closely. The contour of the square seems to be visible throughout its perimeter, although there is no contrast to support the contour for much of its perimeter. (b) Two of Kanizsa's (1955) subjective triangles in which the illusion is stronger than in (a). The subjective triangle appears whiter than the background on the left figure and darker than the background on the right. (c) A subjective contour without differences in brightness due to contrast; the line, however, does have a lustrous appearance. (d) The word FEET spelled using subjective contours. Here, however, the contours do not result from occlusion cues, as in (a–c), but instead arise from cues indicating the presence of unseen objects casting shadows. (e) A simple dashed line. The lines are not filled in perceptually across the gaps, which indicates that subjective contours are not merely a matter of closure or of "filling in." [(a) From F. Schumann, 1904. (b) From G. Kanizsa, Margini quasi-percettivi in campi von stimolazione omogenea. *Revista di psicologia*, 1955, 49.] (c) From G. Kanizsa, *Organization in vision*. Copyright 1979 by Gaetano Kanizsa. Reprinted with permission of Praeger Publishers. (d) From S. Coren, Subjective contours and apparent depth. *Psychological Review*, 79. Copyright 1972 by American Psychological Association. Reprinted with permission.]

Although all of these cases involve the appearance of illusory contours, it is an open question whether they all result from the same underlying mechanism. Coren (1972) has argued that the illusion is due to perceived depth, with the subjective figures in Figure 36.32a and b appearing to lie in front of, and so occluding, the objective inducing areas. Depth is also involved in Figure 36.32d; the physically present contours correspond to the shadows that would result if raised letters (of the same lightness as the background) were illuminated from the upper left.

Subjective contours might be thought of as an instance of the Gestalt law of closure, but this would be incorrect (Coren, 1972). Consider the dashed line shown in Figure 36.32e. Dashed lines are often perceived as equivalent to solid lines, suggesting the operation of closure, but to be consistent with the previous examples of subjective contours, one would have to perceive subjective contours running perpendicular to the dashed line, as though a solid line drawn against a white background were perceived through a white picket fence. (In fact with a little effort, this organization can be achieved.)

To what extent can subjective contours be explained by prägnanz? To be sure, Figure 36.32b is more simply described as three complete black disks and one complete outline triangle partially occluded by a solid white triangle, than as three notched disks and three V's that happen to be oriented with their component edges and line segments collinear. However, Figure 36.33 casts doubt on the prägnanz explanation. Figure 36.33a and b shows irregular, complex figures generated from subjective contours. Figure 36.33c shows that the Penrose and Penrose impossible triangle can be approximated reasonably well with subjective contours (Kanizsa, 1979, p. 218; Pomerantz & Kubovy, 1981, p. 447, Panels c and d). In short, subjective contours can be perceived regardless of whether they simplify the resulting global organization of the distal stimulus.

The likelihood principle can explain Figure 36.32b, since the various occlusion cues strongly suggest the presence of an opaque triangle (Gregory, 1974). The perception of the word "FEET" in Figure 36.32d also is commensurate with a likelihood explanation. Banks and Coffin (1974) questioned whether subjective contours would be perceived by observers unfamiliar with these alphabetic characters; indeed, few observers shown only the leftmost inducing contours of this panel will perceive the letter "F." The complex configurations in Figure 36.33 can be explained either by the operation of local processes that are insensitive to the global figures that emerge from the construction of subjective contours, or by noting that irregular figures are no less probable in the natural environment than more regular ones.

Numerous aspects of the subjective contour problem remain to be explained in detail by either approach. (Some of these are discussed by Ginsburg, Chapter 34.) But the prägnanz approach in particular is greatly damaged by Figure 36.33. The only apparent salvation would require applying prägnanz at the local level (as discussed previously), interpreting terminated line segments more simply as indicating occlusion than as indicating termination without occlusion. But, as we see in the dashed line of Figure 36.32e, terminations do not always lead to subjective contours. Similarly, if one views only a single notched disk from Figure 36.32b, no subjective contours are seen, or will the letter "F" in (d) be seen as such if presented in isolation. Subjective contours are apparently due to *global* organization, not to purely local factors, and so an explanation appealing to a local minimum principle is untenable.

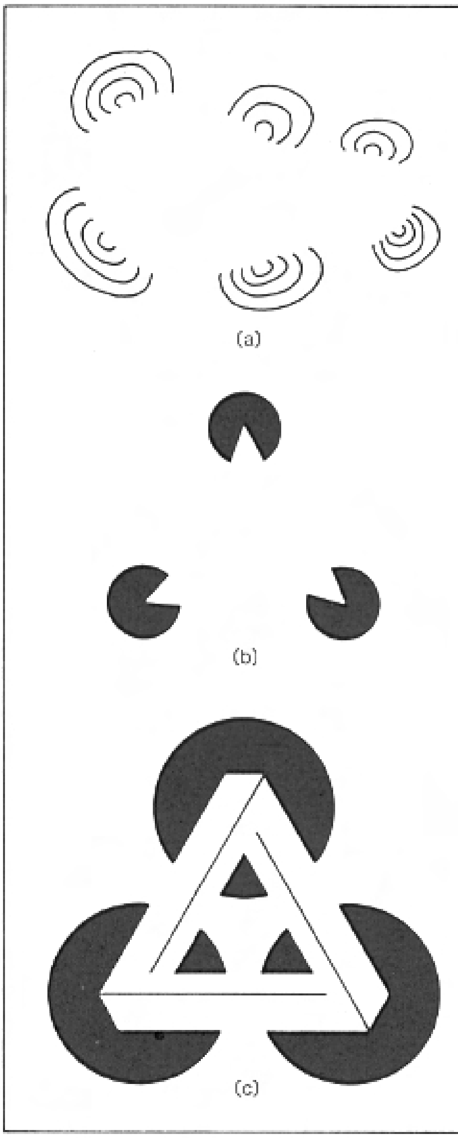

Figure 36.33. Irregular shape and impossible figures produced using subjective contours. (a and b) A demonstration that the subjective figure need not be a good Gestalt for subjective contours to emerge. (c) The Penrose and Penrose (1958) impossible triangle from Figure 36.27 can be approximated reasonably well from subjective contours. These figures all present difficulties for either the global prägnanz or the likelihood principle. [From G. Kanizsa, *Organization in vision*. Copyright 1979 by Gaetano Kanizsa. Reprinted with permission of Praeger Publishers. (c) From J. R. Pomerantz & M. Kubovy, Perceptual organization: An overview. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*. Lawrence Erlbaum Associates, 1981. Reprinted with permission.]

It might also be argued that the perception of subjective contours is mediated by highly specific, ad hoc detector mechanisms unrelated to either prägnanz or likelihood. These detectors might be triggered automatically by specific features in the proximal stimulus, such as the collinearity of edges and of terminated line segments. Figure 36.34, from Bradley, Dumais, and Petry (1976; see also Bradley & Petry, 1977), suggests otherwise. If this figure is viewed as a Necker cube floating in front of eight solid black disks, one sees the cube being completed by subjective contours; but if the figure is viewed as a Necker cube floating behind a white surface containing eight circular apertures, no subjective contours are seen. The figure thus

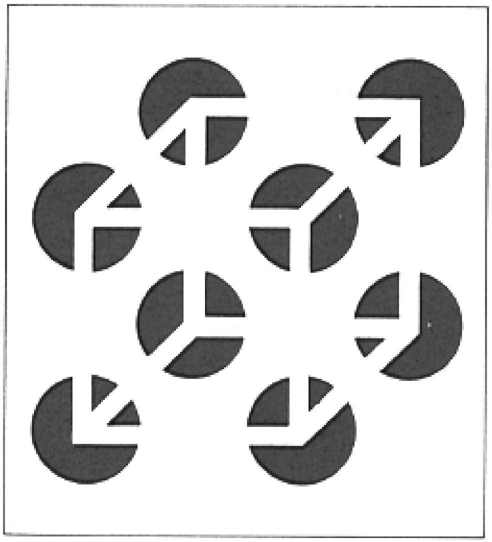

Figure 36.34. The subjective Necker cube. This figure provides a powerful demonstration that subjective contours are not the result of some automatic perceptual process. This figure can be viewed as a white Necker cube floating above a set of eight black disks. If the figure-ground organization is reversed, however, it can be seen as a white Necker cube against a black background as seen through a white screen containing eight round holes. In the former case, subjective contours are seen; in the latter, they are not. (From D. R. Bradley, S. T. Dumais, & H. M. Petry, Reply to Cavonius. *Nature, 261*. Copyright 1976 by Macmillan Journals, Ltd. Reprinted with permission.)

demonstrates that subjective contours are not perceived by the automatic activation of specific detector mechanisms; rather they result from an active ("top-down") and globally oriented organizational process (or schema; see Hochberg, 1978, p. 193).

