# HOW MAPS WORK

## Representation, Visualization, and Design

ALAN M. MACEACHREN
*Department of Geography*
*Pennsylvania State University*

THE GUILFORD PRESS
New York London

# CHAPTER THREE

# How Maps Are Seen

A key aspect of Marr's (1982) approach to vision is his contention that there are three levels of explanation from which to address an information-processing system. The computational level focuses on the what and why. Considering vision at this level, and recognizing that vision involves a series of representations and processes that interpret those representations and build new ones, we begin the task of understanding how maps are seen by asking what the purpose of seeing is. According to Marr, this purpose ultimately focuses on recognizing and identifying shapes in the real world. At intermediate stages, however, there are representation-specific purposes that can be identified. In moving from the initial visual scene as sensed by the retina of the eye to Marr's primal sketch, the purpose can be defined as extracting contrast information (related to differences in intensities and wavelengths) and grouping this information to form edges, regions, and shapes. The purpose of the process leading up to Marr's 2½-D sketch (or Pinker's visual description level) is to make the depth, orientation, and junctions of visible surfaces explicit.

In relation to maps, these two goals imply that the way we establish contrast among map features will be critical at the initial level of vision. At this level, according to Marr, no higher level processes come into play, and therefore the only information available to the map viewer is contrast (from pixel to pixel of the retinal image). Although others have argued that top-down cognitive processes can have an effect even at this early stage of vision, it is clear that applying this top-down control is an effortful process. Sorting out components of a map display will be accomplished most efficiently if the cartographer creates contrasts among those map elements that are most important for the viewer to notice immediately. The second goal, associated with the next stage of perceptual representation, suggests that Gestalt principles of perceptual grouping will play an important role. Again, although top-down processes might be able to facilitate Gestalt grouping (or may interfere with it), the most successful map (at this stage of processing) will be one that elicits grouping that links map elements in logical ways (e.g., areas are seen as homogeneous regions rather than disaggregated individual point features—as might happen with use of a pattern made up of noncompatible elements spaced too far apart).

Cartographically, the goal of research directed to low-level visual–cognitive processes is to understand how the stages of physiological–conceptual representation of a map scene interact with symbolization and design variables. Ultimately, we would like to be able to predict what symbol variables or design choices make differences noticeable (in particular situations or for particular tasks), attract viewer attention, are seen as having equal salience, have an intuitive order, or induce grouping or formation of figures on backgrounds. Vision should, on evolutionary grounds, be good at extracting object shape from the visual scene, assessing depth and relative size, and noticing movement. It must perform these functions from information about contrast on a roughly pixel-by-pixel basis at a retinal level, using neurological hardware to process the retinal image. This hardware appears to rely heavily on spatial filtering and enhancement procedures operating simultaneously at several scales. These filtering procedures take into account sensations received by groups of cells. A key feature of this system is that it emphasizes contrast more than absolute illumination (as it must do if we are to recognize an object as the same at dawn and midday). The system has many more cells devoted to value/brightness difference than to hue or saturation, although those cells devoted to these "color" differences are concentrated in central vision and are agglomerated less as they pass signals to cells in the brain's visual cortex. This concentration means that we have relatively higher acuity for hue (hundreds of differences are discriminable) than for value (tens of differences or less are discriminable). A second key feature of the system is an ability to group the elements that the neurological image processing achieves into "objects," or shapes that higher level processes can match to memory representations.

Within this context, this chapter begins with a brief look at our visual hardware that has evolved to meet the above goals. Once in possession of the basic ideas about how this visual hardware has evolved to meet the needs of vision in general, we can speculate about the limits that it imposes for the abstract task of "seeing" a map. The bulk of the

chapter, then, considers selected low-level perceptual processes and the potential implications of cartographic use of the visual variables (i.e., the building blocks of map design) to create contrasts between and relations among map elements. While there is continuing debate about whether the low-level processes discussed in this chapter operate in a completely bottom-up, preattentive fashion, or are controlled (at least in part) by top-down processes, the key commonality of the processes included here is that they are fast (measured in milliseconds) and probably occur in parallel. It is this fast parallel processing that makes visualization such a potentially powerful tool for science in an era of data excess (see Part III).

## EYE–BRAIN SYSTEM

The intent of this section is to provide a brief sketch of the eye–brain system's major features and to suggest a few examples of how the eye–brain system puts constraints on the way we see symbols and read maps. Knowing the limits, constraints, and idiosyncracies of vision allows us to avoid presenting map readers with processing tasks that are difficult or impossible to perform. Understanding why such limits exist and what our visual system has evolved to accomplish can give us clues about how we might facilitate processing of map information and also clues about the implications of our decisions concerning symbol form, color, size, texture, and so on, for how the information will be processed. The examples provided may also serve to suggest some possible avenues for cartographic research that draws directly upon the quickly expanding knowledge base concerning how human vision works as an information-processing system.

How human vision works is, of course, incompletely understood. What has become clear, however, is that the system does not transfer little pictures of the world from the eye to the brain. Our "perceptions" are constructed (or reconstructed) from a multitude of fragmentary information, some of which is organized spatially (i.e., a direct mapping from positions in the environment to positions in the brain) and some of which is organized according to other attributes of the stimulus (e.g., color, orientation, texture, movement, etc.). Vision is a complex parallel-processing system in which hundreds of millions of sensing cells react to input of light through the lens of our eyes. Through multiple interconnections, these reactions cause subsequent reactions among the tens of billions of cells in our brain that are devoted to vision. Both psychophysical and neurophysiological research indicates that considerable preconscious processing of the signals occurs between the initial incidence of light on the cornea of the eye and the ultimate perceptual experience.

## The Eye

Some common conceptions about how the eye–brain system works evaporate quickly when we take a close look at the structure of the eye. A camera analogy is frequently applied. Like the lens of a camera, the human eye is arranged so that reflected and emitted light passes through a lens and results in an "image" of what is observed on a receiving surface. The extent of the image on the eye's receiving surface is a direct function of the size of the object viewed and its distance from the lens. In comparison to many cameras, the eye contains a rather wide angle lens (focal length of 14–17 millimeters), allowing representation of a scene that extends 60° to either side of the central focus to which vision is directed. Although the camera analogy tells part of the story, it can be very misleading. As the complexity and interconnections of cells in the eye become clear, the camera analogy becomes less useful. The fact that we do not have the sensation of looking at the world through a fish-eye lens is one clue to the complexity of image processing that happens between the eye and our conscious sensation of seeing.

An analogy to image analysis systems used in digital remote sensing might prove useful, at least to cartographers trying to understand implications of the eye–brain system for how map symbolization is "seen." Marr's (1982) computational models of vision will, in fact, sound quite familiar to those conversant with image analysis. His hypothesis is that one of the principle steps in vision is the extraction of "shape contours," and he describes how these contours might be extracted through spatial filtering procedures.

With a camera, a lens focuses an image directly onto a flat piece of film. With the eye, light must pass through a complicated tangle of semitransparent cells on its way to the receptors at the back of the eye, and these receptors lie on a curved surface (Figure 3.1). In addition, unlike a camera, with the eye focusing is achieved by changing the shape of the lens, rather than the distance from the lens to the receiving surface. Receptors in the eye's receiving surface (the retina) vary in density, with substantially more in central vision, and contain distinct kinds of receptor cells that respond to different input.

Two major categories of cells line the retina: rods and cones. The rods are more numerous than cones (about 120 million and 5 million, respectively, in each eye) and will respond to very small changes in intensity of light, but not when light is very bright (Figure 3.2). Rods are insensitive to differences in wavelengths of that light, and therefore to color. Cones need greater illumination in order to react but are sensitive to differences in wavelength. Cones are concentrated in a very small area in the center of the retina (the macula). The fovea, a position that is directly exposed to light entering the eye, is located at the center of the macula. This is the location of greatest visual acuity and contains no rods, only cones. Cones, by their dominance here, are responsible for our ability to see fine detail.

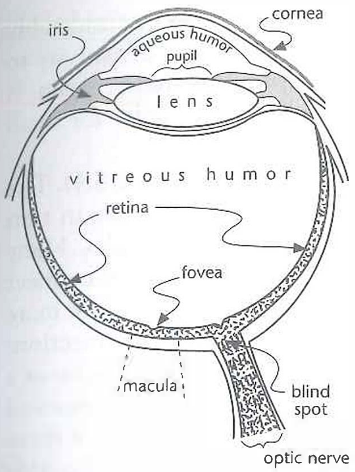

**FIGURE 3.1.** Structure of the human eye.

Cone cells, in persons with normal color vision, can be further distinguished on the basis of the wavelengths of light they respond to. These cone types are generally referred to as L cones (sensitive to long wavelengths), M cones (sensitive to medium wavelengths), and S cones (sensitive to short wavelengths). These different cone cells are unevenly distributed in the eye as well. As a result, the eye's sensitivity to different wavelengths of light varies spatially. Maps of retinal sensitivity to various wavelengths present a complex overlapping picture in which we find, for example, that sensitivity to green is confined to a relatively small hori-

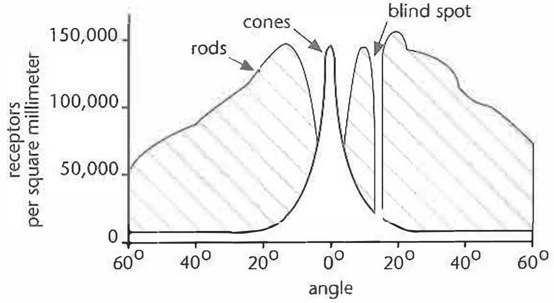

**FIGURE 3.2.** Distribution of rods and cones across the surface of the retina.

zontally extending band, while that to yellow occurs across a considerably larger, and nearly circular, portion of the eye (Figure 3.3). Sensitivity to blue, although covering a greater area of the retina than red or green, is lowest overall (in magnitude), which makes blue a poor color for small map features.

The retina is the first of three cell layers in the eye (Figure 3.4). The second consists of bipolar, horizontal, and amacrine cells. These in turn connect the receptor cells (i.e., rods and cones) to the ganglia. Many bipolar cells form direct connections. Horizontal cells, however, connect receptors with more than one bipolar cell, and amacrine cells link more than one bipolar to individual ganglion cells. These interconnections mean that a ganglion will not transmit an impulse based on stimulus of a single location on the retina; instead, it summarizes the signals received from a number of inputs, the ganglion's "receptive field." Most of these fields, in humans (as well as other mammals), are roughly circular with sufficient overlap for the foveal areas to overlap slightly (Figure 3.5).

To relate the size of receptive fields to the size of discriminable features in the visual field, the "angle subtended" by the feature is referred to. This is the angle formed from the lens of the eye to the top and bottom of the feature attended to. The angle equals that covered by the image of the feature on the retina (Figure 3.6). If, for example, you were viewing one of the pictorial symbols on a National Park Map (4 millimeters high) from normal reading distance of (approximately 460 millimeters), the image on the retina will cover 30 minutes of arc.

Ganglion cell receptive fields vary in size from the fovea to peripheral areas of the retina. Receptive field centers exhibit particularly systematic enlargement from center to periphery. Near the fovea, where the receptive field can be as small as a single receptor cell, the cells are spaced about 0.5 minutes of arc apart (2.5 micrometers). This corresponds to our greatest visual acuity. An example of a single feature that subtends 0.5 minutes of arc is a 0.13-millimeter-wide line on a map at normal reading distance (e.g., representing a road) or a 1-millimeter boundary line on a wall map viewed from about 22 feet away. In contrast to this, receptive field centers near the eye's periphery can be a degree or more. The result is sharply decreasing acuity from the center of vision to the periphery. For maps, this means that a small map symbol, identifiable when we look directly at it, will be less and less clear the further in the periphery it is. For symbols to be recognizable in peripheral vision, then, they need to be larger (Figure 3.7).

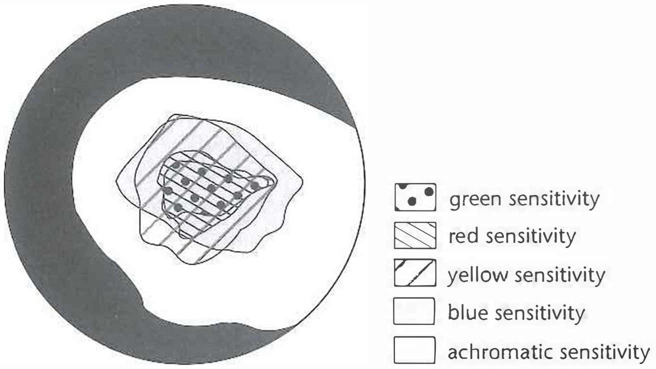

**FIGURE 3.3.** Diagram of overlapping color sensitivity regions in the eye. *After Wade and Swanston (1991, Fig. 3.20, p. 68). Adapted by permission of Routledge.*

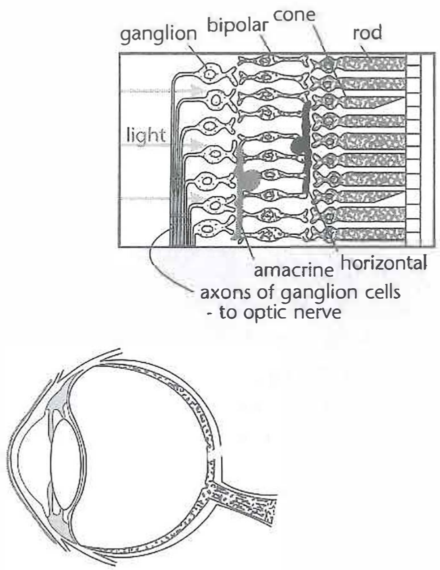

**FIGURE 3.4.** Schematic depiction of the structure of retinal to ganglion cell interconnections. *After Hubel (1988, p. 37). From* Eye, Brain and Vision *by Hubel. Copyright 1988 by Scientific American Library. Used with permission of W. H. Freeman and Company.*

If images or parts of images (e.g., skates on the feet of the figure in the National Park Service symbol for skating area) are to be seen and discriminated in peripheral vision, symbols must be considerably larger than required for discrimination with foveal vision. If differences between two symbols are small, therefore, we will require a "fixation" on the symbol to discriminate it from others and identify what it is.[1] In addition, the color sensitivity maps above suggest that the ability to see and recognize a symbol in specific regions of peripheral vision will vary with its hue.

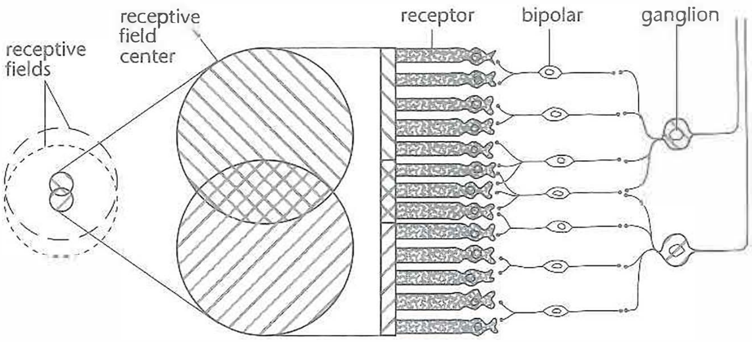

**FIGURE 3.5.** Diagram of a typical ganglion reptive field. *After Hubel (1988, p. 44). From* Eye, Brain and Vision *by Hubel. Copyright 1988 by Scientific American Library. Used with permission of W. H. Freeman and Company.*

This variation in acuity from central to peripheral vision has express implications for designing general reference and topographic map symbols. On a highway map, for example, map users often try to find particular kinds of features (e.g., points of interest, airports, universities, etc.). Symbols that are clearly distinguishable to the cartographer next to each other in the legend (when both are in the foveal area of vision) may not be different enough to be distinguishable when the map user scans across the map looking for them.

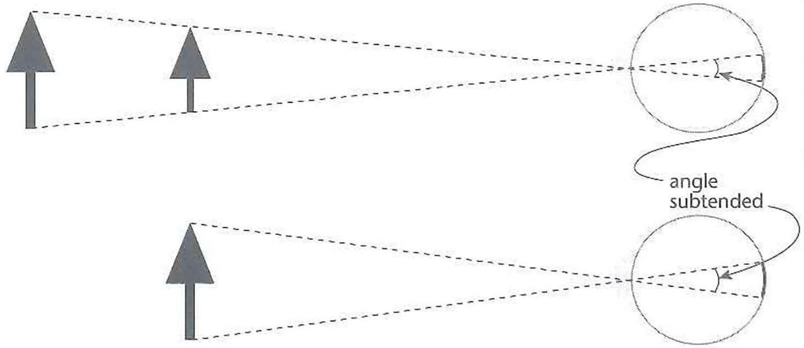

**FIGURE 3.6.** The angle subtended on the retina by light reflected from an object will depend upon both the size of that object and the distance from the eye.

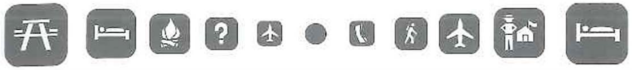

**FIGURE 3.7.** Receptive field centers for gangion cells exhibit systematic enlargement from the fovea to the periphery of the eye. The result is that acuity varies across the retina. If you fixate on the central dot (from about four inches away) all map symbols should be equally legible. *Derived from Hoffman (1989, Fig. 2.4, p. 14).*

Like most neurons in the brain (discussed below), the ganglia collecting signals from receptive fields of the retina generate impulses of a constant magnitude. What varies is their rate of firing. They exhibit a steady (resting) rate until the combined input from their receptive field reaches a threshold, at which point they either cease firing impulses or increase their firing rate. Whether the firing rate of a ganglion increases or decreases will be a function of the kind of ganglion cell stimulated, together with the spatial characteristics of the stimulus. Most ganglia react differently to stimuli near the center and periphery of the receptive field and, as a result, are termed "center–surround" cells. Both ON-center and OFF-center cells exist. With ON-center cells, a stimulus near the center of the receptive field stimulates an increase in firing rate, while a stimulus from the outer cells of the receptive field inhibits firing. With OFF-center cells, this pattern is reversed.

A constant stimulus that covers an entire ganglion's receptive field will result in competing signals that will partially cancel each other with (usually) a net result of slight inhibition on the ganglion's firing rate. If, on the other hand, the cell's receptive field is exactly centered on a small enough stimulus or it crosses an edge of some type, resulting in a different stimulus for the center and surround, the signals of center and surround can reinforce each other.

An interesting example, relevant to selection of area patterns for maps, of how this center–surround system and the size of receptive fields interact is an illusion called the Hermann grid (Figure 3.8). Most people when viewing this grid "see" dark spots at the intersections of the grid,

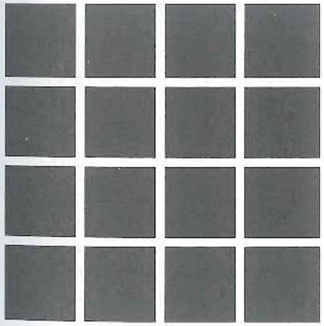

**FIGURE 3.8.** The Hermann grid illusion. Dark gray dots seem to appear at the intersections of the white lines (except at the intersetion you fixate upon). These illusory dots are thought to be the result of center–surround inhibition, with the grid intersections having a greater inhibition, thus the apparent dark spots.

unless they look right at those intersections. If we make use of peripheral vision, the ganglia being used have relatively large receptive fields (about the size of each grid zone). ON-center ganglia with receptive fields centered on the grid intersections will have the same reactions from their central areas as do ON-center cells centered over intermediate points, but an increased inhibition from the surround, resulting in the sensation of a dark spot. If you look directly at the intersection area, the receptive fields for the ganglia now involved are much smaller, and the illusion disappears. While artists sometimes make use of this effect to achieve a feeling of motion or instability in an image, we seldom want such a reaction to maps. We can prevent these distracting effects on maps by avoiding the relatively coarse patterns that match up with peripheral ganglion receptive fields.

In addition to having overlapping receptive fields, ganglia are interconnected and can inhibit each other's firing rate in the same way that a single cell's surround can inhibit the firing rate of its interior. These lateral inhibitory connections are thought to be responsible for the phenomena of "Mach bands," the illusion of shifts in brightness that cause the appearance of two vertical lines in Figure 3.9. Simultaneous contrast is also due to lateral inhibition of ganglion cells. As shown in Figure 3.10, the counties highlighted on the inset map appear to differ in degree of darkness even though they are the same.

Lateral inhibition is important in cartography because it will help accentuate differences between adjacent patterns or between symbols and background. On the other hand, it will make one pattern appear darker when next to a light pattern than when next to a darker pattern. This is one reason that there is an apparently smaller range of gray tones that people can distinguish on a map versus in gray patch experiments typical of gray scale research (MacEachren, 1982). Evidence of lateral inhibition clearly leads to the prediction that fewer shades of gray will be distinguishable on a map (where context within which any particular gray tone appears will vary) than in side-by-side comparison of pairs of gray patches. Only these out-of-context, side-by-side comparisons, however, have been used in formulating gray-tone selection guidelines. There has been a failure to test gray tones on actual maps because of the expectation that the spatial aspects of gray tone perception might confound results. As a result, cartographers have devised some tightly controlled "clean" experiments resulting in gray scales having unknown applicability for use on maps.


(a) (b)

**FIGURE 3.9.** (a) The illusion of Mach bands—the dark vertical line toward the left of the illustration and the light vertical line toward the right. These apparent lines do not exist when luminance is measured with a light meter. (b) This illusion causes layer tints on isarithmic maps to appear to gradually change in value in the wrong direction (i.e., between any two isolines, regions that should have a lower data value will end up with an apparently darker color value than regions with a high data value).

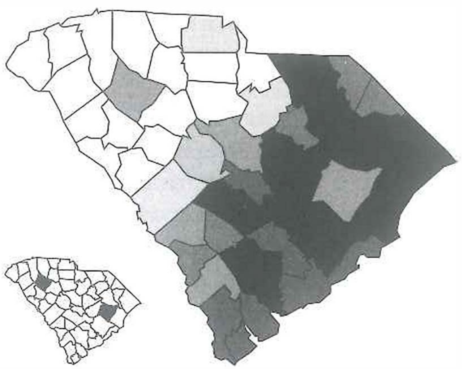

**FIGURE 3.10.** An illustration of simultaneous contrast for two map zones surrounded by predominantly light versus predominantly dark zones. Both counties (highlighted in the inset map) have the same data value and are filled (on the main map) with the same shade of gray).

The only attempt that I am aware of to measure gray tone perception on actual maps was an undergraduate term project by one of my students (Terry Idol) several years ago. The experiment used a gridded 20-class choropleth map with gray tones assigned randomly to cells. Subjects (college students) were asked to estimate the actual value (as a percent of black from 0 to 100) of specified cells. The gray scale derived from this experiment was more linear than any of the gray patch-based scales cited in the literature. Because of some printing flaws in the test maps and the small sample used in the study, I would not consider this isolated map-based gray scale experiment conclusive. If replicated, however, the interpretation would be that 0% and 100% anchor the gray scale and simultaneous contrast on actual maps tends to make light grays look lighter and dark grays look darker, thus at least partially compensating for the apparent perceptual underestimation of differences so often cited in the literature. A much more linear gray scale may apply to choropleth maps than we have suspected thus far, one that bottoms out at about 20% reflectance (or 80% black).

In addition to producing simultaneous contrast effects, lateral inhibition between adjacent ganglion cells has a major role in color perception. Ganglion receptive fields for the three types of cone cell include various opponent relationships of center and surround cells. The three general categories of relationships are (1) red–green opponent cells that include ON- and OFF-center arrangements of L and M cells, (2) blue–yellow opponent cells that include ON- and OFF-center arrangements of S with combined L + M input, and (3) dark–light opponent cells that seem to be stimulated by all three cone types. Proponents of *opponent-process theory* (OPT) argue that these opponent relationships are responsible for our full range of hue sensation (Hurvich and Jameson, 1957). The theory predicts that there are four unique hues (blue, yellow, green, and red) and that all other hues result from mixtures of these four basic colors. The theory was developed in the 19th century with neurophysiological support coming in the latter half of the present century. At least one cartographer (Eastman, 1986) has attempted to apply the theory to selection of hue ranges for choropleth maps. This application will be discussed in the next chapter.

The most comprehensive look at one result of lateral inhibition (simultaneous-contrast or surround-induced changes) in relation to color use on maps is Brewer's (1991) dissertation (which dealt specifically with this topic in the context of color maps). Her initial premise was that induction will cause colors on maps to shift in appearance toward the complement of the surrounding hue. She devised an experiment to determine whether this assumption was in fact true and, if so, whether a quantitative model of simultaneous contrast could be used as an aid to selection of easily identified map colors. The opponent-process approach to color was selected as the best starting point for modeling induction. Brewer found the expected shifts in color appearance toward the complement of the surround (e.g., a red surround makes a central color appear more green). An unexpected finding, however, was that center–surround combinations with low contrast exhibited larger shifts than those with high contrast. This result seemed to be related to color saturation. Saturation shifts turned out to be the largest shift identified in the research. Based on her experiments, Brewer devised a model of the buffer around each color that represents its potential appearance with various surrounds. The model, designed to accommodate 90% of potential map viewers, was judged a success. With this model, a cartographer can ensure that colors for map categories are not confused by selecting only colors whose color–space buffers do not overlap.

As we have with visual acuity, we tend to take for granted the similarity of the color-processing system from person to person. Some of the individual variability may, however, be critical to map design (which is why Brewer chose a 90% target rather than designing for the average map reader). As Judy Olson (1989) pointed out, for example, a significant proportion of the population has some level of color deficiency. The deficiency is usually due to the absence of or the failure to function of one or more of the cone types found in the eye. Males are particularly likely to suffer from some level of color deficiency (about 8% for males vs. 0.4% for females). Recent evidence suggests that females, in addition to being less susceptible to color deficiencies of this type, sometimes actually have an extra category of cone in their retina, thus possibly giving them an extra dimension of color vision not shared by any male counterparts. Although about 12% of females may have this extra category of cone, the necessary experiments to determine what impact it has on their color vision have not as yet been conducted.

## Eye to Brain

From the ganglia, axons extend that connect these composite signals to the next step in the process: the optic nerves. The optic nerves serve as the connecting link between the eyes and the brain. One of their primary functions seems to be the spatial amalgamation of signals from each eye. After leaving the eye, the optic nerves converge at the optic chiasma where they divide so that information from one side of each eye is directed to the same side of the visual cortex in the brain. Since the image on the retina is reversed, this sends information from each half of our visual field to the opposite side of the brain (Figure 3.11).

There are less than 1 million optic nerve fibers leading from the ganglion cells. In the outer part of the retina, up to 600 rods might be connected to one optic nerve fiber through one or more ganglion cells. In the fovea there is close to a one-to-one match of cones and optic fibers. This is one reason for more acute vision at the center of the visual system.

## Brain

In the 1960s neurophysiology predicted the ability to understand thought by understanding our neurophysiological hardware. The slow progress since the breakthroughs that found cells with apparently specific functions (e.g., recognizing edges at specific orientations, and possibly even recognizing faces) has led to a realization that neurophysiological and

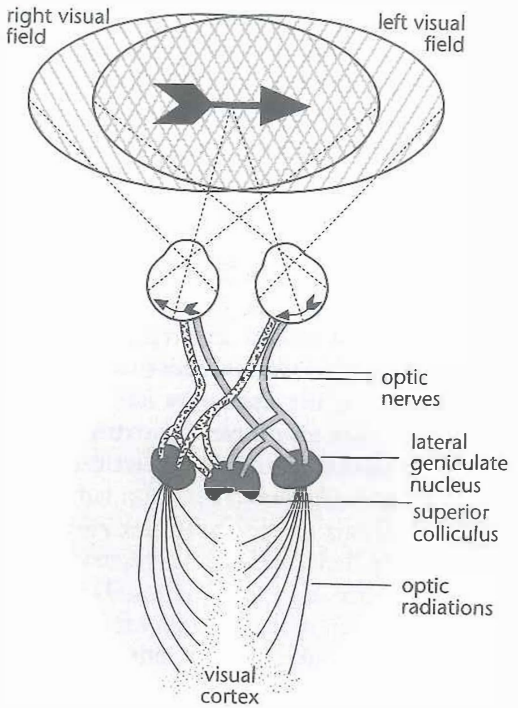

**FIGURE 3.11.** A depiction of the pathways connecting receptor cells in the eyes to the primary visual cortex in the brain. *Derived from Hubel (1988, p. 60) and Wade and Swanston (1991, Fig. 3.22, p. 70).*

neuropsychological evidence about vision will provide only part of the answer to how vision works. Following Marr's ideas, neurophysiological hardware is best considered as a mechanism that has evolved to meet the needs of vision, rather than as a fixed system that our visual abilities were adapted to. It does, however, exert limits on visual tasks that are not part of everyday behavior in the world (relatively recent visual tasks that human vision has not had time to evolve special procedures for). Reading maps seems to be one such unnatural task, with its typically abstract, two-dimensional static depiction. From a cartographic point of view, then, we are interested in features of how the brain processes visual signals not because this knowledge is likely to tell us how maps work, but because these processes put limits on what symbolization and design variations might work.

As indicated above, the signals sent to the brain by the ganglion cells result from a complex interaction of signals from each cell's receptive field together with the inhibitory interconnections of individual ganglion cells. Cells first reach the lateral geniculate nucleus (LGN) in the brain where neurons behave similarly to ganglia. Receptive fields still correspond to concentric regions in the retina. The LGN is arranged in six layers, each of which contains cells that respond to only one eye. For reasons that are as yet not understood, the six layers are arranged, from the top down, in a left, right, left, right, right, left eye sequence.

Once signals reach the visual cortex at the rear of the brain, linkages back to the retina become much less simplistic. Like the network of ganglia, cells within the visual cortex emphasize the signals coming from the macula. Approximately 50% of each side of the visual cortex is devoted to these signals. In contrast to neurons of the eye, however, those in the visual cortex have been found to be more specialized. Some appear to respond to particular visual elements such as line widths, angles, orientations, and so on, and some to the hue and brightness distinctions found with ganglion cells. The overall consensus of recent work in neurophysiology is that the visual system is composed of a sequence of processes capable of initially detecting edges, lines, and patterns (the processes Marr associated with extraction of the primal sketch) and subsequently analyzing these to result in more complex structures (Marr's 2½-D and 3-D representations).

This research has recently begun to result in maps of the brain in which the spatial arrangements of cells associated with specific functions are depicted (Figure 3.12). It seems particularly apt for a book about how maps work to include maps of the brain as a piece of evidence concerning how the brain might process maps.

Clinical observation beginning in the 19th century was responsible for the first crude mappings of the brain's major sections. It was not until the 1950s, however, when single-cell recording techniques began to yield information about individual and groups of cells that the complexity and intricacy of the neural interactions began to be recognized. Kuffer (cited in Hubel, 1988), in 1952, demonstrated the existence of the center–surround cell receptive fields described above. Much of what we now know

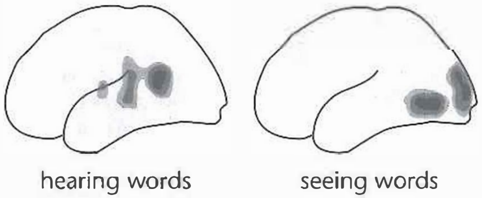

**FIGURE 3.12.** Activity maps (derived from positron emission tomography scans) that suggest the varied landscape of functions mapped out across the brain. *After Raichle (1991, color plate 3-1). Reprinted with permission from* Mapping the Brain and Its Research: Enabling Technologies into Neuroscience Research. *Copyright 1991 by the National Academy of Sciences. Courtesy of the National Academy of Sciences Press, Washington, DC.*

about limits to vision imposed by the brain's architecture is due to extensions of Kuffer's methods to examination of other neurons in the eye–brain system (usually of cats and monkeys).

Among the most startling discoveries about cells in the brain was the somewhat accidental finding by Hubel and Wiesel in 1958 that some of the cells in the visual cortex were orientation-specific (Hubel, 1988). These researchers had not been having much luck trying to determine the stimuli that would cause change in the signals of certain cells in the visual cortex of a cat. They had been using mostly opaque glass slides to block light to all but one location in the retina. Suddenly they noticed that the cell they were monitoring was stimulated as they slid a glass slide into place. Eventually they realized that it was the shadow of one edge of the slide that caused the cell to respond.

Experimental work in the intervening years has made it clear that there are at least two kinds of orientation-sensitive cell in the visual cortex. Simple cells respond only to a stimulus that is at a particular orientation *and* at a particular position on the retina (Figure 3.13). Complex cells also respond to single orientations, but are less sensitive to position of the stimulus (figure 3.14). Subsequent research led to discovery of additional cells that respond to "end stopping" (e.g., the end of a line) and to corners. Additional cells seem to respond only to movement, and some apparently only to movement of an edge across the visual scene.

How cells in the visual cortex selectively react to such specific kinds of features is still not completely understood. One likely hypothesis is that orientation-specific cells are linked to a set of ganglion cells that have receptive fields arranged in linear fashion across the retina. Figure 3.15 provides a schematic depiction of how this system might be connected.

Cells in the visual cortex are arranged in layers, and each layer seems to be somewhat distinct. Cells in some layers are binocular (responding to stimuli from both eyes) and those in other layers are monocular (responding to only one eye). Some layers contain cells that are orientation-selective and other layers have cells that are not. Near the middle of the visual cortex, in the layer know as 4C, cells appear to be arranged in two intersecting slabs, one set in which right and left eye dominance alternates and the other in which orientation selectivity varies systematically (see Figure 3.16).

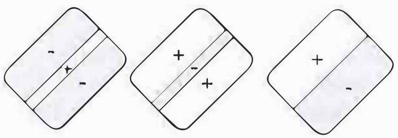

**FIGURE 3.13.** The response of orientation-sensitive cells in the brain to lines of varying orientation. The stimuli that these cells react to (from left to right) are a slit covering the (+) region, a dark line covering the (–) region, and a light–dark edge on the boundary. *After Hubel (1988, p. 72). From* Eye, Brain and Vision *by Hubel. Copyright 1988 by Scientific American Library. Used with permission of W. H. Freeman and Company.*

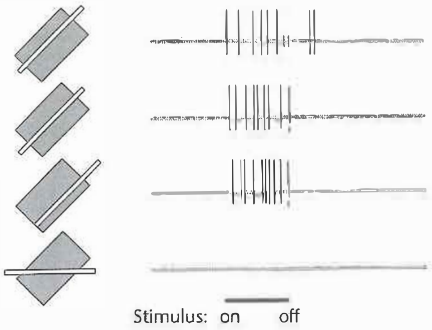

**FIGURE 3.14.** The response of complex cells to position and orientation. The plots at the right show cells reaction (or lack of it) to slits of light at different orientations and positions. *After Hubel (1988, p. 75). From* Eye, Brain and Vision *by Hubel. Copyright 1988 by Scientific American Library. Used with permission of W. H. Freeman and Company.*

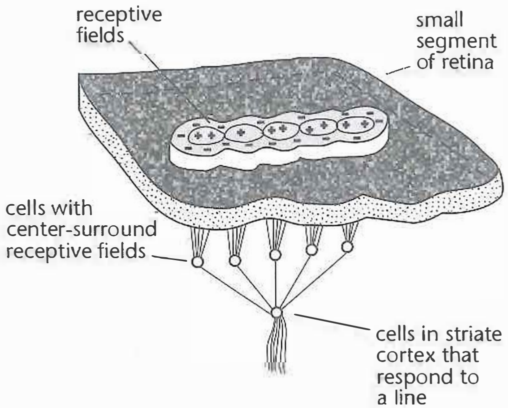

**FIGURE 3.15.** Hypothesized connection of orientation cells in the visual cortex, through ganglion cells to retinal cells. This particular grouping of cells results in cortex cells that are sensitive to linear stimuli. *After Hoffman (1989, Fig. 5.5, p. 57). Adapted by permission of the author.*

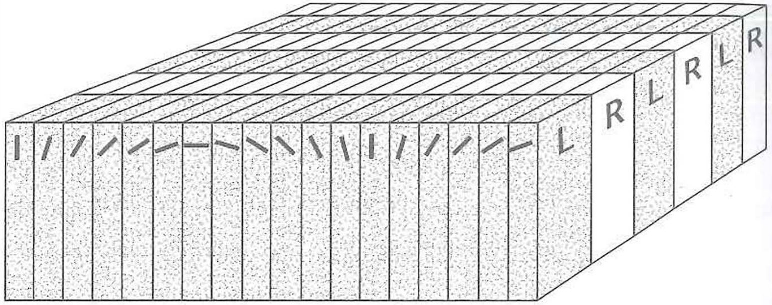

**FIGURE 3.16.** A schematic view of the arrangement of cells in the visual cortex. This depiction of cell arrangement illustrates the dual divisions of cell function for occular dominance and orientation. As Hubel emphasizes, actual cells in the brain are far less regular in arrangement than depicted with this schematic model. *After Hubel (1988, p. 131). From* Eye, Brain and Vision *by Hubel. Copyright 1988 by Scientific American Library. Used with permission of W. H. Freeman and Company.*

## PERCEPTUAL ORGANIZATION AND ATTENTION

If an information-processing approach to vision and visual cognition is accepted as a useful conceptual structure, then derivation of meaning from maps can be viewed as a linked series of processing modules. A number of authors have made this point and offered their versions of how the overall process should be divided. The first such categorization was probably Olson's (1976) level-of-processing approach in which she delineated three levels: (1) comparing symbol pairs, (2) recognizing characteristics of symbol groups, and (3) using symbols as signals to information about what is represented. Both Phillips (1984), with a "low-level"/"high-level" categorization, and Dobson (1985), with a distinction between "visual-search guidance"/"cognitive-search guidance" emphasize a break between preattentive and attentive or perceptual and cognitive processing. Both posit that map design will have more impact on the lowest level of processing because it is at this level that the system is virtually overwhelmed with input and expert knowledge is least likely to apply. As certain optical illusions demonstrate, early perceptual reactions can often be hard (or impossible) to ignore—an indication that top-down processing (i.e., knowledge) has less control at this level and perhaps in some cases no control at all (Figure 3.17).

Although I do not agree with Dobson that cartographers should direct most of their research energy to the influence of symbolization and design decisions on low-level processes—the higher level processes such as derivation of meaning and decision making are what maps are really about—I do agree that failure of a map at this level can make it difficult or impossible to use. A complete understanding of how meaning is derived from maps, then, must begin with an appreciation for the selectiveness of vision in giving us a representation to think about. The *information theory* approach (that treated cartography as a communication system) focused on vision as an information filter with the cartographer's goal being to limit the amount of information that was filtered out in the communication process. This perspective treated perceptual representations as imperfect translations of reality. In contrast, the approach taken here is that perceptual representations are not fuzzy copies of the world, but interpretations of that world. The cartographer's goal is to determine what kind of representation her maps produce and how symbolization and design decisions influence the processes leading to those representations.

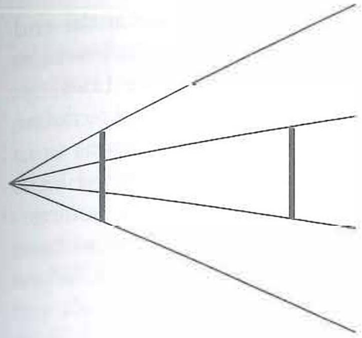

**FIGURE 3.17.** The length of the two vertical lines is identical, but it is difficult to convince yourself of that.

The remainder of this chapter, then, considers these initial visual processes from a perspective of how symbolization and design decisions interact with them.[2]

Gestalt psychologists in the early part of this century laid the groundwork for our current understanding of the perceptual organization of visual scenes. The Gestalt approach emphasized the holistic nature of human reactions to sensation. According to Wertheimer (quoted in Ellis, 1955, p. 2), "There are wholes, the behavior of which is not determined by that of their individual elements, but where the part processes are themselves determined by the intrinsic nature of the whole." More specific attention to pattern is seen in Kohler's (1947, p. 103) statement that "the organism responds to the *pattern* of stimuli to which it is exposed; and that this answer is a unitary process, a functional whole, which gives, in experience, a sensory scene rather than a mosaic of local sensations." Wertheimer's initial emphasis was on defining principles of grouping, and he mentions the segregation of figure and ground only briefly at the end of his article. Kofka and Kohler, Wertheimer's contemporaries, however, extended the initial thoughts on formation of figures. Kohler (1947, p. 145), for example, argued that "sometimes it seems more natural to define a principle of grouping not so much in terms of given conditions as in terms of the direction which grouping tends to take." This viewpoint is specifically linked to figure formation in his statement that "a homogeneous field in visual space is practically uniform and, being without 'points,' there are no relations between 'points' within this field. When Gestalten appear we see firm, closed structures, standing out in lively and impressive manner from the remaining field" (quoted in Ellis, 1955, p. 35).