In summary, the likelihood principle appears to be the only viable contender for explaining subjective contours. The explanation is not complete at this time, however.

5.5. Apparent Motion

This phenomenon, first noted by Exner (1875), formed the cornerstone for Wertheimer's (1912/1961) original notions that launched the Gestalt movement. Several different and perhaps unrelated types of apparent motion exist. Wertheimer's main concern was with the "phi" phenomenon, or objectless motion. If two separated lights are flashed on and off at the proper alternation rate, the observer experiences a strong sense of motion between them, even though no light is seen to move (cf. "beta" motion, in which a moving object is seen traversing the space between the two terminal positions). Wertheimer argued that with phi motion the observer was experiencing "pure" motion (pure in that only motion and no moving object is perceived); thus motion could not be a sensation derived in a structuralist fashion from the combination of the separate sensations produced by the inducing lights. The phi phenomenon seemed to lend itself well to an explanation involving dynamic brain fields (Köhler, 1923; Wertheimer, 1912/1961): each of the two lights produces a specific and local disturbance in the brain, and field effects operating automatically produce a smooth flow of current between them, much in the same manner as when a single light is perceived undergoing real motion.

Can apparent motion be considered a *simple* organization of the distal stimulus event? In the case of beta motion this might be true because one perceives only a single, continuously visible, moving object rather than two flickering, stationary objects. In addition, the trajectory followed by the apparently moving object is often the simplest (straightest, shortest, or most smoothly curved) interpolation between the terminal positions (Foster, 1978); in situations where two or more objects are put into apparent motion simultaneously, the matching (or correspondence) between the multiple objects in the two flashes is often the simplest or most economical one possible (Kolers, 1972; Pantle & Picciano, 1976; Ternus, 1926/1950; Ullman, 1979).

However, two complications arise. First, the perception of beta motion depends on the timing of the stimulus flashes being set within certain limits. If the alternation rate is too fast, two stationary, flickering lights are seen simultaneously; if it is too slow, the lights are again seen as stationary and as appearing and disappearing in succession (Kolers, 1972). Similarly, the kind of motion seen in the Ternus situation just noted is dependent on timing. It is not clear why beta motion would be a simple organization at some alternation rates but not at others. By contrast, it is possible that the temporal bounds on apparent motion reflect some learned or evolved knowledge of the temporal parameters of real motion (Shepard, 1981). Korte's (1915) well-known "laws," which summarize (roughly) the spatial and temporal constraints of apparent motion, could be viewed as the internal representation of these learned boundaries, an interpretation consistent with the likelihood principle. Kaufman, Cyrulnik, Kaplowitz, Melnick, and Stoff (1971) showed that the velocity range over which apparent motion is perceived does not overlap with the range for real motion; apparent motion is perceived only at velocities that are above the upper threshold for seeing real motion. Kaufman (1974, p. 401) has suggested that apparent motion arises from stimulation of detectors whose function is to respond to real motion that occurs at velocities too great for ordinary, real-motion detectors. According to this reasoning, motion (rather than succession or simultaneity) is the organization of the stimulus event that is most likely to be correct.

The second complication is that beta motion can be perceived just as easily between two alternating stimuli that differ in shape as between two identical shapes (Kolers & Pomerantz, 1971; Orlansky, 1940). In the former case, the observer usually perceives the two objects undergoing a plastic deformation such that the moving object transforms its shape smoothly as it moves between the two termini. Again, it is not clear why a single object changing in shape is any simpler an organization than the veridical perception of two different objects flickering in place. Nor for that matter is it clear why the former is more likely than the latter. Many objects change their shape as they move (as with a bird in flight), but often they do not (as with a gliding bird or a rigid object); just which is the more likely state of affairs is anyone's guess, so the likelihood principle neither wins nor loses on this score. However, it is relatively rare that objects abruptly appear out of nowhere (or disappear abruptly), although this does occur occasionally in nature, as with fireflies at nighttime. More likely, an image that disappears from one location and reappears at another signifies a moving object that either has passed behind an occluding surface or has been interrupted by a saccadic eye movement or a blink. Thus apparent motion detectors may have evolved as a mechanism for preserving the perceptual continuity of objects continuously present but not continuously represented in the proximal stimulus.

A likelihood explanation for the standard alternating-lights demonstration of apparent motion might predict that the observer would perceive subjective contours forming an opaque subjective figure occluding the space between the two terminal locations. However, subjective contours are not seen in this situation (although they are seen in other cases where further dynamic cues to occlusion are present). Furthermore, if a solid, opaque object is placed in the path of apparent motion, one might expect the observer to perceive the moving object to pass *behind* the occluding surface rather than in front of it, since the former provides a likely explanation of the disappearance of the object from the proximal stimulus. In fact, however, these two alternatives are perceived with approximately equal frequency (Pomerantz & Kubovy, 1981). Although facts such as these do not present insurmountable obstacles for a likelihood explanation, they do force an element of arbitrariness in any eventual model by requiring it to explain our sensitivity to some aspects of the stimulus event but an insensitivity to other aspects that would seem to be of equal importance.

In sum, our conclusions for apparent motion parallel those for subjective contours. The prägnanz explanation is weak; simplicity does not seem to be relevant to our perception of apparent motion. By contrast, the likelihood approach seems more promising. The stimulus information associated with this phenomenon is a likely indicator of an object moving under conditions that cannot be sensed by detectors of real motion because the velocity of motion is too great, because the moving object was occluded, or for some other reason. Clearly we need more information about the ecological optics of visual motion, but this approach seems far more promising than one based on prägnanz.

6. ASSESSMENT AND RECONCILIATION OF PRÄGNANZ AND LIKELIHOOD

At this point we have considered an extensive sample of the perceptual phenomena on which claims about perceptual organization have been based. Our discussion has necessrily excluded many important effects, especially those in modalities other than vision, and we have focused on qualitative aspects of these phenomena at the expense of quantitative details and models. However, we are now in a position to evaluate the evidence and ask whether there are any common features to be found within our sample of phenomena that could serve as a foundation for a general theory of perceptual organization.

6.1. Gestalt Psychology and Prägnanz

The evidence favoring the prägnanz principle is somewhat thin. Whereas the dominant view of perceptual organization held by psychologists in general is most often centered around the Gestalt approach (as witnessed by most textbook treatments of perception), the Gestalt explanations of Gestalt phenomena are often inadequate, vague, or simply wrong. The various Gestalt laws of organization are on fairly safe ground when considered as descriptions of organizational tendencies, although even here they often fail to give us predictions of everyday perception or (in the absence of a formal model) to tell us when one law will prevail over another. The law of symmetry, which can be regarded as the keystone of prägnanz, lacks the kind of clean and robust demonstration one should expect of so preeminent a principle. It is clear that symmetry can be detected in visual patterns, but so can countless other physical properties; it remains to be demonstrated, however, that symmetry either is actively imposed upon our organizations in soap bubble fashion or is sought after in an extensive search of alternative organizations. To be sure, certain effects described previously lend themselves to a Gestalt explanation, but in many such cases alternative explanations appear to do at least as well.

6.2. The Helmholtzean View and Likelihood

The Helmholtzean approach, with a few exceptions, has provided a more promising and potentially testable explanatory framework for a number of organizational phenomena. To be sure, these phenomena can hardly be said to be *explained* by the likelihood principle unless a specific mechanism can be put forth to model how the phenomenon occurs. In addition, in many cases we lack the necessary ecological survey to inform us about critical information in the proximal stimulus. Despite these limitations, the likelihood approach seems far more promising than the prägnanz approach, which failed to make predictions of any sort for many important phenomena.

Certain robust effects, such as familiar size and illumination direction (Hochberg, 1978), are completely unrelated to simplicity and so demand an account based on adaptations to likely correspondences between proximal and distal stimuli. These adaptations could be based either on ontogeny (i.e., on learning during the lifetime of the organism) or phylogeny (i.e., on evolutionary trends). Most of the Gestalt laws can be regarded as indicators (of varying validity) of what distal stimulus is most likely to have given rise to the proximal stimulus. The major drawback to this approach is the practical difficulty of performing the ecological surveys required to substantiate the claim that our organizational rules capitalize on properties of proximal stimuli that mirror the most likely properties of distal stimuli.

The primary requirement for any plausible theory of perceptual organization is to account for the veridicality of perception. Schemes based on prägnanz must recognize that our distal environment shows only a weak bias toward geometric regularity. A perceiver who, for the sake of economical encoding of the environment, imposes regularity where it does not exist would suffer a distorted, Pollyanna-like representation of the world that would be counterproductive to survival (Attneave, 1982). The likelihood principle, by contrast, is better suited for representing the world veridically.