For several decades while the behaviorists held sway in psychology (particularly in the United States), Gestalt psychology and its principles of perceptual organization were ignored by experimentalists. More recently, particularly in response to the needs of computational vision research and the attention to form and structure in vision that it has stimulated, Gestalt principles are being re-examined. Uttal (1988, p. 146), for example, contends that "human visual perception is powerfully driven by the global organization of form." Recent research in psychology that incorporates Marr's basic contentions (that human vision is an information-processing system and that information-processing systems can only be understood if examined at a combination of computational, algorithm–representational, and hardware levels) have drawn heavily upon Gestalt principles as a source of ideas for understanding grouping in early vision and figure–ground separation associated with object and pattern recognition (see Roth and Frisby, 1986, and Bruce and Green, 1990, for overviews of this work).

Pomerantz (1985) points out that the distinction of Gestaltists, between processes of grouping and of figure–ground separation, is significant from an information-processing perspective. Although there is a clear connection between principles of grouping and formation of figures, grouping of as yet unidentified edges, blobs, terminations, and the like, is a requisite step in deriving a primal sketch. Once edge segments are grouped into contours, then it becomes possible for vision to sort out figure from ground. There is considerable evidence that the initial grouping stages are almost entirely preattentive with little or no input from higher level processes. Research results concerning figure–ground segregation are mixed, with some evidence that figures can spontaneously "pop out" of a background, together with demonstrations that input from stored knowledge or expectations does (in some circumstances) effect both the initial appearance of figure and the stability of the figure–ground relationship.

From a cartographic perspective, low-level issues of grouping seem most relevant for exploratory cartographic visualization in which limited attention will be directed to any one map view and the goal is to notice patterns and relationships. Exploratory visualization implies that the outcome is not know, and that knowledge and expectations therefore may often be absent or wrong. Obtaining an immediate impression, before conscious application of knowledge schemata take over, is likely to play a major role in whether the visual displays lead to insight or simply are used to confirm expectations.

Perceptual organization operating at higher levels is important in those situations where particular information is to be emphasized while other information is suppressed. When goals are to create an imagable map (Peterson, 1987) or to ensure that a particular region becomes the focus of attention (Dent, 1972; MacEachren and Mistrick, 1992), issues of selectivity, associativity, and figure–ground become relevant. Most of the references to Gestalt principles by cartographers have been in relation to figure–ground segregation, with only limited attention paid thus far to the underlying processes of grouping.

## Grouping

The "pattern of stimuli" mentioned by Kohler occurs due to grouping of elements in the sensory field. In relation to Marr's theory, these elements might be edge parts, blobs, and the like. For maps, viewing these edges and "blobs" will occur in the primal sketch representation in response to map symbols such as points, lines, or elemental parts of textures. Their grouping will determine whether symbols are seen as intended and which kinds and scales of patterns are noticed. Wertheimer (1923; translated in Ellis, 1955) set out the rules for such perceptual grouping in his classic paper, "Laws of Organization in Perceptual Forms." He defined the following factors or rules:

1. *Proximity*: Objects close together form groups. In the abstract, the factor holds that in any array of individual elements, those that are closest together will be seen as part of a group (Figure 3.18). Cartographically, as detailed below, proximity has been postulated to account for the appearance of regions on maps (Figure 3.19). A particularly intriguing part of Wertheimer's argument, in light of current interest in dynamic and animated maps, is his contention that the factor of proximity holds for sound as well as sight. Sounds close together in time will form perceptual groups. This issue (without reference to Gestalt principles) is alluded

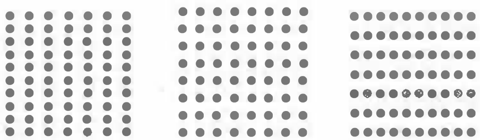

**FIGURE 3.18.** Grouping by proximity: rows (far right) and columns (far left) of dots versus evenly spaced dots (center).

to by Krygier (1991) in his identification of the audio variable *rhythm* as "the grouping and ordering of sounds."

2. *Similarity*: Like objects form groups. As presented by Wertheimer, similarity relates to nonlocational characteristics of perceptual units. He specifically mentions color, form, and sound. From a cartographic perspective, we might consider the similarity of all graphic variables (color hue, color value, shape, etc.), as well as tactual and audio variables (Figure 3.20). Wertheimer (1923; translated in Ellis, 1955) points out that similarity is not absolute, but can occur in degrees. Thus, judging "more and less dissimilar" becomes an issue in how people experience map symbols.

3. *Common fate*: Objects moving together are seen as a group. For this factor, Wertheimer points out that already grouped units that move together may hardly be noticed, but that units from separate groups moving together can be "confusing and discomforting" and will most certainly be noticed. It is posited that common movement of units from separate (static) perceptual groups will override proximity, similarity, or other factors to achieve a new group held together by their "common fate." Cartographically, of course, this factor applies only to animated or dynamic maps. In this context, however, it may play a particularly strong role in what groups are perceived. A corollary to Wertheimer's common fate in relation to map animation is that objects that change together (even when they do not move) are seen as a group. In our map animation research at Pennsylvania State University, we have used this principle to animate static maps that depict existence of a feature with a fixed location (Figure 3.21). Similarly, Monmonier (1992) has employed what he called "blinking" as a method to emphasize the spatial pattern (or lack of it) in the proportion of public officials who are female. Blinking, in this context, involves having a choropleth map class (with values for the United States grouped by quintiles) blink on and off while other classes are turned off. Thus, one at a time, the states in successive quintiles are visually grouped so that the viewer can easily identify regional patterns in exclusion of females from public office.

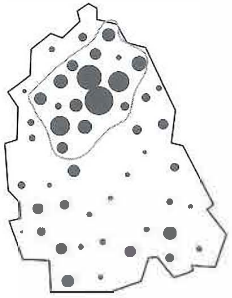

**FIGURE 3.19.** The importance of proximity in region identification on graduated circle maps. The outlined circles represent the consensus region identified by Slocum's subjects. *After Slocum (1983, Fig. 9, no. 14, p. 71). Adapted by permission of the American Congress on Surveying and Mapping.*

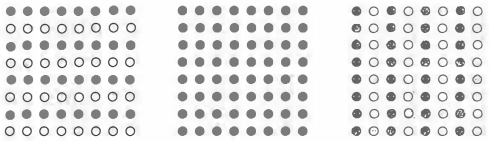

**FIGURE 3.20.** Grouping by similarity: white versus black circles.

4. *Pragnanzstufen*: Perceptual groups are characterized by regions of "figural stability." This factor is difficult to translate, but implies that grouping has discrete cases. In relation to proximity, for example, there will be a relative threshold distance at which units will be seen to group or to occupy space in an undifferentiated way. Wertheimer's example sug-

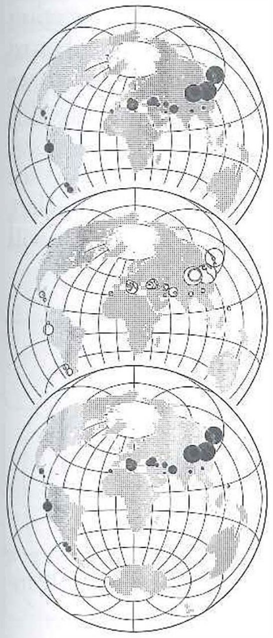

**FIGURE 3.21.** A sequence of maps simulating earthquake epicenters blinking on and off to highlight their clustering. *Reproduced fom DiBiase et al. (1992, Fig. 3, p. 207). Reprinted by permission of the American Congress on Surveying and Mapping.*

gests that for a row of dots we will see an ab–cd–ef–gh–ij grouping, no grouping at all, or an a–bc–de–fg–hi–j grouping at different possible regular spacings of the dots (Figure 3.22). This concept seems to match with anchor-effect theories of magnitude estimation (discussed below) and ideas about prototype categories (discussed in the next chapter).

5. *Objective set*: With change, there will be a tendency toward stable groups. Following from the above factor, the idea here is that if a set of perceptual units is initially seen as a group and that over time the position of those units changes, perception will try to retain the initial group. In addition, there will be a tendency to see a limited number of states (e.g., grouping A, undifferentiated scene, grouping B). Wertheimer's example is based on a scenario in which seven pairs of dots gradually change relative positions (refer to Figure 3.22). Again, cartographically, this factor applies to animation. The implication is that throughout a movement perception tries to maintain a stable state, resulting in a greater likelihood that we will see a constant grouping on a set of change maps if they are presented dynamically than if they are presented on a page as small multiples. This would be an interesting hypothesis to test. The possibility to be concerned with is that "a certain (objectively) ambiguous arrangement will be perfectly definite and unequivocal when given as part of a sequence" (Wertheimer, 1923; translated in Ellis, 1955, p. 80). The issue here is one of visualization quality and how to determine when a pattern is "real" or illusory (MacEachren and Ganter, 1990).

6. *Good continuation*: Elements that follow a constant direction group. This factor applies not only to straight-line arrangements, but to curves, as illustrated in Figure 3.23. Cartographically, this factor allows contours on a black and white map to be seen as separate curved lines differentiated from roads or rivers that they might cross (Figure 3.24).

7. *Closure*: Closed objects form wholes. There is a tendency to see bounded perceptual units as wholes. Even when bounding edges overlap, there is a likelihood that the factor of good continuation cited above will allow us to see the separate bounds as units and apply closure to isolate their edges as groups defining wholes. The critical role of good continuation, and its potential dominance over closure was dramatically illustrated by Wertheimer (1923; translated in Ellis, 1955) (Figure 3.25). Cartographically, closure has clear applications to situations such as graduated symbol maps where it has been demonstrated that circle overlap does not prevent readers from seeing the circle segments as whole circles, or from judging circle size (Groop and Cole, 1978). In addition, a variety of graphic methods for emphasizing the closure of a map region have been examined.

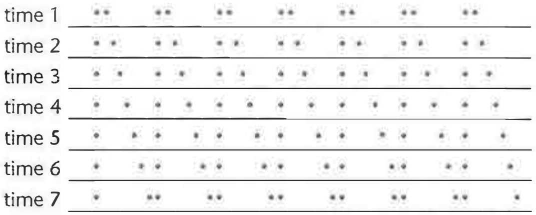

**FIGURE 3.22.** Grouping due to Wertheimer's Pragnanzstufen factor. If shown as a time series, the original groups (top row) will be seen at time 4, even though all distances are equal at this time.

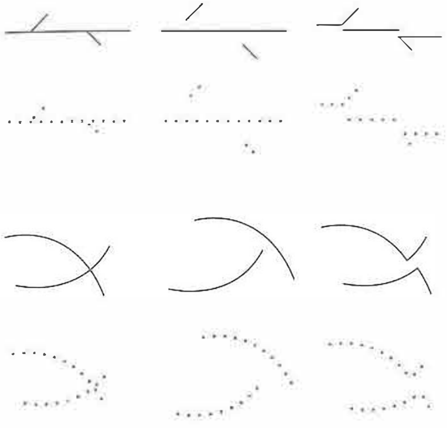

**FIGURE 3.23.** Grouping by good continuation. On the top, we "see" a long line with two short lines attached, rather than a short line with two angular attachments. On the bottom, we "see" two smooth curves crossing.

8. *Simplicity*: Objects will group in the simplest form. Wertheimer did not specify simplicity as a specific factor, but mentioned it in relation to what he called "good Gestalt." This concept was a basis of

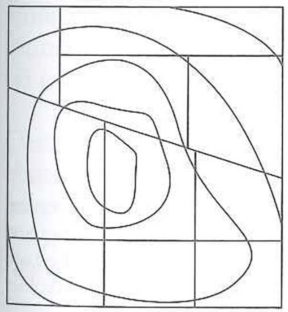

**FIGURE 3.24.** Good continuation helps map viewers sort out intersection lines on maps. Even in the absence of other contrast, we can visually separate the contours from the county boundaries.

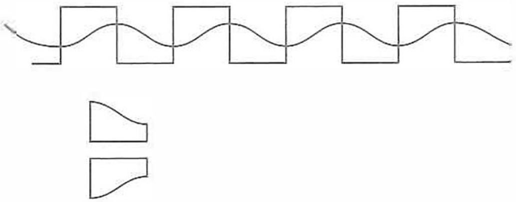

**FIGURE 3.25.** Dominance of good continuation over grouping by closure. Few observers are likely to interpret the figure above as a set of the irregular shapes shown below. Instead, we see a curved line across an angular one. *After Ellis (1955, Fig. 13, p. 82). Adapted by permission of Routledge & Kegan Paul.*

Wertheimer's "Law of Prägnanz," which Koffka (1935, p. 138) described as follows: "Of several geometrically possible organizations that one will actually occur which possesses the best, simplest and most stable shape." An example relevant to cartography is found with the tendency to interpret ambiguous situations (such as Figure 3.26) as interposition of simple figures rather than more complex adjacent figures.

9. *Experience or habit*: Familiar shapes or arrangements form groups. Many subsequent authors have focused on Wertheimer's contentions that past experience was not essential to perception of groups and that proving a role for past experience would be difficult. As a result, these authors have (mistakenly) characterized Gestalt psychology as disallowing the possibility for knowledge to play a role in both perceptual grouping and figure–ground perception. Wertheimer did, however, contend that experience or habit, in the form of "repeated drill," could play a role and at times cause groups to be seen that are at odds with what the other factors might dictate. Although he placed more emphasis on preconceptual processes, Wertheimer did not rule out the possibility of what we consider in the next chapter as "knowledge schemata" playing a role, even at early low-level stages of visual processing.

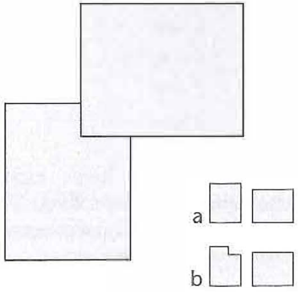

**FIGURE 3.26.** Grouping by simplicity. It is easier to see two squares overlapping (a) than a square next to an L-shaped region (b)—in spite of the fact that the latter could represent a common geographic feature such as Utah and Wyoming.

A number of cartographers have cited the above Gestalt "laws" (Wood, 1968, 1972; Dent, 1972; McCleary, 1981). For the most part, they have been treated as laws, with attention directed to devising logical guidelines for incorporating the laws into map design. Few cartographers have questioned the principles or considered their relative influence on grouping of map elements. This tendency to take the Gestalt laws for granted is even apparent among psychologists (e.g., Kosslyn, 1989, uses some of the laws as given in developing a procedure for assessing graphic acceptability). Pinker (1990), like Kosslyn, contends that Gestalt principles have a role in the process of translating the initial visual scene to a visual description of a graph (a representation of entities and relationships among those entities). He goes on to suggest, however, that we do not at this point understand how to apply these principles because there has been little empirical research about the situations in which they hold or their relative importance.

Among psychologists, Pomerantz (1985) has provided perhaps the most explicit analysis of Gestalt grouping principles and their interrelations, as separate from the issue of figure–ground segregation. He begins with a convincing argument that grouping is "logically prior to figure–ground segregation" (p. 128). We must group perceptual units into objects and regions before a choice can be made among objects or regions concerning which are the focus of attention.

Although no cartographers, to my knowledge, have explicitly mentioned Gestalt principles in relation to perceptual grouping of map elements, a few have incorporated the principles in their work without crediting them to Gestalt psychology. Olson (1976) alludes to the cartographic importance of grouping with her three-tiered hierarchy of mental processing in map use. Her second level deals with recognizing properties of symbol groups. To recognize these group properties, of course, the visual process must provide groups for which properties can be compared by higher level processes. Olson considers (but does not test) the possible impact of symbol scaling (for graduated circle and dot maps) on regions that might be identified, as well as the role of value contrast among different symbols and between them and the background. In the former examples, the variable of proximity is manipulated and in the latter case similarity is used.

In a somewhat more direct examination of the applicability of Gestalt grouping laws to map reading, Slocum (1983) investigated visual clustering on graduated circle maps. Slocum's stated goal, formulated in behavioral terms, was to develop a method to predict perceived map groups using a combination of the psychological principles he felt were relevant to the problem. He hypothesized that proximity, similarity, and good continuation would play a role in the visual groups seen on graduated circle maps.[3] He was unable to devise a measure of good continuation in the absence of eye movement recordings, so it was not actually tested. In addition to grouping factors, Slocum hypothesized that "figure–ground" would play a role in visual grouping. Figure–ground was limited for purposes of his experiment to a measure of value contrast, with dark areas expected to be seen as figure. The incorporation of this measure was based on prior evidence by Jenks (1975) that value difference had an effect on the groups seen, and its interpretation in terms of "figure–ground" seems to be based on Dent's (1972) emphasis on value contrast as a figure–ground variable.

Slocum's (1983, p. 61) experiment involved having subjects outline sections of graduated circle maps that they "saw" as "visual clusters"—"groups of circles that appeared to belong together and form a visual unit." His analysis indicated that a combination of proximity and figure (defined as relatively dark sections of the maps) provided a reasonable prediction. In fact, 92% of the individual circles on his 10 test maps were correctly classified as in or out of a cluster. Similarity of circle sizes had virtually no effect upon groups seen (Figure 3.27).

Eastman (1985b) has also investigated an aspect of perceptual grouping but, unlike Slocum, did so following a cognitive information-processing approach. Specifically, he examined the effect of several design variations of a typical reference map on the perceptual organization of the map.[4] The goal was to determine whether proximity of locations, similarity of symbolization for those locations, regional inclusion (which can be associated with the closure of country boundaries or road segments), or linear linkage (determined by road connections between cities and at least loosely associated with "good continuation") influenced how map items were grouped in memory. Eastman found that the maps stimulated five different groupings, each of which was primarily associated with one or two of the map designs. All five grouping strategies led to hierarchical memory structures. A comparison of subject groups (that grouped map elements differently) did not support Eastman's hypothesis that different grouping strategies would lead to differences in learning speed or memory accuracy. Graphic organization of the maps, however, did appear to have an impact, both on how maps were perceptually organized and on whether organization was easy. Lack of graphic organization led to grouping by proximity or horizontal partitions. A map with regions delineated led to regional chunking (Figure 3.28). The map with no graphic organization also proved to be much harder to learn than the rest.

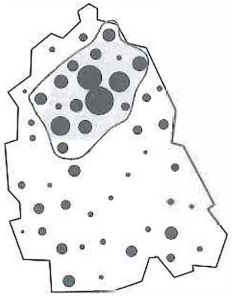

**FIGURE 3.27.** Grouping as predicted by Slocum (gray region) in comparison to grouping produced by his subjects (outlined region). *After Slocum (1983, Fig. 9, no. 13, p. 71). Adapted by permission of the American Congress on Surveying and Mapping.*

In addition to these relatively direct applications of Gestalt grouping principles to cartography, grouping has been considered less directly in studies of map regionalization. Muller (1979), for example, demonstrated that map viewers arrived at similar regions or groups when asked to delineate regions of high, medium, and low population density on continuous tone choropleth maps. In spite of being presented with many more color values than could be discriminated, subjects were able to group similar values into categories in a consistent way. In contrast to Muller, who asked subjects to delineate high, medium, and low regions, McCleary

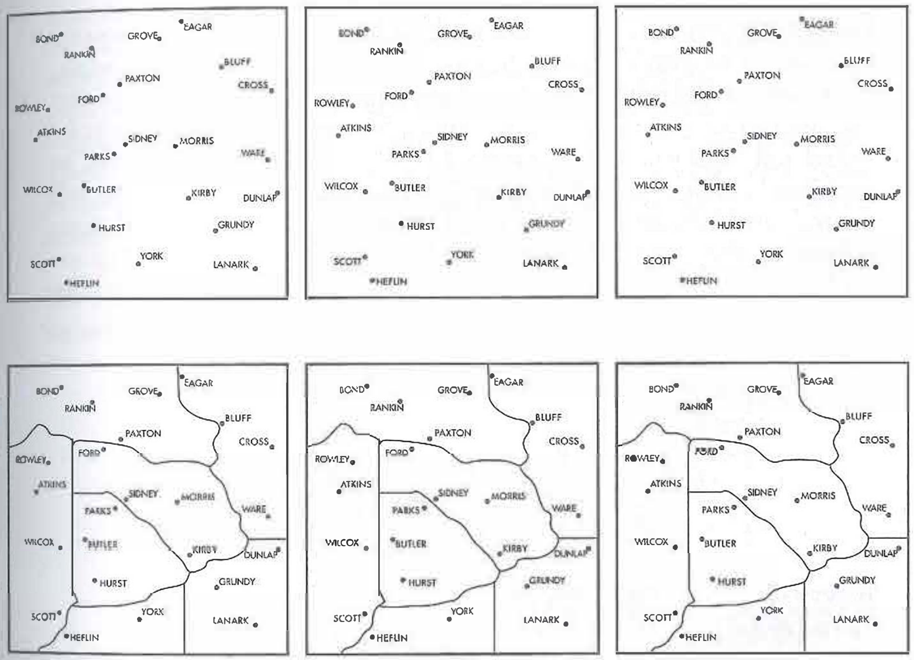

**FIGURE 3.28.** Eastman's graphically undifferentiated map (top) compared to the map leading to the most consistent grouping strategies (bottom). For subject groups viewing each map (at left), the gray shading represents consensus first- and second-order chunks (middle maps and right maps, respectively). *After Eastman (1985b, Figs. 2, 8, 9, 12, and 13, pp. 5, 15, 16). Adapted by permission of the American Congress on Surveying and Mapping.*

(1975) had subjects delineate any regions they saw on a set of dot maps. An intriguing result of this task was that although grouping by proximity seemed to be at work in all cases, his subjects fell into two quite distinct types that he termed "atomists" and "generalists" (Figure 3.29). The atomists focused on local details. According to McCleary (1975, p. 247), they "seemed obsessed with detail and may have lost sight of the overall pattern of density." For the generalists, on the other hand, "lines are schematic and the 'attitude' expressed by the boundary line drawn suggests a reductionist view of the image." This finding has not been pursued in the cartographic literature but has interesting implications for our current concern with use of cartographic visualization for exploratory data analysis. It is important to determine whether McCleary's atomists and generalists represent general categories of map viewers and whether these tendencies are altered with training or expectations.