The likelihood approach is not free of difficulties, however. One problem lies in explaining why certain likely organizations do not occur, as in the case of apparent motion along a path blocked by an opaque object. A second problem concerns the perception of impossible figures, such as the Penrose triangle. Gregory (1974) has maintained that paradoxical percepts pose no special problem for the Helmholtzean view: perceptions, by this approach, are only hypotheses, and "Hypotheses, or descriptions, can be logically impossible. Paradoxes may be generated by following false assumptions, which turn out to be incompatible" (p. 277).

This tack seems overly forgiving, for why should a system founded on likely hypotheses generate ones that are impossible? The only way to justify this stand would be to shift the scope of the perceptual unit to more microscopic levels, where the perceptual system works independently on more local aspects of the total stimulus pattern to arrive at likely hypotheses (Hochberg, 1981a). This would allow sets of simultaneous hypotheses, each of which is highly probable, to contain internal contradictions. When such contradictions are tolerated, it indicates that no executive process or homunculus is supervising the assembly of local hypotheses into a unitary, global representation free of contradictions.

Our assessment leans toward the likelihood principle as the best overall description of how perceptual organization operates. However, an acceptance of likelihood does not mandate a rejection of prägnanz, because the two are not mutually exclusive. As we have seen, certain phenomena support prägnanz, so a compromise could be sought that captures the essentials of both positions without diluting the positive contributions of either or endangering the parsimony of any definitive framework for perceptual organization. Indeed, Mach (1906/1959) was certain that, "The visual sense acts therefore in conformity with the principle of economy, and, at the same time, in conformity with the principle of probability . . ." (p. 215).

6.3. Economical Coding

Information theory (see Attneave, 1959; Garner, 1962) and structural information theory (also known as coding theory; Leeuwenberg, 1971, 1978, 1982; Restle, 1982) may provide a basis for such a compromise. Consider Attneave's (1954, 1981) concept of *economical coding*:

> Suppose that what the system *likes* is short descriptions and that the image is progressively changed, within the constraints of the input, until its description is minimized. This way of looking at the matter, which is considerably different from the classical Gestalt point of view, has the advantage of taking into account not only intrinsic stimulus properties—that is, redundancy, uniformity, or homogeneity in the stimulus itself—but also schemata corresponding to familiar objects. If an input can be brought into conformity with a well-formed schema that is frequently used and to which a short symbol has been assigned, it might be described quite as economically as if it were intrinsically simple (Attneave, 1981, pp. 417–418).

Attneave's observation points to two concepts that are the primary focus of the balance of this chapter: information theory as applied to human perception, and structural information theory, which attempts to devise schemes for representing patterns in a fashion that reflects their perceptual organization. Structural information theory is discussed at length in Section 8. For the moment, let us focus upon information theory.

Information theory allows both prägnanz and likelihood to be translated into the common framework of brief codes. In information theory the information value of a stimulus is inversely related to its predictability, and the lower the information value, the shorter the description or encoding of the stimulus can be. The predictability (or redundancy) present in symmetrical patterns can be exploited in a number of ways (some of which are described later) to produce briefer descriptions. Less obvious is the fact that frequently occurring stimuli are also redundant in that they are more predictable than rare stimuli; this redundancy can be exploited in a similar fashion as well. For example, the most frequently occurring words in English tend to be the shortest words (e.g., a, the, of, is; this principle is known as Zipf's law). Similarly, in electronic communications, more frequently occurring characters (such as the letter E) are best transmitted using fewer bits than rarer characters (such as Q), as in Huffman coding. If this approach is correct, then certain potential conflicts between likelihood and prägnanz can be avoided, and any scheme of perceptual organization that exploits redundancy in the stimulus may be accommodated.

7. INFORMATION THEORY AND PERCEPTUAL ORGANIZATION

Information theory (IT) provides a formal model for measuring the amount of information in a signal or stimulus. The theory was developed outside the field of psychology (in communication engineering) and is completely neutral with respect to psychological issues. In fact, Garner's (1962) book, *Uncertainty and Structure as Psychological Concepts*, avoided using the term "information theory" whenever possible to prevent confusion between IT as a mathematical theory and IT as a descriptive model of human information processing. Thus a caveat is in order for those who would apply IT to human perception.

As Garner (1974) has made clear, IT is most useful as a *normative* model of human information processing, that is, as a benchmark against which to compare human performance, since it prescribes theoretical limits on the efficiency of information encoding and transmission in an ideal information-processing system. IT has not been taken as a literal (or descriptive) model for perception, although there was a time when many psychologists were optimistic that it would serve that function. But IT has proven useful both in providing improved dependent variables for measuring perceptual processes and in guiding the development of psychological models. Recent years have seen a decline in its use for quantifying the amount of information processed by the perceiver, although the earlier research on that topic has not declined in importance over time. Rather, IT has been used successfully as a tool to answer certain questions about human information transmission, and once those questions were answered, interest naturally turned to other matters. At a minimum, IT places constraints on any plausible model of perception, and so no perceptual psychologist should be ignorant of it. IT has also proven useful in studies of manual control (see Wickens, Chapter 39) and of workload monitoring and supervisory control (see Moray, Chapter 40).

7.1. Basic Concepts

The essence of IT is that the information value of a stimulus is related directly to its predictability: a completely predictable or redundant stimulus conveys no information, whereas an unexpected stimulus or event does convey information. Information is defined as that which reduces uncertainty; a completely redundant stimulus has zero uncertainty associated with it, and so when it occurs, no reduction in uncertainty takes place. The connection between IT and perceptual organization derives from the fact that frequent or simple (e.g., symmetric) stimuli are predictable and thus are redundant. But translating likelihood or prägnanz into a language commensurate with IT is not entirely straightforward.

An important concept in IT is that the meaning of a stimulus derives entirely from its alternatives, that is, the stimuli that might have occurred but did not. Given this orientation, IT has proven most useful in understanding how people perceive and process stimuli drawn from clearly defined sets. For example, IT's best-known application has been the channel-capacity experiments (summarized in Miller, 1956) that measure how well observers can identify stimuli drawn from sets that vary in only one dimension, such as circles varying in size, tones varying in amplitude, and color patches varying in lightness.

It is much more difficult to apply IT to describe either an individual stimulus or stimuli that lack a clear dimensional structure or other means from which the parent set of stimuli can be inferred. For example, it is impossible to assess the complexity (or information load) of a scribble drawn on paper without knowing the dimensions along which the scribble was free to vary, or in other words, what other alternative scribbles might have occurred instead of the one actually produced. Scribbles seem quite complex perceptually, and indeed they are hard to encode, describe, copy, and remember. Scribbles are ill-defined patterns produced with very few constraints. Thus every scribble has a large number of alternatives and so carries a high information load. But unless the few constraints that do exist can be stated precisely, the exact information load (or uncertainty) of scribbles is infinite, since with no constraints an infinite number of scribbles can be generated.

Similarly, it is not obvious from IT that so simple a stimilus as a circle has a low information load, since a circle is a unique pattern. Circles are well-defined patterns with numerous constraints; they vary in only one dimension (their diameter), and so they have few alternatives and a low information load. Yet unless the number of possible diameters is specified, an infinite number of circles is possible (albeit a smaller order of infinity than with scribbles), and information load is again infinite. An ellipse, by contrast, is perceived as slightly more complex than a circle, since an ellipse is free to vary in one or two additional dimensions (its proportions and perhaps its orientation as well). In each of these cases, it has been the assumed constraints, inferred by the perceiver, that allow the alternatives to a stimulus to be enumerated. This step is necessary for using IT, but it does not come from the theory itself.

Garner (1962) has discussed the problem of the "unique stimulus" at length. The problem is particularly important for perceptual organization, where the focus of attention is often on the unique stimulus and where the observer's perceptual processes are often different from those engaged by the performance tasks where IT has proven most successful. We return to this problem, and some solutions to it, after reviewing the fundamentals of IT.

7.2. A Primer on Information Theory

This section is provided for readers who are either unfamiliar with the basics of the theory or who may need a brief review of its essentials. It concentrates on only those concepts relevant to perceptual organization. This brief overview cannot substitute for more complete discussions, which are provided in Attneave (1959), Garner (1962), and Coombs, Dawes, and Tversky (1970).

Fundamentally, the information value of a stimulus is determined by its predictability or probability of occurrence. This proposition accords well with the common-sense definition of information as *news*; when one receives a message that is completely predictable, nothing new is learned, and so the information value of the message is zero. Conversely, when an unexpected message is received, we experience surprise, and the information value of the message is high. In the limiting case of a *totally* unexpected message, the information value would be infinite; but such an event rarely if ever occurs. Technically, the information value of a stimulus is specified by its uncertainty U: when you receive information, your uncertainty is reduced by some amount (all the way to zero if the information received uniquely specifies the stimulus in question).

7.2.1. The Guessing Game. To illustrate, let us borrow from the guessing game experiment used by Shannon (1951), the originator of IT (Shannon, 1948). Suppose you are receiving a passage of text over a communication channel (e.g., a computer terminal), one letter at a time. Your task is to predict each successive letter in the sequence before you may proceed to the next letter. Since there are 26 different letters possible at each step in the sequence (ignoring punctuation marks and other nonalphabetic symbols, case distinctions, and the like), you will often experience considerable uncertainty about what letter will occur next. In fact, if we were to begin this game at some randomly selected point in the text, you might resort mainly to guesswork, because you would have little basis on which to form an educated guess.

One obvious but inefficient strategy would be to guess at specific possibilities until you hit upon the correct answer (e.g., is it an A? . . . a B? . . ., etc.). Because there are 26 alternatives, on average you would hit upon the correct answer midway through the list, after 13 guesses or so, depending on chance factors. A more efficient strategy would be to ask more general questions, beginning with, "Does the letter fall between A and M (in the first half of the alphabet)?" Although neither a "yes" nor a "no" answer pinpoints the identity of the unknown letter, you can eliminate exactly 13 alternatives in either case. Following an optimal strategy (in accordance with IT), you can arrive at the exact identity of the letter after 4.7 guesses, on the average. (The first guess reduces the number of possibilities to 13; the second to 6.5; the third to 3.25; the fourth to 1.625; and the fifth to 0.8125, at which point the stimulus is overdetermined. Thus somewhere between the fourth and fifth guess you arrive at just one possibility, which corresponds to a state of complete certainty). Ignoring chance variations, 4.7 guesses is the logical minimum needed to arrive at the correct answer. Accordingly, the information value of a single letter sampled at random from 26 letters is 4.7 bits ("bits" is a contraction of "binary digits," and it reflects the fact that the answer to each question is binary in nature because it must be either a yes or a no). Each additional binary guess allows you to choose among double the number of alternative possibilities. It follows that the uncertainty U of a signal is proportional to the logarithm (to the base 2, for convenience) of the number of equally likely alternative stimuli that could occur. If there were 32 equally likely stimuli, $U = 5$ bits, since

$$U = \log_2(32) = 5 .$$

If there were 16 stimuli, $U = 4$ bits, since

$$U = \log_2(16) = 4 .$$

In the case of 26 equally likely alphabetic characters:

$$U = \log_2(26) = 4.7 \text{ bits} .$$

If you knew that the message you were receiving was a normal English sentence, you could arrive at the correct answer after fewer than 4.7 binary guesses because of the redundancy of the language. This redundancy comes from two basic sources. First, the 26 letters in English do not all occur with equal frequency: the letter E, for example, is much more common than the letter Z. The calculations above all assume that each alternative stimulus is as likely as the next, but when the probabilities of the alternatives vary, uncertainty is reduced and redundancy increases. Second, the sequence of letters in English text is constrained by orthographic rules that prohibit certain sequences (such as zq), mandate other sequences (such as u following q), and less dramatically favor certain ordered pairs over others (gh is more common than hg). Similarly, syntactic and semantic rules constrain what words are likely to follow others in normal text, which in turn increases predictability. In general the average uncertainty that exists when a variety of alternatives (with differing probabilities) is possible is as follows:

$$U = - \sum_i p_i \log_2 p_i , \qquad (1)$$

where p is the probability of occurrence of each of the i alternative stimuli. Stated verbally, the average uncertainty is given by the sum of the logarithms of the alternatives probabilities, each weighted by the negative of the alternative's probability.

7.2.2. Information Transmission and Channel Capacity. The best-known application of IT to perceptual processes came in the well-known channel-capacity experiments (Miller, 1956), which attempted to measure the maximum amount of information that could be conveyed over a single sensory channel. Subjects were presented with stimuli, one at a time, that varied along a single dimension, such as tones varying in amplitude. The task was one of absolute judgment; that is, subjects were asked to respond to each stimulus with a label (typically a digit) that uniquely identified the stimulus. The number of alternative stimuli possible was varied, so the average uncertainty of the stimuli varied as well. Channel capacity was measured by the maximum amount of information that subjects could transmit about the stimuli in their responses.

The procedure for measuring information transmission from such an experiment is as follows. First, a two-dimensional confusion matrix is created to summarize the subjects' data. In this matrix the rows correspond to the different stimuli and the columns to the different responses. The entry in each cell of the matrix is the probability of occurrence of that particular stimulus-response combination (e.g., the proportion of trials on which a given response was produced for a given stimulus). The entries in all the cells in the matrix must sum to unity. The marginal row and column totals of the matrix indicate the relative frequencies of the stimuli and of the responses without regard to their co-occurrence or correlation.

The uncertainty of the stimuli U_s, which is entirely determined by the experimental design, can be computed by applying Eq. (1) in Section 7.2.1 to the row marginals. Similarly, the uncertainty of the responses U_r, which is entirely determined by the subjects' responses, can be computed from the column marginals. Last, the uncertainty of the total set of co-occurrences (called the *joint* uncertainty, or $U_{s,r}$) can be calculated in identical fashion. Given these three terms, information transmission $U_{s:r}$ is then computed as

$$U_{s:r} = U_s + U_r - U_{s,r} . \qquad (2)$$

In effect, information transmission is a measure of the nonmetric correlation between stimuli and responses. If each stimulus is invariably assigned its unique response, then all four quantities in Eq. (2) will have equal values. That is, U_r will equal U_s because the different responses occur with frequencies identical to the different stimuli, and $U_{s,r}$ will have the same value as well because each row and column of the matrix will have only one nonzero value. It follows from simple algebra that $U_{s:r}$ will have the same value as the other three terms, and information transmission will be perfect since all the information contained in the stimuli has been conveyed with perfect, one-to-one correspondence in the responses. However, should each stimulus be associated with more than one response, the correlation will drop; the $U_{s,r}$ term will increase (indicating more alternative stimulus-response pairings), and $U_{s:r}$ will decrease. Finally, should there be no correlation between stimuli and responses, the $U_{s,r}$ term will be equal to the sum of U_s and U_r, and $U_{s:r}$ will become zero, indicating that no information has been transmitted.

7.2.3. Experimental Data. A vast literature on channel capacities for various perceptual dimensions has yielded a remarkable convergence on an estimated capacity of between two and three bits for virtually all unidimensional continua (Garner, 1962; Miller, 1956). This estimate corresponds to perfect performance in identifying between four and eight equiprobable alternative stimuli. For example, Garner (1953) found a channel capacity of 2.1 bits for identifying tones that varied in loudness. This indicates that as long as there were no more than four or five different tones presented, subjects could identify them consistently without error. But beyond five stimuli, information transmission remained at a ceiling of 2.1 bits, indicating that a channel capacity limit had been reached, as depicted in Figure 36.35. (This figure also indicates the unique ability of the information transmission measure to reflect a channel's maximum capacity. Other measures, such as overall percentage correct identification, show perfect performance at low levels and declining performance at higher levels of stimulus uncertainty. Thus the percentage correct measure fails to capture the manner in which a channel conveys increasing amounts of information until its capacity limit has been reached.)

When the stimuli employed in these absolute judgment experiments are multidimensional, higher estimates of channel capacity are obtained. If the information conveyed by each dimension (or channel) could be processed independently, we would expect that the total information transmission estimate would equal the sum of the capacity estimates for each component dimension. However, the total is generally lower than this