## What We Attend To

Perceptual grouping is thought to work, at least in part, at a preattentive level. Based on Marr's speculations, some amount of grouping (into edges, blobs, etc., of the primal sketch) is a prerequisite to all seeing. Grouping will interact with visual attention in complex ways. Where our gaze is directed will limit what can be grouped (only global features of a scene in peripheral vision vs. details in central vision). The results of grouping will control what can be attended to and where our gaze might travel next. Where we direct our attention can, of course, also be consciously controlled. As a complement to issues of grouping, then, we must consider

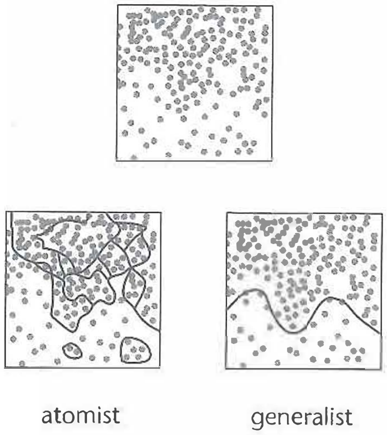

**FIGURE 3.29.** Sample subjects from McCleary's dot map regionalization experiment illustrating the grouping strategies of atomists (left) and generalists (right). *Reproduced from McCleary (1975, Fig. 4, p. 246).*

the combination of processes that fall under the heading *visual attention*. An important issue that Wertheimer considered in relation to grouping is the possibility of more than one factor acting at the same time. Such interaction may enhance visual grouping or may act in opposition to inhibit it (Figure 3.30). In addition to the effect on grouping, the interaction of multiple variables of perceptual units can influence the separability of features of the unit. This has obvious implications for how multivariate symbols are perceived, particularly for which aspects of a multivariate map symbol we can attend to together or separately.

### *Selective Attention and Separability of Visual Dimensions*

Recent research on perceptual organization has emphasized the notion of selective attention as a way to measure the role of different features in the visual scene on perceptual grouping (Pomerantz, 1985). "Selective attention" refers to the ability to attend to one dimension of a display and ignore another. If dimensions or variables can be segregated in this way, they are not grouped. If, on the other hand, it is difficult or impossible to selectively attend to the separate dimensions, they are considered to be perceptually grouped. In a series of experiments, Pomerantz and his colleagues examined selective attention to features of compound stimuli. Their results, in addition to informing us about general perceptual processes relevant to map reading, are likely to be particularly relevant to design of multivariate symbols for maps.

Many of Pomerantz's experiments used sets of simple parenthesis-like symbols that were paired in various ways. These pairings were designed so that some should lead to groups (based upon Gestalt principles

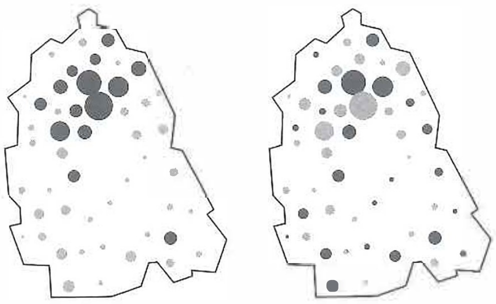

**FIGURE 3.30.** Similarity and proximity acting together to enhance grouping (left) and in opposition resulting in ambiguous grouping (right). *Derived from Slocum (1983, Fig. 9, no. 13, p. 71).*

of good continuation, similarity, symmetry, and proximity) and others should resist grouping. One set of stimuli are shown in Figure 3.31.

A typical experiment would match a control case in which subjects had to sort two stimuli (e.g., the top row of each box in Figure 3.31) versus a selective attention case in which subjects had to sort all four stimuli (e.g., the two pairs on the right of each box in Figure 3.31 in one category and the two pairs on the left in the other) (Pomerantz and Garner, 1973). In both cases the task could be completed by focusing on only the left-hand element of the parenthesis pair. If subjects could selectively attend to this element and ignore the other, both groups should accomplish sorting at the same rate. Subjects in Pomerantz's selective attention group, however, took longer to sort their stimuli. This is an indication that the pairs were processed as groups. The control case subjects had the easy task of sorting these perceptual groups into a symmetrical and an asymmetrical category. The selective attention group was forced to treat the four stimuli separately because, as groups, the two columns of parenthesis pairs do not form Gestalt categories (in fact, symmetrical vs. asymmetrical units form categories counter to the ones required by the sorting task). When the same experiment was run with stimuli that should, according to Gestalt principles, not group, there was no difference in response time between the experimental groups. The parentheses did not form perceptual units, and therefore the right-hand parenthesis could be ignored and sorting accomplished by focusing attention on the left parenthesis only. Thus both groups had a two element categorization task and completed it at the same rate.

Cartographically, Bertin (1967/1983) has focused upon issues similar to those that interest Pomerantz, but has not investigated his contentions experimentally. Bertin's hypothesis (which he treats as fact) is that the visual variables can be independently judged on the basis of what he calls *selectivity* and *associativity*, and that these designations are discrete (i.e., a visual variable is either selective or nonselective in all applications). His selectivity is similar to Pomerantz's selective attention. Where Pomerantz focuses on whether conjunctions of two or more objects proximate to one another are seen as a whole (a group), Bertin is interested in whether objects (map symbols) spread across the map can be formed into visual groups. Visually grouping, or attending selectively to, a particular value, for example, seems easier than attending to a particular shape (Figure 3.32). Bertin's concept of selectivity is limited to grouping by similarity (although he does not define it in these terms). The emphasis is on whether visual grouping is "immediate" (a term that can probably be taken to mean preattentive) for all symbols in a category identified by a specific variation of one visual variable (e.g., all blue symbols on a map compared with symbols in various hues). Bertin posits that location, size, color value, texture, color hue, and orientation (of point and line symbols only) are selective variables.

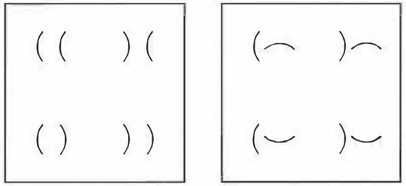

**FIGURE 3.31.** Test stimuli with good (left) and poor (right) grouping. *Reproduced from Pomerantz (1973, Fig. 6.1, p. 129). Copyright 1973 by the Psychonomic Society. Reprinted by permission of the author.*

There is empirical evidence for some of Bertin's claims (although not derived from explicit attempts to test those claims). In relation to orientation, for example, Olson and Attneave (1970) demonstrated that a difference in orientation of simple line symbols can cause regions to be discriminated quickly (Figure 3.33). There is even a neurological (hardware level) explanation for why orientation is selective. Research by DeYoe et al. (1986) with monkeys has demonstrated that there are cells in the monkey's cortex (regions V1 and V2) that respond to pattern edges defined by differences in orientations of the texture elements making up the patterns. For this differentiation to occur, orientation differences must be in the center and surround portions of the cell's receptive field.

Nothdurft (1992) found that with limited variation within pattern areas, differences in orientation of as little as 20% were sufficient for a 75% success rate for preattentive pattern segregation. As variability in orientation of individual elements making up the pattern increased, the necessary difference in mean orientation of line segments in the two regions (required to achieve a 75% preattentive selection rate) increased in a roughly linear fashion. Beyond 30% variability in orientation within the individual patterns, pattern discrimination was unsucessful regardless of magnitude of between-pattern orientation difference.

In contrast to Bertin's sweeping claim, other evidence exists that the

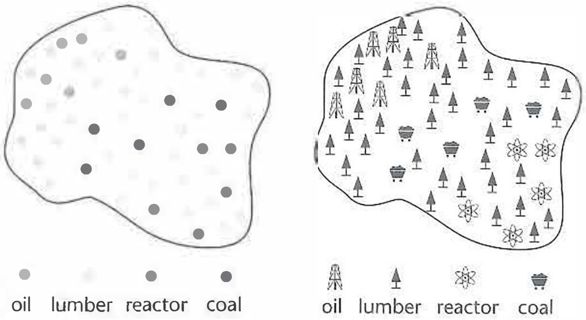

**FIGURE 3.32.** Value (right) seems to be selective while shape (left) is not.

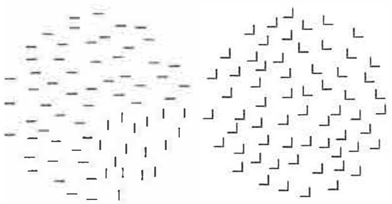

**FIGURE 3.33.** Differences in orientation result in visual groups, but differences in alignment do not. *Derived from Bruce and Green (1990, Fig. 6.16, p. 117).*

key selective variable is symbol slope rather than orientation. If a symbol is used that has an internal orientation, it becomes clear that we can distinguish (mathematically) between orientations that are 180° apart, but that this orientation difference is not selective in Bertin's terms (Figure 3.34). With Bertin's line segment examples it was not obvious that 180° rotations were not selective because they could not even be detected.

Not all of Bertin's visual variables have been tested for selectivity, and the only empirical tests thus far have been by psychologists (who have not specifically set out to test graphic variables but to study the phenomenon of selective attention). In addition to slope, there is evidence for selectivity of color hue and color value. For both, Julesz (1975) found that regions were easily segregated if distinct value or hue differences exist. An interesting factor for these visual variables was that when pattern elements are small, vision seems to respond to an average signal. A region of mostly black and dark gray squares having a few white and light gray ones mixed in (and as a result mostly dark) is easily segregated from a region of mostly white and light gray squares having a few dark gray and black ones mixed in. Similarly, wavelengths of colors seem to be averaged so that a region of red and yellow squares (and a few green and blue) is clearly discriminated from one of green and blue squares (and a few red and yellow). A red–green region, however, is not easily discriminated from a blue–yellow region.[5] Evidence also exists to support Bertin's contention that shape is not selective, at least in the case of different shapes that have the same number of line segments and terminators (Figure 3.35).

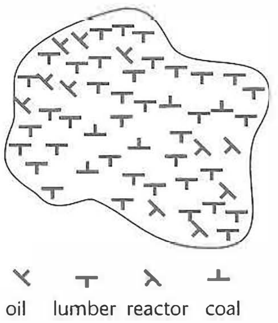

**FIGURE 3.34.** It appears that slope of parts rather than orientation of an overall shape must differ (in some cases) for orientation to be selective in Bertin's sense.

At least one graphic variable that Bertin ignored has also been demonstrated to be selective. Julesz (1965) demonstrated that what he called "granularity" of a pattern (discussed in Part II of this book as pattern arrangement) leads to easy segregation of regions. The general success of Bertin's selectivity claims, along with some discrepancies uncovered by empirical research in psychology, suggests that cartographers need to take a closer look at Bertin's ideas. Studies that empirically test Bertin's hypotheses and investigate the magnitude of differences required along specific dimensions of visual variables (including additional variables that others have added to Bertin's original set) are clearly called for.

Bertin considers visual variables largely in isolation and does not discuss their potential interaction on a map. Experiments along the lines of those conducted by Pomerantz might be used to determine how vision will react to multivariate symbols that are designed to convey redundant information for emphasis and to enhance discrimination or separate information so that interrelationships can be noticed. In the first case, we would want to apply visual variables for which selective attention is difficult; in the latter case we would want the opposite.

No cartographic research (to my knowledge) has been conducted in relation to the issue of selective attention to visual variables in multivariate map symbols. Shortridge (1982) has, however, provided an overview of evidence from psychology and suggested possible applications to map symbolization. In particular, she considered the issue of *integral* versus *separable* dimensions (i.e., visual variables). Separable dimensions are ones for which selective attention is easy; integral dimensions tend to be seen

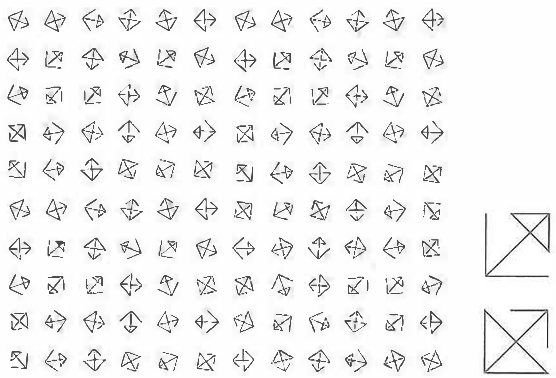

**FIGURE 3.35.** Shape, with other variables held constant, is not selective. *After Julesz (1981, Fig. 6, p. 95). Copyright 1981 by Macmillan Magazines Limited. Adapted by permission of the author and* Nature.

as wholes, and therefore selective attention is hard. As an example, consider a map that uses line size to indicate temperature at weather recording stations and line orientation from horizontal to vertical to indicate precipitation amount (Figure 3.36). If the two dimensions (e.g., symbol size and line orientation) are separable, selective attention will be possible and a viewer should be able to compare two stations on temperature or precipitation quickly—and not be able to judge temperature–precipitation correspondence easily (a contention that seems to be supported by Figure 3.36).

As Shortridge (1982) points out, psychologists began to distinguish between integral and separable dimensions as a way to explain results of visual search tasks that sometimes indicated processing of multiple stimuli in parallel and sometimes in a serial self-terminating manner (i.e., one symbol at a time until a target is found, at which point processing is halted). For serial searches, if a target is not present the search is exhaustive (relatively long) and will increase in length as the number of stimuli in the scene is increased. When a target is present (and a serial search is used), it will be found (on average) after half of the stimuli have been processed (response times will be 0.5 times that of target-absent cases). If stimuli are processed in parallel (all at once) then processing times will not be effected by the number of stimuli that must be processed. Although predicting whether a serial or a parallel process will be invoked does not seem easy, attention to this question led psychologists to notice the differences between compound stimuli that seemed to differ in the likelihood of serial versus parallel processing. Recognition that some symbol dimensions are integral (i.e., difficult or impossible to attend to separately) led to a *holistic* account of processing as an alternative to the serial–parallel possibilities. This account suggests that integral symbol dimensions create a whole that is processed as a single unit. Evidence for such holistic processing comes from research by Lockhead (1970) and Pomerantz and Schwaitzberg (1975), both of whom found that certain conjunction tasks, in which subjects had to discriminate or categorize symbols on the basis of the conjunction of two dimensions, were performed faster than discrimination or categorization on the basis of either dimension individually.

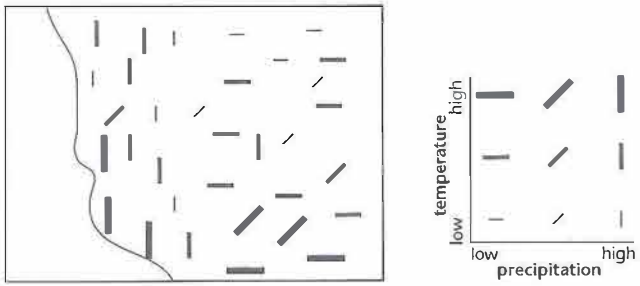

**FIGURE 3.36.** A map of temperature and precipitation using symbol size and orientation to represent data values on the two variables.

### *Divided Attention and Variable Conjunctions*

In their research on conjunction tasks Pomerantz and Schwaitzberg (1975) used measures of *divided attention* as a complement to previous selective attention studies of perceptual grouping.[6] They reasoned that if selective attention to parts failed, implying that they were grouped, viewers should find it easier to attend to the groups as a unit. This hypothesis was tested by having subjects try to sort stimuli (of the kind used in their initial experiments—see Figure 3.31) according to groups that were similar while ignoring elements within those groups that would suggest alternative sortings. They found that when Gestalt attributes of element pairs (e.g., combinations of symmetry, closure, similarity) indicated that grouping of the elements into a single whole was likely, sorting by groups was faster than sorting by individual element. When, on the other hand, individual elements did not form "good" Gestalt groups, sorting by individual elements was easy and sorting by group was extremely difficult.

The above evidence indicates that various combinations of map symbol attributes may lead to integral or separable symbol dimensions which in turn may facilitate divided or selective attention. Knowing which will occur in particular cases is clearly crucial to making effective map symbolization choices. Integral combinations should be useful in univariate map applications where the goal is to enhance discrimination while reinforcing appearance of order for quantitative information. One example would be the combination of color value and saturation for area fills on a choropleth map of population density. By combining these variables to produce a wide range of area fills (e.g., from a light, desaturated blue to a dark, fully saturated blue), it may be possible to extend the practical number of categories that can be used. Multivariate symbols with separable dimensions, on the other hand, seem suited to the depiction of multivariate data (either qualitative or quantitative) in which the viewer will want to extract various components of the data separately. Examples include the temperature–precipitation map cited above or a map showing relationships between soils and geology (e.g., Wakarusa quad; Campbell and Davis, 1979). In the latter case, color was used for one variable and pattern for the other. Each was, in fact, a combination of visual variables and neither the combinations nor the conjunction of the color–pattern sets have been examined for selective attention.

In response to our current lack of knowledge concerning how visual variables interact in multivariate symbols, Shortridge (1982) suggests a program of research to evaluate whether specific combinations of visual variables combine in integral versus separable ways. She considers creating a classification scheme based on symbol properties a useful goal. In addition, she presents a hypothesis that integral versus separable conjunctions of visual variables may not be discrete categories, as presented in most psychological literature to date, but may be two ends of a continuum. This proposal allows for some level of integrality to occur between size and color value, a conjunction that Shortridge used with graduated circles to demonstrate the potential advantages of variable redundancy with a quantitative map sequence (but one that psychologists have labeled as separable). Dobson (1983) provides evidence that this particular conjunction of variables (size and value) does improve processing over using size alone. Dobson conducted three experiments in which subjects viewed a graduated circle map of the western United States and responded to tasks requiring location (counting the number of states in a particular category), categorization (identifying the category for a particular state), and comparative judgment (determining which of a pair of states had the higher data value). A control group viewed a map in which black circles scaled by area was presented and the redundant symbol group viewed a map of the same data in which color value as well as circle area was used to represent data values (Figure 3.37). Response times as well as accuracy of responses both indicated significant processing improvements for the size–value conjunction over size alone, an indication that those variables are at least partially integral.

Some psychologists working with integral versus separable conjunctions to study perceptual grouping have recognized a third category of conjunctions that fits between integrality and separability (Pomerantz and Garner, 1973). This intermediate category is termed "configural." Where integral conjunctions refer to two physical dimensions that correspond to a single perceptual code and separable conjunctions refer to two physical dimensions that lead to distinct perceptual codes, configural conjunctions maintain separate perceptual codes, but also code a relational or "emergent" dimension. Both integral and configural dimensions lead to "filtering interference" (interference of the second, nonrelevant attribute in tasks requiring attention to only one attribute) and "condensation efficiency" (improvement on tasks requiring both attributes to be considered as a unit). Integral dimensions differ from configural ones, however, in exhibiting "redundancy gains" (improvements in speed of performance on tasks in which both attributes provide the same information).

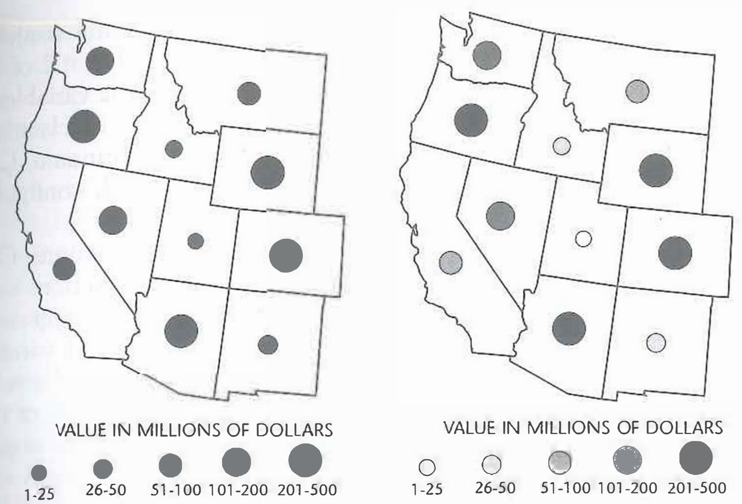

**FIGURE 3.37.** A pair of Dobson's maps. *Reproduced from Dobson (1983, Fig. 7.1, pp. 156–157). Reprinted by permission of John Wiley & Sons, Ltd., from* Graphic Communication and Design in Contemporary Cartography. *Copyright 1983 by John Wiley & Sons, Ltd.*

Based on the above definitions, the size–value conjunction that improved performance on Dobson's experiments would be considered integral, but evidence from at least five psychological studies that Shortridge (1982) cites indicates that size and value are at least configural (if not separable). One difference between the psychological studies and Dobson's research is that Dobson's subjects had to assign stimuli to one of five rather than one of two categories. Another difference was that Dobson's subjects had to locate a named state and its circle from the map display containing 11 circles, while the subjects in the psychological studies only saw one stimulus at a time. The apparent redundancy gain in Dobson's experiment may therefore be associated with search time rather than with categorization time. Another possibility, of course, is that Shortridge's continuum hypothesis is correct and that a size–value conjunction is somewhere between the separable and integral extremes.

The concept of an integral–separable continuum of symbol conjunctions has found some support in the psychological literature. Cheng and Pachella (1984) in particular have argued that most phenomenon already categorized as integral or separable actually exhibit "degrees of nonseparability." Further support comes from a recent study by Carswell and

Wickens (1990). They examined 13 stimulus sets involving conjunctions. All were derived from existing graphics. They found that 2 of the 13 commonly used symbol conjunctions contained separable variables, 2 contained configural variables, and the other 9 could not be classified. Rather than interpreting their results as support for a continuum, Carswell and Wickens favor three distinct categories: integral, configural, and separable.

In addition to examining separability of variable conjunctions, Carswell and Wickens (1990) considered whether or not conjunctions were *homogeneous* or *heterogeneous* and whether they used object integration. Homogeneous conjunctions are those in which the same visual variable (e.g., location in space, as on a graph, or orientation as in a wind rose) is used for both (or all) variables. Object integration is the merging of two attributes into a single object (Figure 3.38). Garner (1976) has argued that object integration is more likely to lead to integral or configural conjunctions than will two distinct spatially contiguous objects (e.g., paired bars on a bar chart). Following from these ideas, we might expect that the Carr et al. (1992) bivariate $NO_3$–$SO_2$ map (which uses homogeneous conjunctions and object integration) would result in configural conjunctions for which individual attributes and their relationships can be easily extracted from the line slopes, their direction agreement (both up, both down), or the angle between them (Figure 3.39).

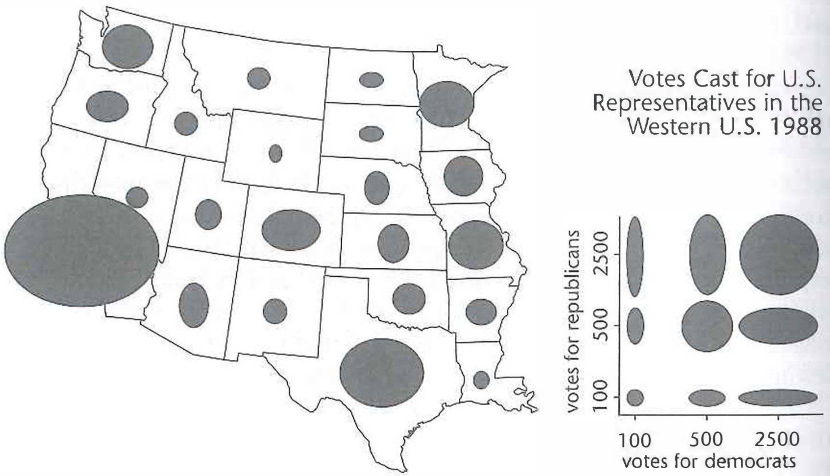

**FIGURE 3.38.** An example of the use of an ellipse as a map symbol in which the horizontal and vertical axes represent different (but presumably related) variables.

**FIGURE 3.39.** Bivariate map of $NO_3$ and $SO_4$ trends. The original Carr et al. version of this map used a wheel with eight spokes, rather than a simple dot, as the center of each glyph. When large enough, this added feature facilitates judgment of specific values. *After Carr et al. (1992, Fig. 7a, p. 234). Adapted by permission of the American Congress on Surveying and Mapping.*

### *Associativity of Graphic Variables*

As described in Chapter 2, *associativity* exists for a visual variable if variations within that variable (or, in Bertin's terms, the "levels" of the dimension) can be ignored, allowing the units using that visual variable to form a perceptual group. Bertin demonstrates the difference between associative and disassociative variables with a bivariate map composed of point symbols that vary in size (which he considers a disassociative variable) along one axis and shape plus orientation (a pair of associative variables) along the other (Figure 3.40). As is clear here, and for Bertin's original somewhat more complex conjunction of three variables, it is easier to attend to different shapes of the same size than different sizes of the same shape. Bertin's claim is that different levels of particular visual variables retain sufficient similarity that symbols to which these various levels are assigned can be seen as a visual group regardless of proximity. Bertin contends that for his associative variables, this grouping will occur "immediately."

Just as Bertin's (1967/1983) contentions about the selectivity of the visual variables are related to psychological work on selective attention,

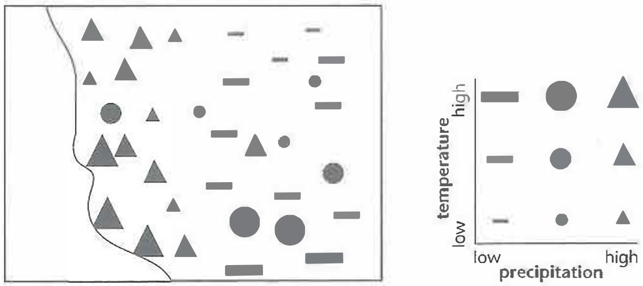

**FIGURE 3.40.** The bivariate temperature–precipitation map of Figure 3.36, this time using point symbols that vary in shape and size to represent the two quantities.

his arguments concerning associativity are related to research on divided attention. In the case of Pomerantz and Schwaitzberg's (1975) divided attention study, divided attention was easy for pairs of shapes that were in close proximity and formed Gestalt groups. As distance between the elements increased, however, attention to the feature pairs as units became more and more difficult (after 2° of arc separation, response times for divided attention rise markedly). This evidence makes Bertin's contentions about associativity seem unlikely. At the least, associativity will depend upon proximity, decreasing as proximity among symbols increases. At this point, we have no information to suggest the shape of this relationship (whether it might be linear, geometric, or stepped with one or more thresholds), nor do we know whether the associativity–proximity relationship will look the same for all of the visual variables that Bertin claims are associative. Shortridge (1982) suggested that we examine whether the integrality or separability of pairs of visual variables is a discrete phenomenon or is better represented as a continuum. We should perhaps extend this suggestion to all aspects of visual variable combinations and examine whether Bertin's selectivity and associativity concepts also represent continua.

### *Indispensable Variables*

There seem to be differences in dominance among both visual variables and Gestalt grouping principles in various contexts. That position, in both space and time, has a dominant overall role in perceptual organization is the contention of Kubovy's (1981) concept of "indispensable" variables. Both Pinker (1990) in relation to graph understanding and Bertin (1967/1983) in relation to map understanding have chosen to ignore time, the second of Kubovy's indispensable variables. Considering the current attention to map animation and dynamic visualization, however, we can no longer afford to do so.

In a map context, Slocum's (1983) analysis of proximity and similarity as factors in groups seen on graduated circle maps supports the contention that spatial location (in the form of proximity) is a more dominant variable than similarity (of size). In Slocum's study, in fact, circles of different size were more likely to be seen in the same group than those of the same size. His subjects attended to relative location of circles and ignored similarity of size. In the context of multivariate dot maps, however, Rogers and Groop (1981) found that proximity did not overpower color hue. Their subjects were able to identify univariate regions as effectively on trivariate dot maps using different color dots for each of the three variables as they could on individual dot maps. While this result does not necessarily counter the claim that location is an indispensable variable, it does indicate that grouping by a conjunction of color hue plus proximity works as well as grouping by proximity alone.

Humans appear able to segregate the visual scene in terms of both position in X–Y (or the plane of the retinal image) and position in Z (or depth). Research on visual search for objects having conjunctions of two or more variables, for example, has demonstrated that perception can segregate a scene on the basis of depth planes and position in these planes. Nakayama and Silverman (1986) presented subjects with displays in which stereo disparity was used to produce a near and a far visual plane containing colored items. In their experiment, all nontarget items in each plane were a single color hue (e.g., near = red and far = blue). Targets were the opposite color of the depth plane in which they appeared. Subjects were told to locate the colored target and their response times were measured for displays having various densities of nontarget items. The display density did not affect search times, indicating that search was accomplished in parallel (i.e., all potential targets were attended to at once). Since the depth plane that did not contain a target had items of the same color as the target, this result means that subjects were able to direct their attention to one position in Z and to ignore the potentially distracting objects at another position in Z.

Following from these results for position in 3-D space, we might predict that position in space–time will be easily distinguishable (and more noticeable) than position in static space or aspatial time. This makes sense on evolutionary grounds. Our ability to attend to moving objects can be thought of as an ability to focus attention on position in space–time. If the position of an object changes over time, it is very difficult to avoid attending to it. This "fact" is the basis for the Gestalt principle of common fate, which Wertheimer (1923; translated in Ellis, 1955) argued was often dominant over grouping by proximity. Humphreys and

Bruce (1989) cite a number of related studies in which various conjunctions of locational with nonlocational visual variables were tested. It is clear from these studies that visual scenes can be segregated by disparity in both depth and motion (across space over time) and that these aspects of location are dominant over nonlocational variables such as color, form, orientation, or size. Both motion and disparity in depth also seem to dominate position in the plane as a factor in forming perceptual groups. One counterpoint to the argument that disparity in depth is more noticeable than differences in color, texture, and the like, is that natural camouflage of animals and artificial camouflage of military equipment both seem to be effective in concealing, in spite of the presence of depth due to binocular parallax—until movement occurs.

## Where We Attend

In relation to visual attention, we began by considering what humans attend to when we look at a particular map (or spatial display) location. In this section, we move on to consider various factors that determine where we look when viewing a map. Two aspects of this question are considered, location within the visual scene and the scale of attention.

### *Location*

Attention to items in the visual scene has been likened to a *spotlight* that highlights a small area making it more visible than its surroundings (Posner, 1980) (Figure 3.41). This spotlight can be directed away from our fixations to objects or events in peripheral vision (without changing the direction of fixation). It is therefore somewhat independent of eye movements. In fact, it is probable that the ability to change the spotlight (or location) of attention without eye movement may be how the visual system determines where the next eye movement should fixate (Humphreys and Bruce, 1989).

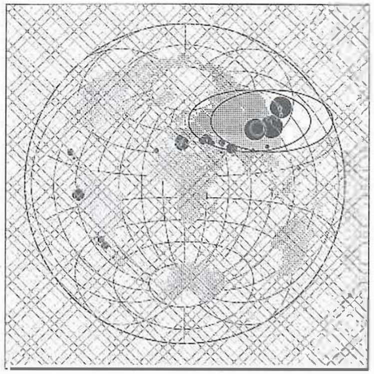

**FIGURE 3.41.** The attention *spotlight* as a viewer scans a map. Each ellipse represents the region of emphasis by each eye.

The view of attention as a spotlight with a focus, a margin, and a fringe can be traced to William James in 1890 (1890/1960), but has regained popularity due to a variety of recent response time studies that show response time decreases for appearance of stimuli at locations anticipated due to a cue and increases for appearance at locations away from location cues (Humphreys and Bruce, 1989). Tsal and Lavie (1988) have also demonstrated that when subjects were prompted to locate targets (letters) of a given color or a given shape, other nearby letters were more likely to be recalled than letters of the same color or shape that were not adjacent to the target. They contend that their findings "strengthen and extend the notion that attention operates as a spotlight" (p. 19).

We seem able to narrow the focus of the spotlight more in the foveal area of vision than in the periphery (Downing and Pinker, 1985). Evidence also suggests that the focus of attention may begin with a wide aperture (but low resolution) and gradually change to a more focused, higher resolution (Eriksen and Murphy, 1987). As a consequence, the analogy of a *zoom lens* has been offered as an improvement on the original spotlight analogy. This multiscale feature of attention with its apparent tendency to begin with a broad view corresponds to evidence for a dominance of global versus local processing of visual scenes (Navon, 1977) and to Marr's contention that 3-D model representations are hierarchical, making recognition of membership in a category possible before recognition of individuals (see Chapter 2 for details).

Cartographically, a key aspect of the way attention works is that initial views, if they take in large segments of a map, will be able to process only gross features. This processing will then guide the narrowing of attention to particular features and objects in order to examine details. Particularly in a visualization context, therefore, graphic design impacts upon the initial wide-scope global view of the map and may dictate what specific details are seen. Also, in the case of reference and travel maps, the ease with which point features, labels, and other small map items can be found by scanning across the map is likely to be controlled to a large extent not by the discriminability of separate features of symbols or text, but by the higher level appearance of symbols and words as a whole and by the overall map structure that may influence attention and thereby guide where attention will be directed (Phillips and Noyes, 1977—see discussion in the "Scanning the Visual Scene" section below).

There is evidence that attention can be directed to objects as well as locations. Duncan (1984), for example, demonstrated that subjects could more easily attend to two attributes of one object (rectangle length plus position of a gap in the rectangle or line type and its orientation) than to elements of two different objects (length of a rectangle plus line orientation). In his experiment, the objects were superimposed, and therefore location was the same. That we can attend to features of objects when location is restricted to foveal vision, however, does not discount the role of location in attention. As mentioned in Chapter 2, when more than one object at more than one location is presented to us, we are more able to attend to a particular location (and to all objects at that location) than to a particular category of objects regardless of position. To attend to an object, we must first attend to its position. It is only then that the features of the object begin to be clear enough to guide our attention to them.

### *Scale*

In the previous section, it was suggested that visual attention may begin with a broad extent but relatively coarse resolution, and then, based on cues obtained from this initial perspective, be redirected to another location or focus in on a particular area or object. Humphreys and Bruce (1989) contend that overall spatial structure is probably available more quickly than is the structure of local details. Humphreys and Quinlan (1987) suggest that both pattern and object recognition might rely on descriptions available from relatively low spatial frequencies—the global features—because patterns at this frequency are more stable over time.

Neurophysiological evidence complements the view of multiple spatial scales of visual processing. Wilson et al. (1990, p. 240) cite research with cats and macaque monkeys indicating that "at each stage of the visual pathway cells with receptive fields in the same part of the visual field can respond to different ranges of spatial frequency." In addition, they contend that "spatial frequency selectivity becomes progressively narrower moving up the system from retinal ganglion cells to LGN cells to simple cortical cells."

The idea of multiple scales of attention is closely associated with research concerning global–local precedence—whether global holistic properties or local components or parts are perceived more readily (Watt, 1988). The divided attention studies of Pomerantz and his colleagues provide one piece of evidence for global structures taking precedence over individual features. As noted above, their results demonstrate that in categorization tasks, certain arrangements of parts are processed more quickly as a unit (a whole) than are either of the individual parts (Pomerantz and Schwaitzberg, 1975). These results seem to support global precedence for "good" Gestalt groups, and local precedence for "poor" groups.

The experiments, however, focus only on the issue of global versus local processing of small perceptual units that are easily attended to because they appear in foveal vision, exactly where they are anticipated. Their concern, then, is not spatial but object-based and their research has little to say about the relative spatial scope of attention and how it is controlled.

In examining a map or other visual scene, one role of visual attention is to determine where to look. Because attention can be directed to various locations at various scales, the issue of whether perception usually begins with a spatially global or a spatially local perspective becomes important. Most studies of global–local precedence seem to implicitly accept the roving zoom-lens analogy for visual attention and have focused on the question of the scale of feature that is most easily attended to initially.

In a now classic study that has stimulated much of the subsequent research, Navon (1977, p. 354) investigated the postulate that "perceptual processes are temporarily organized so that they precede from global structuring towards more and more fine grained analysis [local structuring]." His experimental stimuli (compound letters), were selected so that global and local components could be manipulated independently (Figure 3.42). The stimuli were composed of small letters organized in arrangements to create large letters with the small and composite letters being either the same or different. Subjects were asked to identify either the local stimuli (small letters) or the global stimuli (large letters), and the speed with which they could do so was measured. What Navon found was that identification of global features was faster than identification of local features, and that conflicts between local and global letters interfered with identification of local letters, but not with identification of global letters. Navon's interpretation of his findings was that global processes must necessarily be prior to local ones. What is not clear from this research is whether identification of global stimuli requires a prior grouping (of as yet unidentified local stimuli).

Subsequent research by Paquet and Merikle (1988) has considered situations in which the visual scene is composed of more than one element set. They again used compound letters as stimuli, but presented sub-

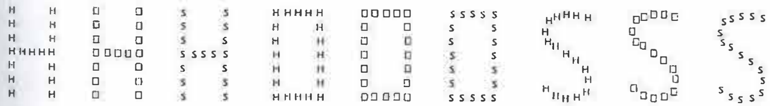

**FIGURE 3.42.** An example of the kind of compound letters used by Navon and other psychologists studying global–local precedence. *Derived from Navon (1977, Fig. 5, p. 365).*

jects with pairs of them rather than a single set (Figure 3.43). Subjects were asked to attend to one of the pair (identified by a surrounding circle or square) and, as in Navon's study, to identify either the local or the global letter. Their results confirmed that the global letters were identified faster and that the global aspect of the attended form was harder to ignore than the local aspect. Beyond this confirmation, however, they found that both global and local aspects of the unattended stimulus could influence identification speed, with local features having an influence if a local identification was requested and global features having an influence if a global identification was requested. Further, Paquet and Merikle (1988, p. 98) found that "it was impossible for observers to ignore the category of the global aspect of the nonattended object." This latter finding seems to add even more support to the idea that space is an indispensable variable—because we anticipate features near one another to be related.

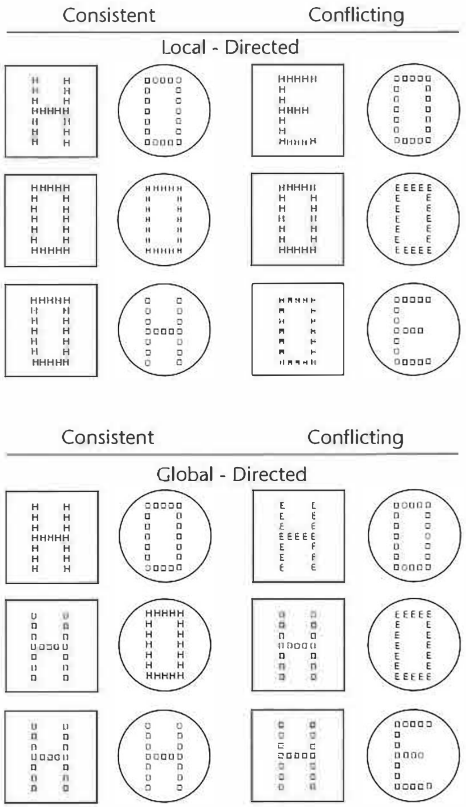

**FIGURE 3.43.** Sample compound letter stimuli. *Derived from Paquet and Merikle (1988, Fig. 1, p. 91).*

If the zoom-lens analogy for visual attention and the idea that attention varies in acuity from central focal point to its fringes are correct, we can anticipate extensions and modifications to Navon's original ideas about global–local precedence. First, since it is clear that people can attend to local details when directed to, we might anticipate that global precedence will be strongest when we are not already cued to expect some local feature. Second, global precedence can be expected to be stronger on the periphery of attention where the resolution of attention (and of vision) is not sufficient to resolve local details. Third, we might expect to find limits on scale of global elements that will be attended to quickly—if they are too large, the elements will be beyond the bounds of our attentional zoom lens, and if they are too small, they will be local details.[7]

All of the above possibilities have been supported to some extent by empirical research. In an experiment using compound letter stimuli similar to Navon's, Pomerantz (1983) dealt with the first issue. Half of his subjects had to respond to either the global or the local letter when it appeared on the screen at random locations. For the remaining subjects, presentation was always at the center of the screen. For both certain and uncertain presentation locations, global letters were easier to identify than local ones, but the difference was greater for uncertain than for certain locations. In a related experiment, Lamb and Robertson (1988) had subjects fixate on the center of the screen before presentation of a compound letter. Presentation could be central or to either side of center. They found that the global-identification speed advantage was greatest for peripheral presentations.