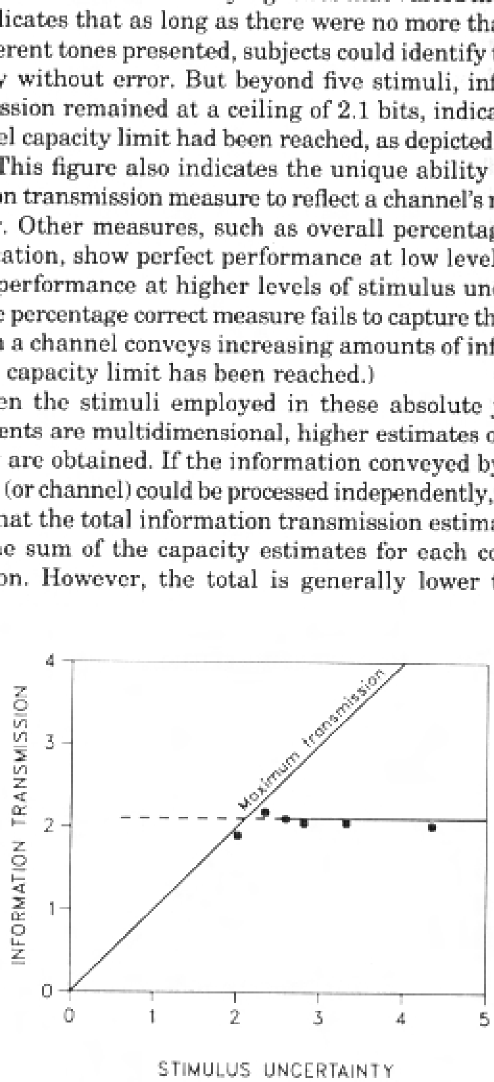

Figure 36.35. Channel capacity for auditory loudness judgments. This figure shows a typical function demonstrating a channel capacity of between two and three bits for absolute judgments of stimuli varying along a single sensory continuum. In this case the stimuli were tones whose intensities ranged from 15–110 DB, and 4–20 stimuli were used in each judgment task. The abscissa indicates the amount of stimulus uncertainty (in bits) over the different tasks; the ordinate indicates the amount of information transmitted (in bits) for each task. Although no tasks were tested in this experiment using fewer than two bits of stimulus uncertainty, the figure shows an idealized, bilinear function that levels off abruptly at about 2.1 bits. In other words, subjects continue to identify all test stimuli correctly as long as there are no more than four stimuli to be judged; beyond four, information transmission does not continue to increase. (From W. R. Garner, *Uncertainty and structure as psychological concepts*. John Wiley & Sons, Inc., 1962. Reprinted with permission.)

(Miller, 1956). For example, Egeth and Pachella (1969) had subjects judge either the horizontal or vertical position of a dot on a card and found a channel capacity of 3.4 bits for each dimension. But when subjects judged both dimensions, total capacity was measured at 5.8 bits, considerably higher than 3.4 but lower than the 6.8 predicted by additivity. One contributing reason for this underadditivity is that the dimensions of horizontal and vertical position were not being processed independently, which is consistent with evidence that these two dimensions are *integral* (Garner & Felfoldy, 1970; see also Treisman, Chapter 35).

7.3. Pattern Perception and Redundancy

A good Gestalt, as we have seen, is one that possesses a simple structure and whose components are (1) few in number, (2) regular in shape, and (3) arranged in a symmetrical fashion. Whereas poor patterns are presumed to be complex in structure, good ones are redundant; for that reason good patterns are thought to be perceived more quickly and remembered more accurately than poor patterns. Many experiments on pattern perception inspired by IT have failed to corroborate these claims. Mixed results have arisen in part because of IT's focus on sets of stimuli rather than on the individual stimulus and in part from confusion about the meaning of redundancy.

7.3.1. Experimental Evidence. Let us consider three experiments that Garner (1962) analyzed in detail to explain this problem. The first experiment, by Bricker (1955), required subjects to learn verbal labels for the visual patterns shown in Figure 36.36 (Garner, 1962, p. 188). The stimulus display consisted of five pairs of lights (denoted as a, b, c, d, and e in Figure 36.36, which shows all eight displays tested), one light above the other. Patterns were created by illuminating either the upper or the lower light of each pair. The dependent variables were the speed of learning the pattern-label pairs and the speed of responding to the briefly flashed light patterns. In one condition, only the three rightmost light pairs (c, d, and e) were illuminated, while in another condition all five pairs were used. As Figure 36.36 reveals, the three rightmost lights are illuminated in all possible combinations over the eight stimuli. Thus these three-light displays involve three bits of uncertainty, one for each light pair. The five-light displays also involved just three bits of uncertainty, even though they contain five lights. The reason may be explained in two equivalent ways. First, light-pairs a and b are completely redundant because pairs c, d, and e alone uniquely identify each of the eight patterns; thus the two bits of information provided by pairs a and b are redundant and need not be attended to. A second way to understand this logic is that while five pairs of lights can be used to generate 32 possible patterns (with an uncertainty of five bits), only eight patterns were actually used in the experiment (with an uncertainty of just three bits); thus, the selection of just eight of the 32 possible patterns creates two bits of redundancy.

Bricker's results were clear-cut: the redundant, five-light patterns were both learned and responded to more slowly than the three-light patterns. In short, redundancy hurt performance rather than helped it. The only way to increase redundancy in a set containing a fixed number of patterns is to increase the number of dimensions along which they are free to vary. Increasing redundancy thus makes the patterns more *complex*, not more simple. The added redundancy will often make the patterns more discriminable from each other (because they will differ in more ways than will less redundant sets), but their added complexity may require *more* time, not less, to perceive and to learn, as Bricker apparently discovered. In a similar experiment by Deese (1956), complex figures were responded to more slowly but also more accurately than the simple ones (see Figure 36.37). In this case redundancy did lead to greater discriminability, but only at the cost of increased processing time. That is, subjects could exploit the redundant information in the complex patterns, but it took them extra time to do so. In short, the effects of redundancy on human perception and performance are not always obvious.

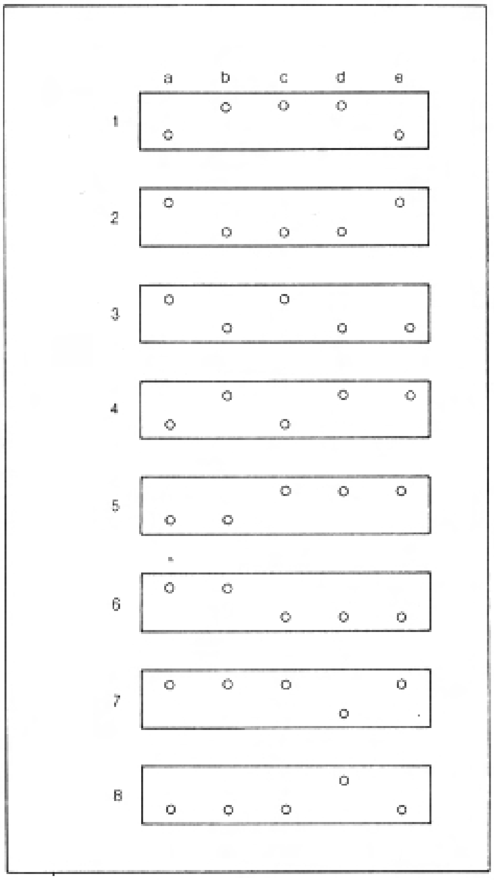

Figure 36.36. Bricker's (1955) stimuli. The stimuli were presented in a display containing ten lights, arranged as two rows of five lights each. Patterns were created by illuminating one of the two lights in each of the five columns. In one condition only the lights in the three rightmost columns (c, d, and e) were illuminated; in the other condition, the lights in columns a and b were added, but these were completely redundant with the other three lights. The dependent variables were the speed of learning verbal labels for the eight patterns used, and the speed of responding with these labels to the briefly flashed light displays. (From W. R. Garner, *Uncertainty and structure as psychological concepts*. John Wiley & Sons, Inc., 1962. Reprinted with permission.)

A third experiment by Attneave (1955) is especially useful for contrasting the information-theory meaning of redundancy as applied to sets of patterns and the less formal Gestaltist meaning of the term as applied to individual patterns. This experiment required subjects to identify (by labeling or by reproducing) dot patterns that were created by filling in cells of two-dimensional matrices. The cells of the matrices were filled in either at random or in a fashion constrained to produce symmetrical (mirror-reflected) patterns. Attneave's main results were that the symmetrical patterns were identified more ac-

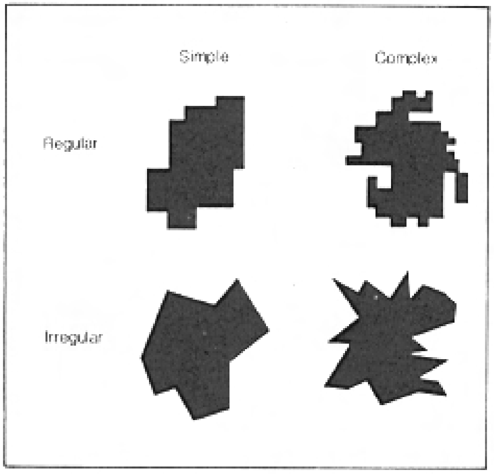

Figure 36.37. Deese's (1956) stimuli. The four patterns shown are a representative sample of Deese's stimuli, which varied in (1) "simplicity," measured by the number of sides the pattern contains, and (2) "regularity," measured by whether the pattern was constrained to contain only right angles or could contain angles of any magnitude. Subjects were shown individual patterns and then were required to select that same pattern from a group of patterns of the same level of simplicity and regularity.

curately than the random (asymmetrical) ones, and that patterns created within smaller matrices were identified better than those created with matrices containing more cells.

From the Gestalt viewpoint, Attneave's results are quite reasonable, because they show superior performance with symmetrical patterns and with less complex patterns created with smaller matrices. From an information-theory viewpoint, Garner notes, quite different conclusions may be drawn. First, Attneave's symmetrical patterns are no more (or less) redundant than his asymmetrical patterns in the technical sense of the term. That is, *any* pattern created within a matrix of a given size is as predictable as any other, regardless of its symmetry, for the same reason that in flipping coins a sequence of five heads in a row is no more or less predictable than any of the other 31 possible sequences. Selecting only symmetrical patterns has no bearing on the *amount* of redundancy, which is solely a function of the number of patterns actually selected relative to the maximum number that could have been selected. Instead, selecting symmetrical patterns affects only the *form* of that redundancy. Second, the decrease in performance as the pattern matrices grew larger shows the same deleterious effect of redundancy on performance that was demonstrated by Bricker (1955). Because the same number of visual patterns were employed at each matrix size, the patterns from larger matrices were more redundant (that is, they differed from each other in more ways; or equivalently, they were drawn from a larger parent set) than those from the smaller matrices.

7.3.2. The Unique Stimulus. A major obstacle in applying IT to pattern perception has been that IT allows us to measure the redundancy of sets of patterns but not of an individual pattern such as a single circle or a particular polygon. Indeed, redundancy (like correlation) applies only to sets of items and in particular to constraints on the way in which they vary. A single item has no variability and therefore can have no measure of redundancy.

Several ways around this obstacle have been attempted with some success, although no general solution to the problem has been proposed. One approach has been to abandon the framework of IT for other formal methods of measuring pattern complexity; this approach is discussed in Section 8 under the topic of structural information theory and elsewhere in this volume in the context of autocorrelation techniques. The other kind of approach, which stays within the framework of IT, was proposed by Garner (1962) and developed by Garner and his colleagues (Garner & Clement, 1963; Royer & Garner, 1966). The essential idea is that when a simple, regular, and symmetrical pattern is said to be redundant, this means that the perceiver infers that the pattern has been drawn from a small subset of alternative patterns. That is, the perceiver makes an inference about what dimensions were free to vary in the pattern, or equivalently what constraints were used in the process of generating the pattern.

7.3.3. Inferred Subsets. This concept of inferred subsets has been developed most fully for simple dot patterns like those shown in Figure 36.38 (Garner, 1974, p. 12; see Palmer, 1982, for a recent expansion of Garner's ideas). These patterns are produced by filling in selected cells of an imaginary 3 × 3 matrix. Because the matrix contains nine cells, there are 512 (i.e., 2^9) patterns possible. However, only patterns containing at least one dot in each row and each column of the matrix were used. This constraint reduces the number of possible patterns to 120. The two stimuli in the leftmost column of Figure 36.38 are judged by subjects to be the "best" Gestalts and are most likely to be described as redundant. Further, they are the most symmetrical patterns possible because they show symmetry around four axes (horizontal, vertical, and the two diagonal axes) and are rotationally symmetrical as well. The patterns

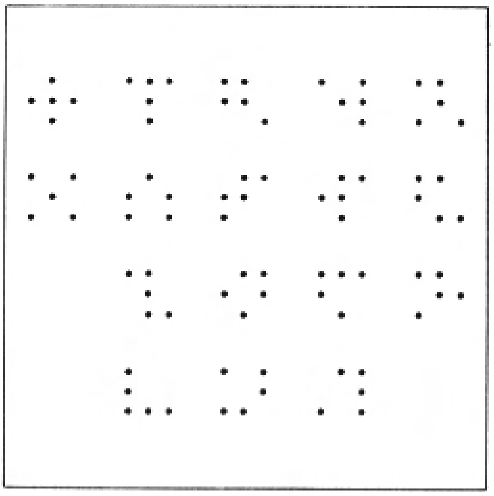

Figure 36.38. The Garner-Clement dot patterns. The patterns shown are a representative sample of the stimuli used first by Garner and Clement and subsequently by a large umber of investigators. The patterns are created by filling in five of the nine cells of an imaginary 3 × 3 matrix, subject to the constraint that each row and column in the matrix contain at least one dot. The patterns in the leftmost column are the "best" configurations and come from an R & R subset of one, which means that only one pattern results as the stimulus is rotated in 90° increments or is reflected. The patterns in the next two columns are of intermediate goodness and come from an R & R subset of size four. The patterns in the two rightmost columns are poor and come from R & R subsets of size eight. The results of numerous experiments show that the better patterns are easier to remember than the poor ones and so show superior performance in a variety of perceptual information-processing tasks. This figure demonstrates the concept of inferred subsets. (From W. R. Garner & D. Sutliff, The effects of goodness on encoding time. *Perception and Psychophysics*, 1974, *16*. Reprinted with permission.)

in the next two columns are judged to be less "good," and physically they are symmetric around fewer axes than the previous stimuli. Finally, the patterns in the two rightmost columns are judged to be the "poorest" patterns, and objectively they contain no symmetry.

According to Garner, perceivers infer that the better patterns are drawn from smaller subsets of equivalent stimuli than the poorer ones. He also has shown that a pattern's equivalency subset can be regarded as the set of patterns produced when that pattern is subjected to all possible combinations of rotation (in increments of 90°) and of reflection. When the two best patterns in Figure 36.38 are so rotated and reflected, the patterns remain unchanged. Thus they are said to come from a subset of just one pattern. In Garner's notation, their R & R (rotation and reflection) subset size is one. The patterns of intermediate goodness yield four different variations when rotated and reflected, and so they come from an R & R subset of size four. Last, the poor patterns yield eight variations and so come from an R & R subset of size eight.

The notion of an inferred subset takes us from the realm of unique stimuli to the realm of sets of stimuli, and so the information-theory concept of redundancy now becomes meaningful. The best patterns are maximally redundant, whereas the poorer patterns possess progressively less redundancy. This kind of redundancy *does* facilitate pattern discrimination and pattern memory in a variety of information-processing tasks; as we demonstrate later, good patterns are better remembered and more rapidly discriminated from one another than are poor patterns.

7.3.4. Perceived Subsets. Whereas the R & R method was used to identify *inferred subsets* of equivalent dot patterns, a second method uses the number of perceived organizations that a given stimulus can possess to identify *perceived subset* sizes. The more alternative organizations an observer can impose on a stimulus, the larger is that stimulus's perceived subset size. Royer and Garner (1966) first applied this method to temporal auditory patterns. These patterns are produced by alternating between two distinguishable sounds (e.g., two tones of different frequency) according to a specified pattern. This experiment used patterns of eight elements, as shown in Figure 36.39 (Garner, 1974, p. 50), where the two auditory elements are represented by X's and O's. The patterns were presented in cyclical fashion wherein the eighth element would be followed by the first element in an unbroken tempo. Subjects were asked to describe the patterns by tapping them out on a pair of telegraphic keys as they were hearing them.

Because the patterns were presented repeatedly in an uninterrupted loop, there is no inherent starting point to the pattern. In principle, subjects could begin tapping at any of the eight positions within the sequence. The data show, however, that subjects perceive clear starting points for most such patterns because they start tapping only at certain of the eight possible positions. Subjectively, too, the patterns appear to be strongly organized, with distinct beginnings and ends. A major finding from this research is that some patterns have only a small number of perceived starting points, whereas other patterns have many. This fact provides the necessary link to measure the size of a pattern's perceived subset and thus to assess its redundancy. In agreement with the results on dot patterns, Garner and his colleagues found that the more redundant temporal patterns (that is, those with few perceived alternatives) were perceived more rapidly and accurately than the less redundant ones.

```
...X X X O X O X O...
  ...X X O X O X O X...
    ...X O X O X O X X...
      ...O X O X O X X X...
        ...X O X O X X X O...
          ...O X O X X X O X...
            ...X O X X X O X O...
              ...O X X X O X O X...
```

Figure 36.39. Royer and Garner's repeating auditory patterns. Shown is one binary auditory pattern of length eight, which consists of a continuing alternation of two auditory stimuli (e.g, a bell and a click), represented here by an X and an O. Because the patterns are repeated indefinitely, they can be described using any of eight possible starting points, as shown here by the eight rows. The subject's task is to learn the patterns and to respond to them by tapping out the sequence on a pair of telegraph keys. Although subjects could begin tapping out the pattern at any of the eight starting points, for many patterns subjects used only one or two, indicating that they had perceptually organized the pattern with a clear starting and ending point. For other patterns, however, many more starting points were used, indicating that these patterns were less organized perceptually, or in other words were poorer patterns. In agreement with the work on visual configurations, the poorer patterns were more difficult to learn and to perceive. This figure demonstrates the concept of perceived subsets. (From W. R. Garner & D. Sutliff, The effects of goodness on encoding time. *Perception and Psychophysics*, 1974, *16*. Reprinted with permission.)