That size of perceptual units has an impact on global–local precedence is supported in a variety of studies. In research with compound stimuli composed of geometric shapes rather than letters, Kimchi (1988) found that the number of local elements making up the global shape interacted with the strength of global precedence (Figure 3.44). Specifically, when the number of elements was small, thereby making the global figure small, global processing was faster whether or not local and global shapes agreed. With larger global stimuli, composed of more local elements, global processing was faster in situations where there was conflict, but not in situations where the shapes agreed. More direct evidence that the size of a global figure must be within some limit in order to receive attention precedence can be found in research by Kinchla and Wolfe (1979) with compound letter stimuli and research by Antes and Mann (1984) with pictorial stimuli. For the compound letters, Kinchla and Wolfe found that compound figures larger than 8° of visual angle (roughly the size of the United States on a page-sized map of North America at normal reading distance) resulted in a reversal of attention to local prece-

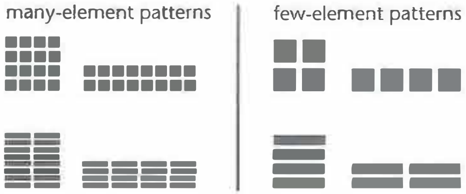

**FIGURE 3.44.** Nonalphanumeric compound symbols used in testing for global versus local precedence. *Derived from Kimchi (1988, Fig. 1, p. 191).*

dence over global. For pictures, Antes and Mann found that global precedence existed for pictures subtending 4°, local precedence occurred for pictures subtending 16°, and neither global nor local precedence was found for scenes subtending 8°.

An interesting feature of the Antes and Mann study is that their picture stimuli contain global–local dependencies (at a semantic level) not found in the compound letter stimuli. Identification of global scenes in pictures can depend (in part) on identification of local details (e.g., one of their pictures is a farm scene that would be quite difficult to distinguish from many other landscapes unless a local detail, the barn, is recognized). In spite of this interdependency, results concerning the effect of scale on global–local precedence is consistent with that for the compound letter stimuli, for which local and global levels are independent. This comparability of results suggests that global–local precedence effects processing of representations at multiple processing levels (e.g., the low-level primal sketch and subsequent 3-D model representation where recognition can occur).

Most global–local research in psychology, like virtually all cartographic research, has been directed to static displays. There is, however, evidence of global–local processing in the temporal as well as the spatial dimension. The research on temporal context might be considered an extension of the Gestalt grouping principle of "objective set" (that humans tend toward stability over time in perceptual grouping). Objects grouped at one time remain grouped over time, even when changes in proximity would result in no groups or different groups in a static scene. Palmer (1975) examined a similar idea in relation to the effect of temporal context on identification of objects. He hypothesized that global scenes in a temporal sequence would influence the identification of local details presented in subsequent scenes. Palmer found that "appropriate" prior scenes facilitated object recognition and "inappropriate" prior scenes (e.g., a kitchen prior to presentation of a mailbox) impeded recognition. When the subsequent object was similar in appearance to one more logically part of the scene (e.g., a loaf of bread) misidentification was likely.

At this point, we can only speculate upon the implications of global–local research for map understanding. There has been no cartographic research to date that has extended directly from these studies. As Mistrick (1990) points out, however, results of the global–local processing research support the concept that extraction of meaning from visual scenes uses hierarchical structuring of information at multiple spatial scales. This view corresponds quite well to Marr and Nishihara's (1978) ideas concerning how primal sketches are derived from a retinal array and to research directed at higher levels of processing that indicates hierarchical structures for memory encoding of spatial knowledge. As discussed in Chapter 8, issues of global–local precedence may have particular relevance to exploratory visualization with maps—a situation in which an analyst is not entirely certain what patterns to expect and a situation for which dynamic manipulation of display scale will be a significant part of the analysis.

## Scanning the Visual Scene

Both what is attended to and the scale of that attention interact with the process of visually exploring a map or other graphic display. It seems clear that both global views and peripheral attention act to steer eye movements toward important information in a visual scene and away from unimportant information. As early as the 1970s, cartographers investigated the use of eye movement recordings as a tool for understanding the visual–cognitive process of map reading and the impact of both changes in map design and training on that process (see Steinke, 1987, for a comprehensive review). In addition, cartographers have studied a variety of visual search problems on maps in an effort to determine how to facilitate search for specific kinds of features (e.g., placenames, point symbols, etc.). Much of the early work in both visual search and eye movement analysis by cartographers suffered from a lack of theoretical perspective. As Steinke (1987, p. 57) noted in relation to eye movement research, early work seemed driven by a "let's see what happens when we put a map in front of somebody and photograph their eyes" approach. Only recently has cartographic research dealing with visual scanning of maps begun to build on a firm theoretical base grounded in perceptual–cognitive theory. Looking back from an information-processing perspective, however, we can identify some links between early cartographic eye movement research and the perspectives, presented above, on vision as a modular system for processing increasingly interpreted representations.

Both Marr's model of vision and Gestalt principles suggest that edges are important elements of visual scenes that are processed early in vision.

Steinke (1987) cites eye movement research by both Thomas and Lansdown (1963), with medical images, and Gratzer and McDowell (1971), with landscape photos, in which foveal attention to edges in the scene was demonstrated to be a common tendency, in spite of highly individualistic scan paths. In the Gratzer and McDowell study, edges defined by the skyline, ridgelines, shorelines, and vegetation boundaries all received particular attention. Not all edges, however, seem to be equal in attracting foveal attention. Mackworth and Morandi (1967), for example, provide evidence that simple contours are noticed using peripheral vision and largely ignored by foveal fixations. It is unpredictable edges, or edges with unusual details, that seem to attract foveal attention.

Although cartographers employing eye movement techniques are clearly aware of psychological research showing attention to edges, no cartographic research based on this knowledge seems to have been done. On maps, eye movement techniques might be used to determine the relative "goodness" of the contour established by different methods of creating contrast between map regions. Alternatively, analysis of eye movements might be used to assess techniques for enhancing the identification of map regions. Although Jenks (1973) found little similarity in map viewers' scan paths when viewing a dot map, he did identify commonalities in relative attention paid to different parts of the map. He did not, however, examine fixation times in relation to regional edges delineated by his subjects. The one example that he does provide of a subject's fixation times for map cells suggests that more attention was given to the edges of dot clusters than to the core of those clusters (Figure 3.45). Dobson (1979a) also found correspondence among subjects about relative attention to different parts of a test map. In examining his data, he found a high correlation between the "informativeness" of map sections and visual attention to those locations. Both Jenks's and Dobson's results suggest that eye movement analysis might be applicable to assessment of the distinctiveness of regions on thematic maps. Measuring attention to transition zones around a region or between regions could be used to assess the strength of regional edges.

One common feature of early cartographic research using eye movement analysis was that subject attention to maps was observed in the absence of defined tasks. Castner and Eastman (1985) point out that we should expect quite different perceptual and cognitive processes to be at work in this situation, which they called "spontaneous looking," and in "task-specific viewing." With spontaneous looking, location of attention will be influenced primarily by the properties of individual map symbols, Gestalt properties of symbol groups, and reader attitudes or expectations about the stimulus and the experimental situation. For this kind of viewing (perhaps typical of early exploratory visualization), then, map design changes are likely to be particularly important. On the other hand, Castner and Eastman contend that for task-specific viewing, cognitive processes will exert a much stronger control over eye movements and attention. In these cases, eye movement analysis might be used to distinguish different problem-solving strategies (or application of different schemata). Recent evidence by Morita (1991), however, demonstrates that while task-specific viewing can lead to greater similarity in eye movement parameters than spontaneous looking, design alterations can still play a major role. He found very distinct differences in patterns of visual search on schematic maps in which numerical information was symbolized by seven different graphic variables (see Figure 2.14).

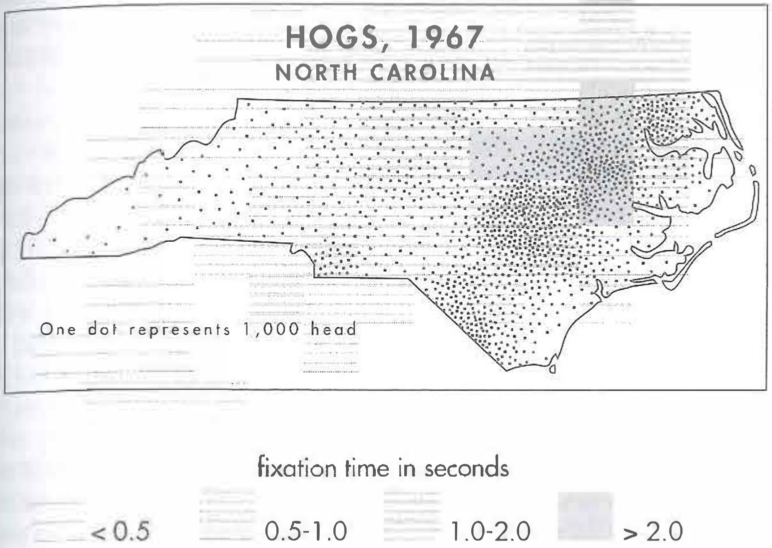

**FIGURE 3.45.** A map of attention to locations on a dot map. *Reproduced from Jenks (1973, Fig. 9, p. 34). Reprinted by permission of Universitätsverlag Ulm GmbH.*

Most cartographic research using eye movement analysis has focused on questions of where foveal vision is directed. The technique can also provide information relevant to the issue of global–local precedence. A graduate research project by Guyot (1971; cited in Steinke, 1987) provides one example. Guyot was interested in how map patterns are compared. In an experiment in which subjects were required to select one of two maps that was most like a third, eye movement recordings were used to measure which of the two maps was attended to first, as well as which received the most and the longest fixations. The finding that the first map looked at was generally selected as the most similar to the referent map suggests that peripheral vision played a key role in pattern comparison for these maps. This in turn suggests that global processing of map patterns is quite sophisticated and that it directs eye movements, at least in a task-specific viewing situation.

Recent uses of eye movement analysis in cartography have focused on task-specific viewing. This focus is driven, in part, by the lack of consistent results of past "spontaneous looking" experiments and by the desire to achieve particular application goals. The map-use task to which eye movement evidence seems most applicable is visual search. Phillips and his colleagues in the Psychology Department of University College in London (with input from cartographers Bickmore and DeLucia) were among the first to use eye movement techniques to examine strategies of visual search for information on maps. They focused on the influence of map design on search and gave particular attention to the problem of searching for names on maps (see Phillips and Noyes, 1977). They proposed that reducing either the number of fixations or the duration of individual fixations would speed searches for place labels, and that reducing the number of fixations should be more effective. Three map-design procedures were suggested to reduce the number of fixations: generalization that reduced the total number of names on the map, categorization and visual coding of names so that only a subset had to be attended to, and use of a map grid that could direct attention to a limited section of the map. Adjustments to type placement were used to control individual fixation times. Both procedures reduced search times, but the three techniques that reduced the number of names considered were (as predicted) more effective. Of these, using small map grids had the most dramatic effect.

Eye movement analysis has not proved to be as powerful a tool for cartographic research as originally anticipated by Jenks and others (Steinke, 1987). This limited success is due to both practical and conceptual issues. From a practical point of view, eye movement analysis has been difficult and expensive. Conceptually, it suffers from the problem that there is no simple way to determine whether a fixated location is also attended to. Because of these problems, the question of visual scanning (particularly visual search) has been investigated in a number of other ways. The two most commonly used measures in visual search experiments (other than eye movement analysis) are the accuracy and the speed with which targets are identified. Accuracy is assessed by the number of correct identifications compared to the number not found and/or misidentified.

In the only cartographic study to consider the role of figure–ground in visual search, Lloyd (1988) compared perceptual and imagery processes involved in determining the presence or absence of pictorial point symbols on a simple map like display (Figure 3.46). The display had a central green area surrounded by a blue area and symbols in yellow were dispersed relatively evenly across the map's surface. The central green area (presumably due to centrality, surroundedness, and slightly smaller size) was expected to be a figure on the blue background. Subjects in the perception condition had the task of determining whether or not two different symbols both appeared on a display. Symbols were either both absent, both on the figure, both on the ground, or one on the figure and one on the ground. Lloyd predicted that when both symbols were on the figure they would be found more quickly because the figure was expected to draw the subjects' initial attention. When both symbols were on the ground he expected opposite results, and with one symbol on the ground and one on the figure he expected intermediate results. Although mean response times exhibited this ordering, differences were not significant. The explanation provided, and one that could be predicted from Gestalt principles, is that the yellow point symbols (due to small relative size and greater contrast with either area than the areas had with each other) were the dominant figures seen and that for most subjects both the green and blue areas became ground.

Treisman et al. (1990) (building on work by Cavenagh, 1987, 1988) proposed a model in which five independent visual pathways—luminance, motion, binocular disparity, color, and texture (Figure 3.47)—process different attributes of the visual scene. This model is supported in

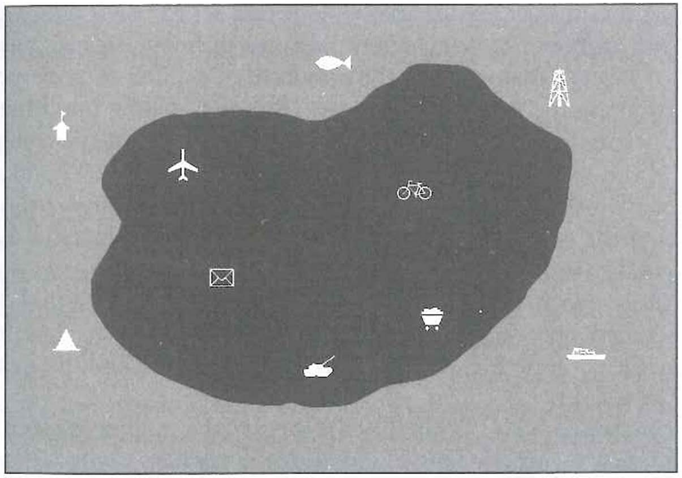

**FIGURE 3.46.** Simulation of the test map from Lloyd's experiment. *After Lloyd (1988, Fig. 4, p. 366). Adapted by permission of the American Congress on Surveying and Mapping.*

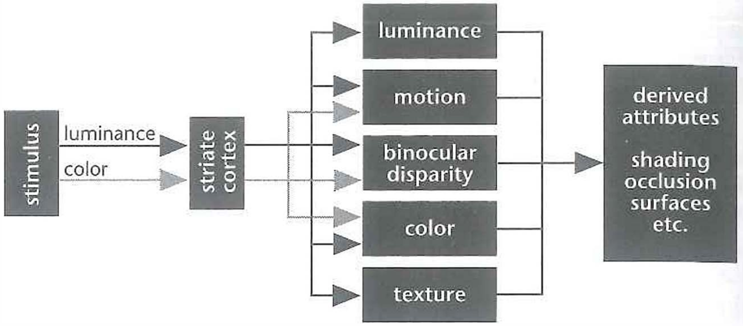

**FIGURE 3.47.** Treisman's five-pathway model for processing of the visual scene. *Derived from Treisman et al. (1990, Fig. 16, p. 294).*

part by neurophysiological findings that provide evidence for separate color–form and motion pathways and neuropsychological evidence from brain-damaged patients that shows luminance operating separately from either color or motion. Experiments using visual search tasks, examination of after-effects, and identification of shape from contours have confirmed the separate pathways for color, luminance, and motion, and have provided evidence that early vision must also process texture and binocular disparity separately. Each of these visual pathways seems able to code size and orientation information which, in turn, may function as shape primitives.

The model led to Treisman's (1988) feature integration theory (FIT) that posits a series of "feature maps" with one set for each pathway plus one for size and one for orientation. Each set of feature maps consists of individual layers (analogous to the structure of a GIS) to code possible variations along the specific dimension (e.g., red, yellow, and blue color maps or orientation maps for various angles). If the Treisman model is correct, it can explain why some visual search tasks can be conducted in parallel while others require serial processes. If a target differs from other elements of the scene along one dimension or feature, a parallel holistic process can be used. When the search is for a conjunction target that can share features with nontargets, however, location of potential targets must be relied upon to link separate feature maps before determining whether the conjunction occurs. This location-based linking can only proceed in serial.

Recently, Cave and Wolfe (1990) suggested a modification of Treisman's FIT that seems relevant to visual search for map symbols. As originally conceived, FIT predicts that if a target (such as a map symbol) differs from all distractors by a single feature, parallel processes should be able to detect it. If, on the other hand, the target is defined by a conjunction of features, FIT predicts that parallel processes play no role and that the serial processing of each individual element proceeds until the target is found. Cave and Wolfe's "guided search" theory posits that both parallel and serial processing are used in conjunction for all visual search tasks. They suggest that a fast parallel (i.e., holistic) processing stage identifies the basic visual features that are present (just as Guyot, 1971, found that global processing of maps in peripheral vision could determine map similarity). This first stage is followed by a slower serial stage that combines the features identified to produce object representations. With conjunction symbols on maps, then, we would predict that a parallel search procedure would act to limit the serial search stage to elements sharing one of the conjunction variables. Guided search, then, is thought to use parallel processing to steer search by identifying the likelihood that a feature is a target and by allowing vision to bypass those features with low target likelihood.

Visual search remains incompletely understood. Cheal and Lyon (1992), for example, demonstrate that neither FIT, nor guided search, nor similarity theory (Duncan and Humphreys, 1987) can handle all aspects of how parallel and serial processes might interact in the search for even relatively simple shapes. Another alternative for combination of parallel and serial search was actually suggested in earlier work by Treisman (1985). If the visual scene can be subdivided into regions (as it often is on maps) within which a conjunction target differs from nontargets along only one feature, search can be parallel within each region, while remaining serial from region to region. The advantage of regionalizing search is supported for map reading by Phillips's placename search research described above (Phillips and Noyes, 1977; Phillips et al., 1978).

## Figure–Ground

The combined information concerning perceptual grouping, visual attention, and global–local processing of visual scenes provides a firm base from which to understand the concept of *figure–ground*. Segregation of figure from ground requires that perception organize the visual input sufficiently for elements of that input to group and attract attention to themselves. In the environment, as well as on maps, distinctions between significant and insignificant elements of a scene need to be made at a coarse holistic level so that they can guide further attention to specific details. figures that attract our attention are distinctive from the background and often appear to be in front of that ground. Hochberg (1980, p. 90) seems to sum up the general concept of figure–ground, and illustrate the fuzziness of figure–ground as a concept, when he states that "figure is thinglike and shaped, ground is more like empty space, amorphous and unshaped."

Figures are defined by their contour or the boundary between object and nonobject. The contour operates on only one field or the other, but not both (Arnheim, 1974). When it is unclear which region of a scene the contour belongs to, an ambiguous figure results in which first one, then the other, part of the field becomes figure (Figure 3.48). Such ambiguity is what cartographers strive to avoid.

Most cartographers who have considered the issue of figure–ground segregation on maps have turned to Gestalt psychology for guidance. This is a good place to start because Gestalt psychology has made the greatest contribution to our current understanding of figure–ground. Gestalt psychology alone, however, does not answer all of the cartographically relevant questions. The discussion that follows will begin with a brief overview of some of the more important Gestalt ideas relating to figure–ground and will then explore more recent psychological and cartographic research that has attempted to extend from this base.