The perceived subset approach has been shown to be applicable to dot patterns as well as to temporal patterns (Pomerantz, 1981). When subjects are asked to indicate their perceived organization of a dot pattern by drawing connecting lines between the dots to indicate the pattern's perceived structure, the "good" patterns yield many fewer alternative organizations than do the "poor" patterns. Thus the perceived subset technique may have some generality for assessing the redundancy of many different types of pattern.

7.4. The Role of Pattern Goodness in Information Processing

At this point, we have established that information-theoretic constructs can be adapted to capture the psychologically relevant properties of good Gestalts such as simplicity and symmetry. We have also seen that patterns that are good Gestalts can lead to superior perceptual and memory performance. Let us now examine in more detail the precise nature of these performance effects to see when they do and do not appear and to gain a clearer conception of how and where they arise in the perceptual and memory systems. This discussion is limited to dot patterns because these are the patterns that have been studied most exhaustively in the laboratory.

7.4.1. Memory. Attneave (1955), discussed briefly before, showed that good (symmetrical) dot patterns are not necessarily remembered better than poor ones. Subjects reproduced patterns in an empty matrix immediately following their presentation. One set of patterns was produced within a 3 × 4 matrix. A second set was produced within a 5 × 4 matrix, either by filling in cells at random or by reflecting the pattern of the 3 × 4 matrix about the vertical axis to yield a pattern with vertical symmetry. The third set was produced within a 5×7 matrix, either by filling in cells at random or by reflecting the pattern of the 5×4 matrix about the horizontal axis to yield a pattern with both horizontal and vertical symmetry. The results showed superior performance for symmetrical patterns over asymmetrical patterns within the same-sized matrix; and superior memory for patterns in the smaller matrices than in the larger. But the doubly symmetric patterns within the 5×7 matrix were remembered more poorly than the singly symmetric ones within the 5×4 matrix, which in turn were remembered more poorly than the random (asymmetric) patterns in the smallest 3×4 matrix. Thus symmetry does improve memory when the total matrix size is held constant; but a small, asymmetric pattern is remembered better than a larger pattern produced by reflecting it about an axis.

7.4.2. Perceptual Discrimination. Clement and Varnadoe (1967) tested the discriminabilities of pairs of dot patterns. They used the Garner-Clement patterns shown in Figure 36.38 in a speeded card-sorting task. The dependent measure was the time required to sort a shuffled deck of cards into two piles (one for each of the two different patterns in the deck). The main finding was that sorting times were shortest when two good patterns were discriminated and longest when two poor ones were discriminated; discriminating a good from a poor pattern required an intermediate amount of time.

Subsequent experiments have attempted to localize the stages of processing responsible for Clement and Varnadoe's discriminability effect. Although their result might suggest that the effect is due to slower pattern comparison for poor patterns, it could also be due to either slower encoding of poor patterns or to a more difficult memory load in tasks involving poor patterns. (Because subjects must keep the patterns to be discriminated in memory while performing the task, a greater memory load for poor patterns could impair performance.) Garner and Sutliff (1974) repeated the Clement and Varnadoe experiment using a discrete reaction time (RT) procedure to obtain RTs to individual patterns (something not possible with the card-sorting technique). Besides replicating the earlier findings, Garner and Sutliff found faster RTs to the good pattern than to the poor one in the tasks requiring discriminating one good from one poor pattern. This asymmetry suggested that good patterns were perceived or encoded faster than poor ones. Consistent with this interpretation is Bell and Handel's (1976) finding that under conditions of poststimulus masking, good patterns are identified more accurately than poor ones, even though no such difference is apparent when no mask follows the stimulus presentation. The logic here is that good patterns are perceived more quickly and thus stand a better chance of being encoded fully before the appearance of the mask.

Checkosky and Whitlock (1973) used Sternberg's (1969) additive factors method to localize the effect of pattern goodness in a memory-scanning task. In this task subjects memorized sets of two or three patterns, which comprised the memory set. The patterns in the memory set were either all good patterns or all poor patterns. Probe patterns were then presented one at a time; subjects were to make one response if the probe was a member of the memory set and a different response if it was not. In addition, probes were presented under conditions of either high or low contrast (visibility). Checkosky and Whitlock found an interaction of memory set goodness and memory set size (two versus three patterns) such that the slope of the function relating RT to the size of the memory set was greater when poor patterns comprised the memory set. Following Sternberg (1967), this indicates an effect of goodness on the speed of memory comparisons, and it is consistent with the findings reviewed previously showing that good patterns are remembered better than poor ones.

Checkosky and Whitlock further found no interaction between probe goodness and probe contrast. Following additive factors logic, two factors that affect the same stage should yield interactive effects on RT; presuming that probe contrast affects the encoding stage, this logic dictates that pattern goodness does *not* affect encoding. However, Garner (1974) analyzed Checkosky and Whitlock's data further and discovered that probe goodness affected the y-intercept of the function relating RT to memory set size, with good probes showing smaller intercepts than poor ones. Such a result indicates an effect of goodness in some stage of processing separate from memory operations, perhaps in the encoding stage.

Pomerantz (1977) attempted to clarify this conflict about whether pattern goodness affects speed of encoding with a discrete RT experiment in which the positive (memory) set consisted of only a single pattern (good or poor), while the negative set contained either two or four patterns, which could be all good or all poor. The results were as follows: the goodness of probes that were in the positive set had a large effect on RTs (indicating a goodness effect on memory), but the goodness of probes in the negative set (which had to be encoded but were not held in memory) did not matter. Accordingly, Pomerantz's (1977) con slusions were in agreement with those of Checkosky and Whitlock (1973) that good patterns are perceived no more quickly than poor ones. Good patterns are processed better than poor ones in a variety of performance tasks, but the advantage held by good patterns appears to be due to the memory component of processing.

This conclusion is important for theories of perceptual organization because it is simple to imagine models in which good patterns would be encoded more quickly than poor ones. Good patterns are symmetrical, contain fewer perceived parts, and contain other structural properties that might allow them to be encoded more rapidly than their poorer counterparts. Perhaps this conclusion that goodness does not affect encoding speed should be tempered by noting that these discrimination time tasks may not require as detailed an analysis of the stimulus as do other tasks that have a more substantial memory requirement. That is, in discrimination tasks patterns must be encoded only to the point where the subject can determine whether they match another pattern held in memory. In Bell and Handel's (1976) experiment, subjects presumably had to encode patterns into a more lasting form of memory representation so they could reproduce them a few moments later. This experiment showed a substantial goodness effect, which suggests that good patterns can be recoded into memory more quickly than poor ones.

7.5. Summary

As noted at the outset of this section, IT has proven a useful, and in several applications a necessary, tool for understanding the nature of the stimulus and how stimulus information is processed perceptually. IT describes limits on the processing capabilities of an ideal observer who retains and integrates information without error and who knows the objective, a priori probabilities of occurrence of the various alternative stimuli. IT does not address how information is processed in human perception and so cannot be used as a descriptive model for the human perceiver. But it can serve as a useful guide once we know how subjects derive their own sets of alternative stimuli. More important, IT has provided essential measurements of perceptual performance and has helped us understand what is meant by such otherwise ill-defined concepts as simplicity and redundancy. Those concepts are difficult enough to make operational even with the aid of IT; without such a formal model, our level of confusion would be far greater than it currently is.

Ultimately we want a model that describes both the functional information load carried by a stimulus (and its alternative organizations) and the processes through which the stimulus is organized and perceived. The next section describes an approach that addresses the first of these two goals. Known both as structural information theory and as coding theory, this approach is not a descriptive model of the processes by which perception operates. But it does provide a framework for developing models for how complex stimuli may be described, models that would make testable predictions of how these stimuli are organized perceptually.

8. STRUCTURAL INFORMATION THEORY

Previously called *coding theory*, structural information theory (SIT) was developed by E. L. J. Leeuwenberg and H. F. J. M. Buffart, and in recent years by the late Frank Restle. The theory is a descendent of the work of Simon and his colleagues (Simon, 1972, 1978; Simon & Kotovsky, 1963), of Restle (1970b), and of Vitz and Todd (1969). Our presentation does not follow in all particulars the accounts of the theory written by Leeuwenberg and his co-workers because as it has evolved the theory has undergone changes both in form and in substance, ranging from the notational system used to the predictions generated. SIT is more difficult to evaluate than information theory, in part because the theory is still in a state of flux, because some of the theory's components have not yet been formalized, and because some of the main papers appear not to have been adequately translated from Dutch into English.

The key papers we have used are Buffart and Leeuwenberg (1981, Note 1); Leeuwenberg (1978); Buffart, Leeuwenberg, and Restle (1981); and Leeuwenberg (Note 2). We also consulted Leeuwenberg (Note 3) and Butler (1982). (For additional discussion of Leeuwenberg's work, see Chase, Chapter 28.)

8.1. Overview

SIT is designed to deal with a class of problem illustrated in Figure 36.40. The pattern in the center of the figure (a) is ambiguous: there are at least four ways to decompose it into parts. The most common interpretation is that labeled (b), two overlapping squares. Why is this interpretation more frequently perceived than the others? We have discussed some earlier approaches to this problem, most notably Hochberg and McAlister's (1953) analysis of the perception of perspective drawings of cubes. SIT differs from Hochberg's approach only in its degree of elaboration and formalization and its strategy of translating spatial patterns into list form.

The theory consists of the following five parts (note that these parts are not hypothetical stages of perceptual processing because SIT is not a process model):

1. Preliminary data. A set of mutually exclusive interpretations of a stimulus is obtained from phenomenological reports of

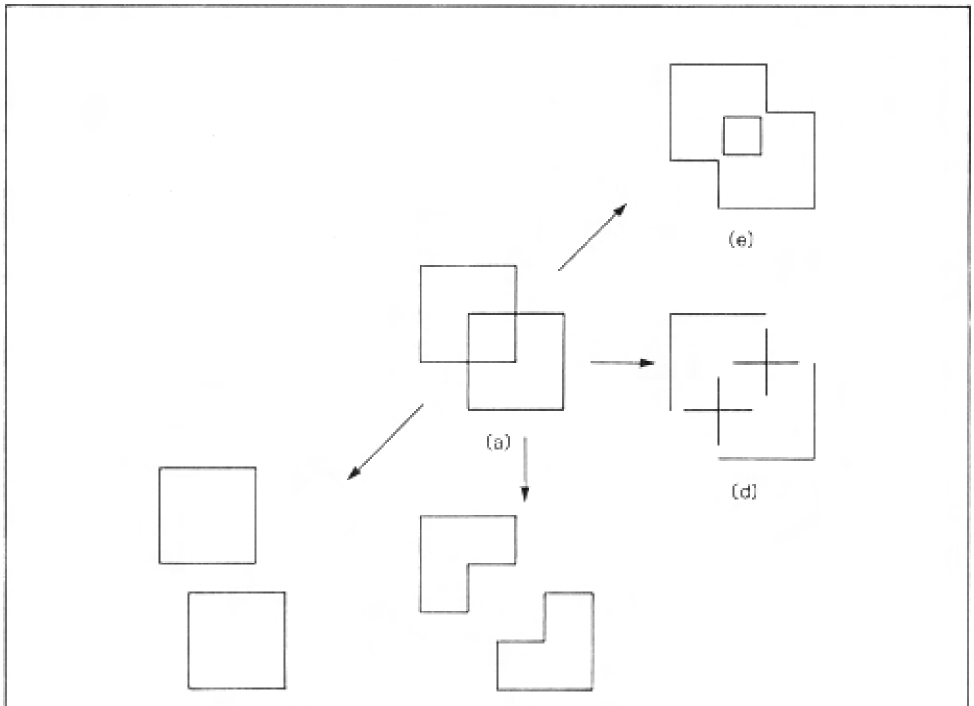

Figure 36.40. Four interpretations of a visual pattern. The central pattern shown at (a) can be interpreted by structural information theory in at least four ways (b–e). These four interpretations will result in four unique codes as described in the text. (Adapted from Leeuwenberg, 1978, Figure 1.)

observers. SIT assumes that the perceptual system considers more than one interpretation before the stimulus is perceived in a fashion consistent with one of these interpretations. The process by which the perceptual system constructs these alternative interpretations is called *perceptual inference*.

2. Semantics. A primitive code is computed for each interpretation. The procedure for generating a primitive code from an interpretation is embodied in the *semantics* of SIT, which consists of a set of *semantic operators* and *rules* for their application.
3. Structural information load. The number of parameters in the code, or the code's *structural information load* (SIL), is calculated.
4. Syntax. Rewrite rules are used to extract all redundancy from the primitive code of each interpretation. This is equivalent to reexpressing a primitive code to minimize its SIL. The product of this process is called an *end-code*.
5. Perceptual decision. A procedure to calculate the *strength of interpretation preference* underlies SIT's prediction of the outcome of the *perceptual decision* process, that is, the likelihood of a given interpretation's being assigned to the stimulus.

8.2. Semantics

Each interpretation of a stimulus can be coded, that is, represented symbolically. An interpretation of a line drawing is usually coded in terms of two kinds of *primitive elements*, line segments and angles, after an origin (a starting point and a starting direction) has been specified. The resulting string of symbols is called a *primitive code*. Each symbol in a primitive code can be viewed as an instruction to a draftsperson, and the whole string can be viewed as a program to reconstruct the entire figure. For instance, Figure 36.41a, is coded

$$\overline{1}\ \hat{90}\ \overline{1}\ \hat{90}\ \overline{1}\ \hat{90}\ \overline{1}\ , \qquad (3)$$

where $\overline{1}$ represents the length of the side of the square, and $\hat{90}$ is a counterclockwise 90° angle.

It is unclear from available accounts of SIT whether the code of an interpretation is or is not unique. For example, in Figure 36.41b, the starting direction associated with the origin is different from that in Figure 36.41a. The corresponding primitive code is

$$\hat{90}\ \overline{1}\ \hat{90}\ \overline{1}\ \hat{90}\ \overline{1}\ \hat{90}\ \overline{1}\ . \qquad (4)$$

Similarly, the primitive codes for Figure 36.41*c* and *d*, respectively, are

$$\overline{i}\ \hat{90}\ \overline{1}\ \hat{90}\ \overline{1}\ \hat{90}\ \overline{1}\ \hat{90}\ \overline{j}\ , \qquad (5)$$

and

$$a\ i\ \hat{90}\ j\ i\ \hat{90}\ j\ i\ \hat{90}\ j\ i\ \hat{90}\ j\ , \qquad (6)$$

where i and j represent line segments such that $i + j = \overline{1}$, and a is an acute counterclockwise angle. Given the current state of the theory, it is unclear whether these four expressions should be considered as primitive codes for a single interpretation or for four different interpretations of the square; on the basis of parsimony, we assume the latter.

The semantics of SIT are not limited to the coding of straight-line patterns. For a brief treatment of the coding of drawings that contain curved segments see Buffart, Leeuwenberg, and Restle (1981). Restle (1979) developed the semantics for the coding of motion displays of the kind used by Johansson (1950). In these dynamic displays each of a small number of lights was displayed either in uniform circular motion or in a parallel projection of a uniform circular motion. When the trajectory of a light was not circular, it was either elliptical (in which case the dot moved most quickly at the minor axis and most slowly at the major axis) or linear (in which case its velocity was a sinusoidal function of time). Figure 36.42 illustrates the parameters required to specify these motions. The projected uniform circular motion that underlies each light in the display requires three parameters to be specified completely: its amplitude (the radius of the circular trajectory, a, in units of length); its phase (the position of the light on its trajectory at some reference instant, ø); and its frequency (f in Hertz). Two more parameters (the angles b and t) are required to specify the orientation of the plane in which the underlying circular motion is taking place relative to the picture plane. Thus if the motion of a system composed of a single light p is interpreted as the

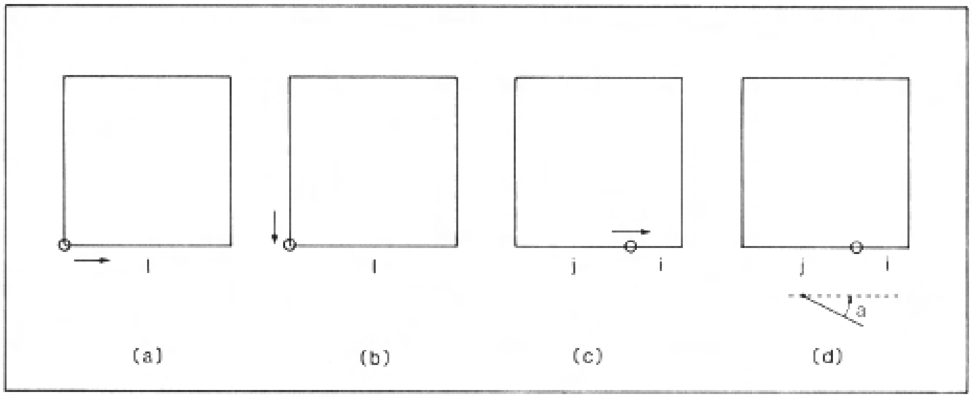

Figure 36.41. Four ways to code a square. The coding process begins with the selection of an origin, defined as a starting point and a starting direction. The four codes shown differ in their origin. As the text explains, it is unclear whether these different origins will all result in unique, different codes. The origin is indicated by a circle. The primitive elements are the line segments *l*, *j*, and *i* and angles a. The arrow shows the starting direction.