Gestalt psychologists' initial approach to figure–ground began with principles of perceptual grouping because, to see a figure, a perceptual unit must exist and grouping produces perceptual units. All grouping factors have a potential role in defining regions of a visual scene, and the scene must be differentiated in order for figures to appear. Extending from the fundamental grouping principles, a set of related principles has been devised to deal with the visual strength of perceptual groups as figures segregated from a background. Some were suggested directly by Gestalt psychologists and others were derived more recently by researchers following similar logic. The factors below seem to be the most relevant to establishing symbols and regions as figures on maps.

1. *Heterogeneity*: A visual field must be differentiated to form groups before one part of the field can stand out as figure. While not one of Wertheimer's (1923; translated in Ellis, 1955) grouping principles, this idea was offered (with no label assigned) at the end of his paper. The main guiding principles offered for establishment of figure were that an enclosed shape for which there was a color difference between the shape and the background will stand out as a distinct figure. This basic idea was elaborated by a number of other Gestalt psychologists. Among them was Koffka (1935) who introduced the concept of articulation as a figure–ground principle (see below for more details).

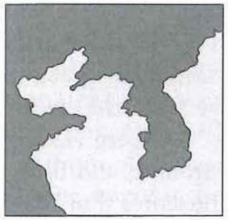

**FIGURE 3.48.** An ambiguous map in which subjects were asked to determine which area was the figure. Half of the subjects picked light and half picked dark. *Derived from Mistrick (1990, Fig. 3.3, p. 60).*

2. *Contour*: Objects are more easily seen as figure the more definite the edge between object and nonobject. The establishment of contour follows directly from establishment of heterogeneity. A noticeable difference between areas creates an edge, boundary, or contour between them. If differences are due to relatively coarse features, the edge will be rather fuzzy and the contour (and the experience of figure) may be weak (Figure 3.49). If the differences consist of fine-textured fills or solid colors, however, (or a distinct line separates the regions) the contour will be stronger, as will the experience of Figure (Figure 3.50).

3. *Surroundedness*: Completely surrounded objects tend to be seen as a unit and thus as figure (Bruce and Green, 1990). That is why, even with weak contour, the white area of Figure 3.49 is seen as figure. This principle is probably the single most useful in creating figure–ground distinctions on maps. As a number of conflicting experiments recounted below make clear, it is hard to avoid ambiguity about figure versus ground in displays that do not have a surrounded figure. Centrally located surrounded shapes will enhance figure formation.

4. *Orientation*: Objects with a horizontal or vertical orientation are seen as figure (Bruce and Green, 1990). A rotation of areas on a hypothetical map illustrates this point (Figure 3.51). It is likely that this tendency has to do with relationships between visual displays and real-world visual scenes in which most figures are upright (as are humans, trees, etc.) or aligned with the horizon.

5. *Relative size*: The smaller of two areas is more likely to be seen as figure. This factor is essentially a corollary to the factor of surroundedness cited above. It is particularly relevant when the larger area completely surrounds the smaller. Assuming the object has sufficient size to be easily detected, the smaller the object relative to its surround, the more it is seen as a figure (Figure 3.52).

**FIGURE 3.49.** A central figure that fades into the background due to a weak contour.

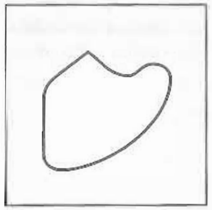

**FIGURE 3.50.** The previous figure–ground difference enhanced by a stronger contour.

6. *Convexity*: Convexity will be seen as figure. A schematic map of England, for example, has greater convexity than a map using a detailed but smoothed boundary. As a result, it should be more likely to stand out as figure (Figure 3.53).

### *Heterogeneity*

Of the factors thought to lead to "good" figures, it is the first, heterogeneity, that has been given the greatest attention. Issues of contour, surroundedness, orientation, symmetry, and convexity have typically been considered as complements to a primary focus on heterogeneity. Both psychologists and cartographers have investigated how segregation of figure and ground is influenced by various methods of creating heterogeneity between areas. Psychologists have primarily been interested in what figure–ground reactions to differences in texture, value, temporal frequency, and so on, tell us about the visual and cognitive processes underlying perceptual organization of visual scenes. Cartographers, on the other hand, have been most interested in developing guidelines for use of area fills on maps that will lead to consistent identification of specified portions of the map as figure. Psychological research has used a rather limited range of possible area fills, limiting the generality of their findings, and cartographers, while they have tested more kinds of area fills, have not linked their research to psychological theory beyond that of the Gestalt principles developed in the 1920s and 1930s. The discussion below provides an overview of some of the work from both disciplines (with an emphasis on the role of brightness differences) and uses an experiment by one of my graduate students as an example of how we might integrate these two research streams more effectively.

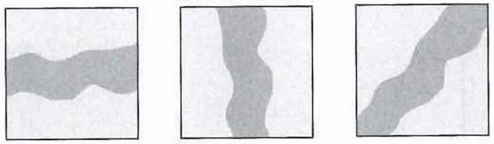

**FIGURE 3.51.** When value contrast is relatively equal, map regions that are horizontal or vertical are more easily seen as figure than regions that are diagonal to the map border.

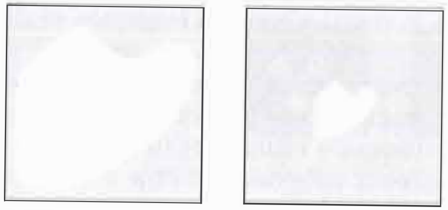

**FIGURE 3.52.** Small, surrounded map areas are more easily seen as figure.

Both cartographers and psychologists have given considerable attention to the role of brightness (i.e., color value) differences in figure–ground. On maps, it is a particularly important tool because all other means of creating heterogeneity between areas (with the exception of color hue) have a tendency to interfere with other map information (e.g., texture differences require one area to have a coarse enough texture to be noticeable as texture—and coarse-textured area fills make text difficult to read). In psychology, attention to brightness as a figure–ground factor began in the 1920s, with Wever (1927) among the first to address the issue.

In relation to brightness, Wever (1927, p. 222) contended that "a minimum brightness difference is necessary for the experience of figure and ground. As brightness-difference increases, the 'goodness' of the experience increases, though at a constantly diminishing rate." This contention was based on research in which subjects viewed irregular forms presented tachistoscopically. In his experiment subjects viewed 1,060 different black forms on white backgrounds with illumination varied, result-

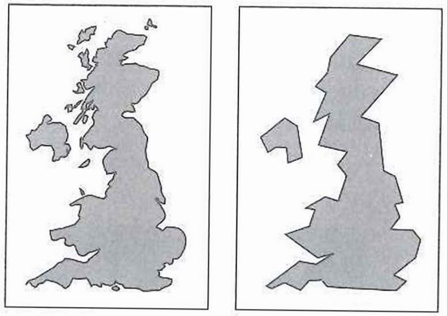

**FIGURE 3.53.** A detailed depiction of the coastline of England, Scotland, and Wales compared to a schematic depiction. The latter exhibits greater convexity.

ing in differences in contrast between brightness of the white and black areas.

Much of the psychological research subsequent to Wever has used variations on a simple pie-wedge stimulus. This stimulus allows researchers to experimentally manipulate the actual value of parts of each stimulus, the number of components to the stimulus, their relative size, and the background upon which the stimulus appears. One of the first uses of this stimulus was made by Goldhamer (1934) who examined the influence of relative size on the appearance of white versus black wedges on a gray background (Figure 3.54). He showed that for equal-sized regions, black tended to be seen as figure, but when size varied, it was the smaller shapes that were regarded as figure, regardless of brightness. Oyama (1960) used similar stimuli but came up with somewhat conflicting conclusions. The difference in his experiment was that the surrounds for the stimuli were either white or black rather than gray. Half of the wedges were gray (of varying shades for different stimuli) and half were the opposite of the surround (Figure 3.55). He found that the sectors opposite in brightness to the surround tended to be seen as figure, and that this tendency was strongest when the alternate wedges were closest in brightness to the surround. The situation in which surround and one set of wedges are similar resulted in the other set of wedges appearing as small shapes on a relatively homogeneous background. That the wedges of opposite brightness to the background appeared as figure in this case agrees with Goldhamer and with general Gestalt principles about small surrounded areas being seen as figure. When the alternating wedges were closer (in color value) to each other than either was to the background, however, the white wedges with black surround (and grey alternate wedges) were seen as figure more often than the black wedges with white surround (and gray alternate wedges). This disagrees with Goldhamer's finding that black shapes stand out as figure when size does not differ.

Overall, studies of brightness difference using the pie-wedge stimuli have produced equivocal results. There has been a consistent finding that the smaller of the areas are seen as figure regardless of brightness and that horizontal–vertical wedges are more likely to be seen as figure than diagonal wedges (Bruce and Green, 1990).

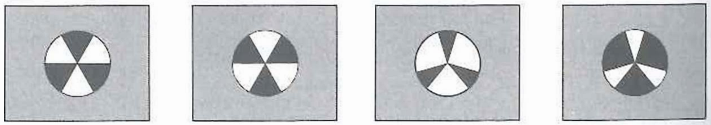

**FIGURE 3.54.** A sample of the kind of stimuli used by Goldhamer (1934).

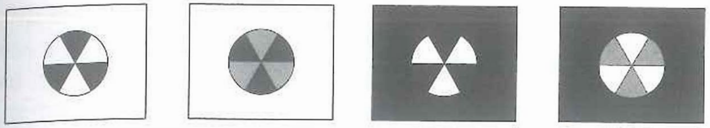

**FIGURE 3.55.** Examples of the variations on the standard pie-wedge stimuli used by Oyama (1960).

Most of the remaining psychological research on brightness as a figure–ground variable has made use of stimuli similar to the Rubin's vase–face ambiguous figure (Figure 3.56). This stimuli is somewhat more similar to the situation on a map in which land–water areas are adjacent and the cartographer wants one to be seen as figure. Harrower (1936) created an experimental setting similar to Goldhamer's, but with the vase–face figure on a surround. He used the following combinations: black surround, black face, white vase; black surround, white face, black vase; white surround, black face, white vase; white surround, black face, white vase. Whether there was a light or a dark vase or face did not seem to matter. The figure proved to be whichever differed from the surround, although there was a slight tendency toward face as figure. Harrower also examined a range of brightness differences assigned to surround, vase, and face. In this case, the face part of the stimulus was mounted on a track that allowed the halves to be pulled apart. The subject's task was to attend to the face as figure as long as possible. Results indicated that the face was held as figure longer as brightness difference was increased. Again, it did not matter whether the vase or the face was darker.

More recently, Lindauer and Lindauer (1970) used the Rubin's vase–face figure in a similar experiment involving brightness contrast. They compared a control (an unshaded outline drawing of the vase–face figure) with versions in which one area was white and the other was filled with a 20%, 40%, 60%, 80%, or 100% black pattern. In the unshaded control, the face was seen as figure, a finding that the authors attribute to

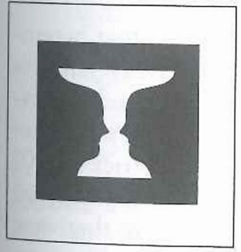

**FIGURE 3.56.** A typical vase–face ambiguous figure.

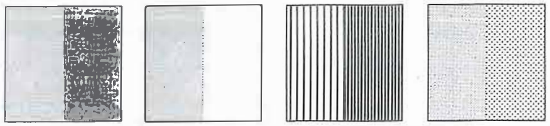

**FIGURE 3.57.** A typical stimulus from Dent's study. *Derived from Dent (1972, pp. 208–221).*

familiarity. For the test stimuli, responses to the shaded area as figure increased as the contrast between areas increased.

Building upon the rather mixed psychological findings, there have been a small number of published cartographic studies of the influence of brightness on figure–ground segregation. The first was by Dent (1972) as one component of his dissertation. He used bipartite squares, stimuli that were simpler than any of those used in previous psychological studies (Figure 3.57). As part of a larger study, his stimuli used various combinations of area fills on the two sides of the square. These were created with both dot and line patterns, in many cases texture was apparent because Dent was testing for it as well as for brightness. Dent's experiment, unlike the psychological research, was not based on response times (to identify figure or to hold particular areas as figure). Instead, his subjects were asked to examine each stimulus and to mark the side of the square that they saw as figure (or that visually stood out). In general, Dent found the coarser areas to be seen as figure (as the Gestalt concept of articulation would predict). When both areas were shaded with similar-sized dots, however, the finer textured, darker pattern was seen as figure—an indication that brightness might be more important than texture.

On a follow-up to the figure–ground test, Dent assessed subject preferences for maps in which a central focus area (e.g., North America) was shaded and the surround was white. For four different maps he found preference for the shaded maps over unshaded maps ranging from 71% of subjects to 96%. Wood (1976) also examined the "most desirable" brightness differences for maps. In this case all maps used some shading with relative brightness of figure, ground, and surround varied across the range of 16 test maps. Regardless of surround, maps with figures lighter than the ground were preferred.

The two preference studies by Dent and Wood suggest that heterogeneity of value for areas is preferred to homogeneity, but that preferences for lighter or darker areas as figures are inconsistent. Neither study directly addressed the issue of how value differences on maps influence the likelihood of map areas appearing as either figure or ground. Mistrick (1990) designed a study to do just that. In essence, she replicated a portion of Dent's (1970) initial figure–ground choice study using stimuli that were more maplike than his bipartite squares.

Mistrick's test maps, as we reported in our paper on brightness contrast as a figure–ground variable (MacEachren and Mistrick, 1992), depicted the land–water border along the coast of Korea (see Figure 3.48). The test maps were cropped so that the map border was square and the land and water areas of the map both occupied 50% of the area included. A pretest had shown that without labels the coastline was recognized by few students at Pennsylvania State University (the source of subjects for the experiment). The test stimulus, then, was a real map but due to the lack of familiarity with it, prior experience was not a factor in determining figure–ground segregation. In addition, the lack of familiarity allowed the map to be presented at four orientations with any change in response to it attributable to relative position of features and their concavities rather than to variations in recognizability of the map.

Mistrick's (1990) test map was created at four orientations with six combinations of area fill (white–dark gray; dark gray–white; white–light gray; light gray–white; light gray–dark gray; dark gray–light gray) applied to the land and water areas respectively (Figure 3.58). Each subject viewed only one map and indicated the region seen as figure. Two hundred forty subjects participated, half of whom saw unlabeled maps and half of whom saw maps with the labels "land" and "water" outside the map border next to these respective areas. The land–water labels had no effect on identification of one area as figure in relation to the other.

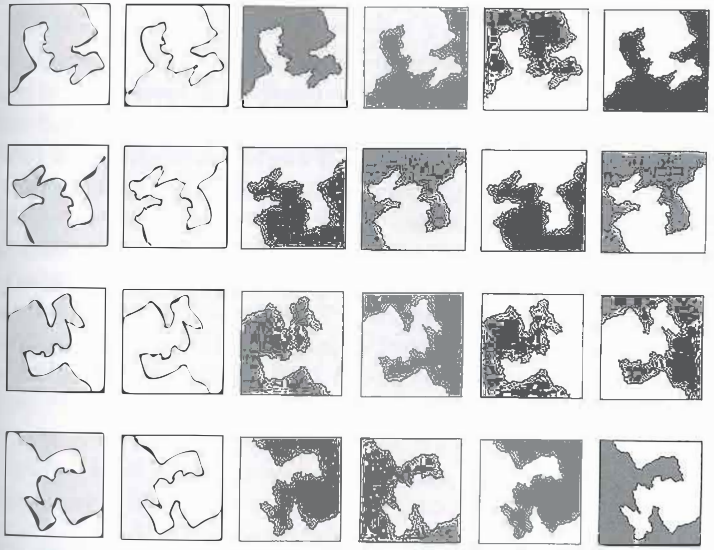

**FIGURE 3.58.** The complete set of maps used in Mistrick's study (at greatly reduced size). *Reproduced from MacEachren and Mistrick (1992, Fig. 6, p. 96). Reprinted by permission of* The Cartographic Journal.

Somewhat surprisingly, in relation to Dent's earlier findings with bipartite squares, the darkness of area fills had no effect on figure–ground segregation either. Exactly half of the subjects identified the relatively darker area as figure and half identified the relatively lighter area. This "negative" finding does, however, seem to agree with previous cartographic interpretations of Gestalt "rules" by Wood (1968), Spiess (1978), and McCleary (1981) that advocated relative value contrast (heterogeneity) between figure and ground but did not suggest that absolute value of area fills is relevant to what is seen as figure.

That contrast influences figure–ground segregation independently of the direction or sign or the difference is logical on "hardware" grounds. Shapley et al. (1990, p. 438) cite neurophysiological evidence that "fundamental neural mechanisms of pattern recognition are essentially nonlinear because these mechanisms ignore the sign of the contrast." They go on to contend that "form depends on the magnitude of contrast at a border while brightness depends on its sign." They demonstrate the implications of this fact with an illusory contour demonstration which reveals that contour-sensing processes do not rely upon the sign of contrast (Figure 3.59). In a similar vein, Barlow (1990) cites evidence that different sets of brain cells are specialized for light, for dark, and for value contrast. These contentions of independent brightness and form processing at a neurological level support the argument that Mistrick and I make (MacEachren and Mistrick, 1992) that contrast should be much more important in establishing figure on maps than specific brightness relationships.

In addition to research on heterogeneity due to brightness differences in relation to figure–ground, a number of psychological studies have considered heterogeneity created by contrast in spatial and temporal frequency of areal fills. This research has been closely linked to information-processing and computational-vision assumptions about how early vision represents visual scenes.

The influence of spatial frequency (or focus) on figure–ground segregation has been addressed in several psychological studies. Wong and Weisstein (1983) found sharp (high spatial frequency) targets (i.e., symbols) to be detected better against figural regions, while blurred (low to middle spatial frequency) targets were best detected in ground regions. Klymenko and Weisstein (1986), for regions that were otherwise ambiguous, examined the dominance of fills having differing spatial frequency (defined by horizontal sine wave patterns at frequencies of 0.5, 1, 2, 4, and 8 cycles/degree). They used the percentage of viewing time that each region was judged to be figure as a measure, and found high frequency patterns to be consistently seen as figures. In addition, the greater the frequency difference, the stronger the effect was. Significantly, in light of cartographic research based upon bipartite, or divided, squares, they found that a greater frequency difference was required for figure dominance in this situation than with more complex regions (i.e., a vase–face figure or an eight-element pie-wedge figure). This may be due to the importance of contour, or edges, in establishing figure, and because contour is more apparent when it has defined form. Kylmenko and Weisstein (1986, p. 324) link their results to Gestalt principles of articulation by suggesting that "greater detail is more or less correlated with the presence of higher spatial frequencies in the stimulus." They also cite a geometrical demonstration by Pentland (1985) that shows "gradient of blur" to be an ecologically valid indication of relative distance in depth. As a result, it can be postulated that blurred patterns will be seen as ground because they appear to be physically in the three-dimensional background. The idea that sharply defined patterns are seen as figure and fuzzy patterns as ground is relevant to the issue of visualizing uncertainty on maps through manipulation of *focus* (renamed "clarity" below) (MacEachren, 1992b).[8]

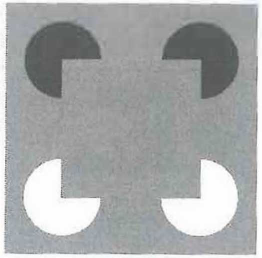

**FIGURE 3.59.** Two Kanizsa squares illustrating the dominance of contrast over brightness in establishing figure–ground.

### *Bottom-Up versus Top-Down Processing*

Of the Gestalt grouping factors empirically considered in relation to figure–ground, the role of "experience" (i.e., top-down processing), or lack of it, has generated the most lively commentary. Initial Gestalt views on figure–ground segregation treated vision as an immediate process acting on "wholes." As such, Gestalt psychologists discounted any role for top-down processing. Gregory (1990) is in general agreement with this perspective. He contends (based on his own research) that illusory contours occur at a processing stage between the retina and the parietal cortex, thus before a level at which object knowledge could be accessed. Kienker and Sejnowski (1986) cite a study by Ullman (1984) who reported that subjects presented with a closed outline and a small spot that could be either inside or outside that outline were able to determine that in–out location within a few hundreds of milliseconds. They agree with Ullman that subjects must have separated figure from ground, then made a decision about in or out. This contention derives from the Gestalt view that only figures have form, and thus an inside or an outside. Kienker and Sejnowski (1986, p. 198) go on to contend that the speed of processing Ullman found compared to time scales for neural processing and "suggests that figure–ground separation is computed in parallel over the visual field." Based on this evidence, figure–ground segregation is (or at least can be) a preattentive, bottom-up process.

There seems little doubt that figure *can* be found without input from higher level processes. Marr's information–processing approach suggests that the extraction of figure is one of the functions of early visual processing which, in the form of the 2½-D sketch, provides higher level processes with the input about object surfaces and orientations needed in order to recognize the objects. He demonstrated that, computationally, bounding edges and surfaces could be extracted from a scene with no input of prior knowledge (Marr, 1982).

The fact (if it is one) that figure–ground segregation can occur preattentively with no input from higher level processes, of course, does not prove that attention and top-down processing cannot sometimes play a role in figure–ground segregation. The ability of most people to consciously control figure–ground reversal when viewing Rubin's vase–face or other similar figures attests to this. In relation to ambiguous figures, Tsal and Kolbert (1985) have demonstrated empirically that figure–ground segregation is affected by attention.

Another related piece of evidence concerning the possible influence of top-down processes on figure–ground segregation comes from research by Peterson et al. (1991). She and her colleagues began by addressing the potential role of object recognition in figure–ground reversals of ambiguous figures. They had subjects view inverted and upright versions of Rubin's-like ambiguous figures. Their test figures were designed so that one orientation (upright) had highly denotative surrounds (viewers agreed on a specific shape represented) while an inverted version was not denotative (they were not recognized as a particular object) (Figure 3.60). Subjects continually reported which of the two regions appeared as figure during 30-second trials. They found that surrounds were more easily held as figure when they were upright (when the surround orientation was seen as denoting a particular identifiable object). Convergent evidence from four experiments led to the contention that "figure–ground reversal computations weigh inputs reflecting the goodness of fit between stimulus regions and orientation-specific structural memory representations" (Peterson et al., 1991, p. 1086). This finding agrees with Marr's model of shape recognition that posits that perceptual descriptions of shape structure (at the 2½-D sketch level) are matched to the best fitting structural

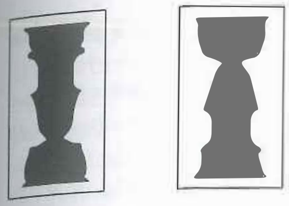

**FIGURE 3.60.** A sample pair of upright and inverted Rubin-like test stimuli. The surround on the left denotes a woman while that on the right has no clear denotation. *Reproduced from Peterson et al. (1991, Fig. 2a, p. 1077). Copyright 1991 by the American Psychological Association. Reprinted by permission of the author.*

memory representation (defined as a memory representation that specifies parts of a shape and their relative locations with respect to a canonical orientation for the object). Peterson et al. also interpret their findings to demonstrate that orientation-independent shape representations had no influence on figure–ground reversals.

Although Peterson et al. (1991) support Marr's views on matches with structural memory representations, there is disagreement about whether this matching can occur before figures are isolated. Peterson et al. (1991) contend that their results suggest a mechanism by which object recognition may facilitate initial figure segregation, as well as figure reversals. As they point out, this view seems to create a paradox: How can experience with shapes influence figure–ground organization given that no shape description should exist until figure–ground organization is determined? Their hypothesis is that contours may be evaluated from both sides simultaneously before figure is determined. A parallel set of experiments by Peterson and Gibson (1991) support this contention. In this set of experiments, full and half-versions of figures with denotative surrounds were used and presented in both upright and inverted orientations. The figures created were designed to meet Gestalt principles for establishing the central area as figure in the full versions but to be ambiguous (according to Gestalt principles) for the half-versions. Their expectations were that (1) for the ambiguous half-figures, there would be an exposure duration at which the denotative region of upright versions would be chosen as figure more often (indicating that shape recognition input can facilitate figure–ground segregation when Gestalt variables are missing or ambiguous), and (2) for upright stimuli a dominance of Gestalt variables would lead to an initial identification of the center as figure; a lack of Gestalt variables would lead to initial identification of the surround as figure—for the upright cases; and if both Gestalt variables and shape recognition work together there will be equally many identifications of center and surround as initial figure. Interactions of display time, uprightedness, and Gestalt goodness of the central shape were found. Evidence indicated that shape-recognition routines required about 150 milliseconds and that if Gestalt variables were not strong enough to have isolated figure from ground in this time that shape recognition played a role.

Both Marr (1982) and Gregory (1990) concede that cognition (top-down processing) can be employed to deal with ambiguous situations. Gregory contends that there would be an evolutionary advantage to a system that worked in parallel with preattentive perception coming up with a quick interpretation that is usually (but not always) correct and conceptual processes (sometimes) modifying that initial impression. Such a parallel system seems to be well supported by Peterson et al. (1991) and is the process behind the pattern identification model of cartographic visualization cited above (MacEachren and Ganter, 1990). Based on this model and ideas about grouping and figure–ground segregation detailed here, we can predict that manipulating any design variables that influence strength of contour or heterogeneity of regions will have a dramatic effect on the patterns noticed. This issue will be considered further in Chapter 8.

## Visual Levels

Closely related to the issue of figure–ground segregation is the concept of *visual levels* in graphic illustrations. The theory behind visual levels is that a viewer of a graphic depiction can group sets of objects into common wholes that are seen as occupying different visual (or conceptual) planes. The concept is related to Bertin's (1967/1983) selective and associative principles in that viewers are believed to see different objects as sufficiently similar that they become a unit and different units as sufficiently different that they are visually segregated. Robinson (1960) may have been the first to discus the principles involved in creating visual levels on maps (although he did not use this term). To introduce novice map designers to how a map might be structured so that items will appear at "differing position on the scale of visual significance," Robinson (1960, p. 223) used an analogy to hierarchical outlines for organizing written essays. He then went on to discuss how contrasts of lines, shapes, colors, and value can be manipulated to achieve this structuring.

Michael Wood (1968) took one of the first systematic looks at the idea of visual levels applied to cartography. Wood's (1968, p. 61) objective was to derive principles that would allow a cartographer to place information "on an imaginary scale of distance planes." The "distance" between these planes (i.e., levels), according to Wood, should be based on the similarity of the data occupying them. The separation into planes is required to "provide for focused attention and a good 'gestalt.'" Wood interpreted the concept of visual planes directly in terms of depth perception (which allows humans to segregate a three-dimensional visual field into many depth planes). Wood's goal was to draw on the psychological literature to derive principles for simulating depth planes on two-dimensional maps. Building on research discussed by Vernon (1962), Gibson (1950), and others, Wood proposed ways in which graphic variables such as texture, hue, and value could be used to create depth cues on two-dimensional maps. Although he was able to develop some suggestions for creating visual planes on maps, Wood saw these suggestions as an interim solution to a question that required empirical research to answer more fully. In subsequent work, Wood (1972) specifically considered the application of several Gestalt principles of figure–ground segregation as they relate to separation of visual levels on maps. He cautions, however, that the map viewer's knowledge and assumptions "can easily reverse" the levels intended by the cartographer.

A number of cartographers followed Wood's lead in looking to psychological literature for ideas about how visual levels might be created on maps (e.g., Dent, 1970; Spiess, 1978; McCleary, 1981). Dent (1970) also looked to the graphic design literature, particularly Bowman's *Graphic Communication* (1968). Dent cites Bowman's concept that graphic depictions can have one of three categories of visual depth organization: planar, multiplane, or continuous. Dent then proposes that most maps should use a multiplane strategy in which information is organized into a small set of discrete visual planes. Following Bowman's lead, Dent suggests that techniques employing contrast, aerial perspective (Bowman's dissimilar focus), and overlay can be used to segregate two-dimensional map information into multiple visual levels. He even goes as far as to propose a formula for predicting how contrast and contour sharpness (an aerial perspective cue) interact to produce position in visual depth:

$$PVH_j = f(I_j, ES_j)$$

where $PVH_j$ is the position in the visual hierarchy of object $j$, $I_j$ is the intensity of object $j$, and $ES_j$ is the edge "sharpness" of object $j$. This hypothesized formula was never empirically tested. Based on more recent evidence concerning brightness contrast cited above, at least one major flaw in the hypothesis seems apparent without testing. An understanding of visual organization suggests that perception of individual features on a map will not happen in isolation from other map elements. Perceptual representations are inferences based on processing relative intensities in the visual field. The intensity (i.e., brightness) of an object, therefore, is probably not related directly to prominence in the visual field. Contrast

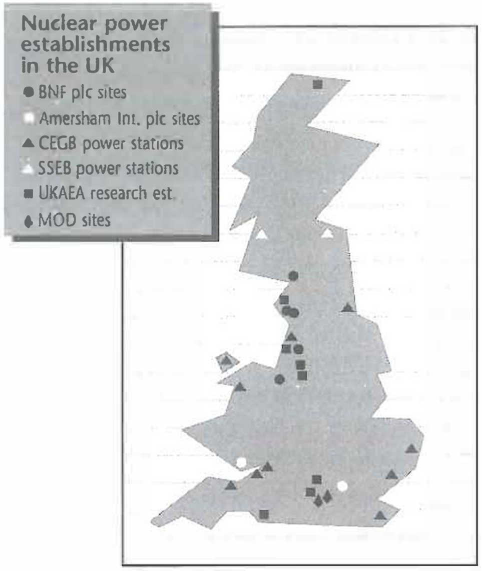

**FIGURE 3.61.** An example of a map with four (or more) visual levels: a base, areas on the base, point symbols on the areas, and a legend (which has its own levels) in "front" of all map elements.

in intensity of an object from its surroundings should, instead, be used. In spite of this flaw, the formula Dent offered is intuitively appealing and could be tested quite easily, but no one has yet done so.

Another untested hypothesis concerning visual levels on maps was proposed by Spiess (1978) and included in the ICA-sponsored text on *Basic Cartography* (Spiess, 1988). Spiess (1988) contends (as if it were a fact) that no more than three visual levels should be attempted on maps. This view challenges earlier cartographic proposals by both Wood (1968) and Dent (1970) who suggest that at least four visual levels are possible and Bowman's (1968) contention in relation to information graphics in general that continuous as well as multiplane organization is both possible and useful in some cases (Figure 3.61). Particularly in light of recent technological developments that allow binocular depth cues to be added to the cartographic tool kit, three visual levels for maps seems unduly restrictive. Whether we claim three, four, or more visual levels as a practical maximum, however, visual hierarchies can clearly exist within each level (e.g., a road crossing a stream). We could easily make a case (based on this kind of evidence) for ignoring visual levels altogether and treating visual hierarchy as a continuum. As considered in Part II, however,

grouping categories of features into a small number of levels facilitates a semiotic approach to development of a map syntactics (see Chapter 6).

## PERCEPTUAL CATEGORIZATION AND JUDGMENT

Underlying perceptual organization are processes of categorization. Hoffman (1989, p. 84) contends that "whatever else it does, the brain must be able to simplify and categorize the structures in the patterns it processes." At the most fundamental level, a bipartite categorization into same–different (i.e., discrimination) is required for early vision to isolate perceptual objects and organize them into groups. All Gestalt principles for grouping or figure–ground segregation are dependent upon heterogeneity of perceptual objects. If there are no differences between objects and background, we will see no objects. If differences are absent among objects, there will be no groups (Figure 3.62).

When vision discriminates between elements of the visual scene, it also generally orders those elements in some way (e.g., one is lighter than another, of coarser texture, longer, closer, etc.). This tendency to order is related to perceptual organization research devoted to visual attention. Items that appear higher on some ordered scale are likely to be more noticeable and thus are probably more often attended to. Eye movement research with maps, for example, has documented the intuitive notion that large map symbols attract more attention than small ones. In addition to ordering perceptual units, psychophysical tests suggest that the judgment of magnitude is also a preattentive process, at least for size and brightness.

The output of perceptual organization can also be the input for further categorization processes. Detection and discrimination among perceptual objects and groups as well as identification of order or relative magnitude are required if features of a visual scene (i.e., map symbols) are to be assigned to more specific categories. From the perspective of an in-

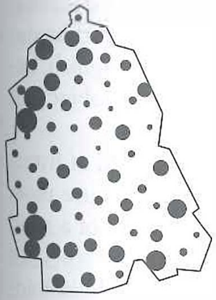

**FIGURE 3.62.** A map from research by Slocum for which subjects produced no consistent regionalization (i.e., groupings of circles). In this case, differences exist, but they are small. *Reproduced from Slocum (1983, Fig. 9, no. 13, p. 71). Adapted by permission of the American Congress on Surveying and Mapping.*

formation-processing approach to vision and visual cognition, then, detection, discrimination, and ordering on maps are important in relation to perceptual organization of map marks making up symbols, patterns, and regions and in relation to categorization of map symbols (an issue addressed in more detail in Chapter 4).

Cartographers have directed attention to the empirical examination of map symbol discriminability and have devoted considerable energy to a search for "laws" by which judgment of order and estimation of magnitude are performed. The driving force behind these efforts has been the map engineering goals of determining "least practical differences," making the order of ranked information intuitively obvious, and scaling symbols to match perceptions. These goals, first identified by Robinson (1952), have been approached from a communication perspective—for example, if a cartographer wishes to communicate that there is a categorical difference between two map elements, what physical difference is required to ensure that most people will notice it (and interpret it correctly). In spite of this limited perspective, past results can, to some extent, be put in a broader information-processing context where they might inform work in perceptual organization of maps as well as work dealing with symbol categorization, identification, and interpretation. No attempt is made here to provide a comprehensive review of cartographic research on discriminability, apparent order, or magnitude judgments. What is provided are a few key examples to illustrate how this research might be fit into a broader (thus potentially more useful) context. A sampling of recent ideas from the psychological literature is also described to help link the psychophysical approach taken in much of the cartographic research with the overall cartographic-representation perspective presented in this book.

## Detection

Discrimination is the ability of vision to recognize a difference. Detection is, in essence, a discrimination problem in which a viewer must discriminate between some signal and the background on (or in) which that signal appears. For some purposes, however, it is useful to distinguish between detection (the ability to notice the presence of an object or feature) and discrimination.

For areas, there is at least one detection issue that has been of interest: detecting texture of area fill patterns. This is an issue because most area fills on maps to be printed (even if the final appearance is of a solid color) are made up of a textured pattern.[9]

One of the things that has become clear from both neurophysiological and psychophysical research is that the visual system is relatively insensitive to high frequency (fine) patterns. This allows us to create the impression of flat gray tones from patterns made up of fine dots. Castner and Robinson (1969) were among the first cartographers to investigate the perceptual thresholds involved. If patterns are coarser than about 40 lines (or dots) per inch, we see them as predominantly a pattern (at normal reading distance) (Figure 3.63). Patterns between 40 and 85 dots per inch are ambiguous (as with ambiguous figures in figure–ground research, these fills can be seen as either gray or textured, but not both at once). We can recognize the pattern easily, however, it does not dominate our impression, and we can see a value difference as well. Above 85 dots per inch we no longer notice the texture (unless we are consciously trying to see it—as you probably are now). We see these patterns generally as a color value or a gray tone. If a viewer is consciously trying to detect a texture, however, we have to go to almost 300 dots per inch before our visual system becomes incapable of detecting individual dots in foveal vision.

In contrast to Castner and Robinson's thorough examination of texture detection, cartographers have given little attention to questions of point or line symbol detection. Keates (1982) discusses the issues and suggests that detection will be a combined function of symbol size and contrast with the background it appears on (generally a contrast of color value or hue). Viewing distance is an additional factor that is often ignored because it is assumed to be normal reading distance (an incorrect assumption for route maps posted in public places, slide presentations, etc.). At normal reading distance with high-contrast symbols (e.g., black on white), size is not an issue because lines or symbols too small to be seen cannot be consistently printed (at least on normal porous paper). It is when color hue or value differences are used that detection of point and line symbols can be compromised. Although Keates (1982), Spiess (1988), and most authors of cartographic texts caution cartographers to use symbols with sufficient contrast to the backgrounds they appear on, no guidelines exist because little empirical testing has been done.

Noting the importance of low-level vision to any higher level pro-

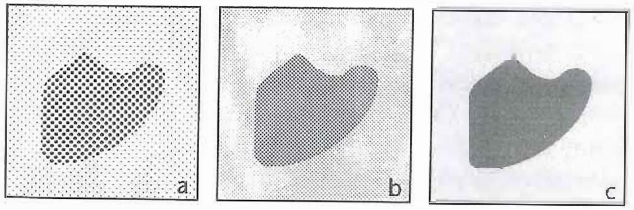

**FIGURE 3.63.** Map regions with 35 lines/inch dot fills (a), 65 lines/inch dot fills (b), and 133 lines/inch dot fills (c).

cessing of map information, Dobson (1985) advocates a concerted effort by cartographers to investigate questions of symbol "conspicuousness." Vision research offers a few hints about potential detection problems on maps, as well as about particular issues that deserve empirical research. The structure of our eye results in a rapid decrease in visual acuity from the fovea to the periphery. Detection will obviously be best for features (e.g., map symbols) that we are looking directly at. With map search tasks, however, in which the map user wants to find an occurrence of a particular feature, symbols must be detectable with peripheral vision if the search is not to be painfully slow. Engel (1977) has demonstrated that increased contrast can increase detectability in peripheral vision.

For maps, black symbols on white backgrounds are usually detectable. It is when color is added that detection problems become likely. Travis (1990), for example, points out that about 8% of men are congenitally red–green color deficient. For these map viewers, detection of symbols on backgrounds and discrimination between symbols can at times be impossible if this problem is not taken into account. Thus far, Olson (1989) seems to be the only cartographer to explore the issue of color deficiency empirically. Based on her research, she devised some guidelines for color choice that should limit hue detection and discrimination problems for the color deficient.[10]

Beyond issues of color deficiency, all humans have differences in acuity for different hues. Because our eyes have no blue cones in the fovea, our ability to detect blue map symbols is reduced over other hues. As Robinson (1967) noted, the traditional choice of blue for coastlines and depth contours is a poor one in situations for which quick discrimination of these features from their backgrounds is critical (e.g., navigation charts). For logical reasons, and some not so logical, cartography is a tradition-bound discipline. We are therefore unlikely to see a sudden change in color of depth contours or coastlines, in spite of empirical (and neurophysiological) evidence favoring such a change. In producing a detailed map of Georges Bank, for example, cartographers at Woods Hole Oceanographic Institution conducted extensive tests of detectability and discriminability of various point and line symbol colors on a range of background colors, but did not even bother to test anything other than blue for depth contours (Woods Hole Oceanographic Institute, 1982).

Humans are particularly sensitive to motion. We can detect motion of a few seconds of arc (Mowafy et al., 1990). This ability can be put to use in animated displays—if time is available. Similarly, humans are quite sensitive to aspatial change. As Travis (1990, p. 431) notes, "From neon signs in Las Vegas to the blue light atop an ambulance, flickering lights are used in our society to gain attention." On otherwise static maps, blinking symbols can be used in symbolic ways to highlight important features (MacEachren, 1994b). Examples of blinking symbols have been used by a number of animation authors (DiBiase et al., 1992; Monmonier, 1992).

## Discrimination

Discrimination (in its usual sense of noticing a difference between two perceptual units rather than between one unit and its background) has been investigated by both psychologists and cartographers using two experimental paradigms, one based on same–different tasks for stimulus pairs, the other on visual search tasks for a target on a background of similar features. Visual search tasks are a less direct measure of discrimination in situations for which the target is both conceptually and visually different from the nontargets. When differences are strictly visual, however, search tasks provide a direct measure of discriminability (Uttal, 1988).

Using visual search methods, psychologists have uncovered a rather unexpected aspect of how vision discriminates. There appear to be natural categories (e.g., circles) from which vision will "notice" differences more readily. Discrimination seems to be asymmetrical. Within a dimension or feature class, some values appear more likely than others. Deviations from these standard values are more noticeable in relation to the standard than the reverse. For example, curved lines among straight lines are discriminated more quickly than straight among curved lines as are tilted lines among verticals, circles with gaps among whole circles, and ellipses among circles (Treisman et al., 1990).[11]

### *Text Discrimination*

For maps, the same–different experimental method has been particularly useful in studies of place labels. Shortridge (1979) used this method to develop guidelines for the minimum point size difference necessary for text labels to be noticeably different (Figure 3.64).[12] In collaboration with Welch (Shortridge and Welch, 1980, 1982), she went on to investigate how both the experimental methodology and the features of town labels other than point size influenced discriminability. In the first follow-up study, they focused on discriminability of dot symbols of different size used to indicate town location. They found that larger size differences were required for discrimination if a simple same–different task was posed than if the task was for subjects to indicate the larger dot. In addition, if subjects were led to expect some dot pairs to be the same (as they would when viewing a typical map), the difference required for discrimination

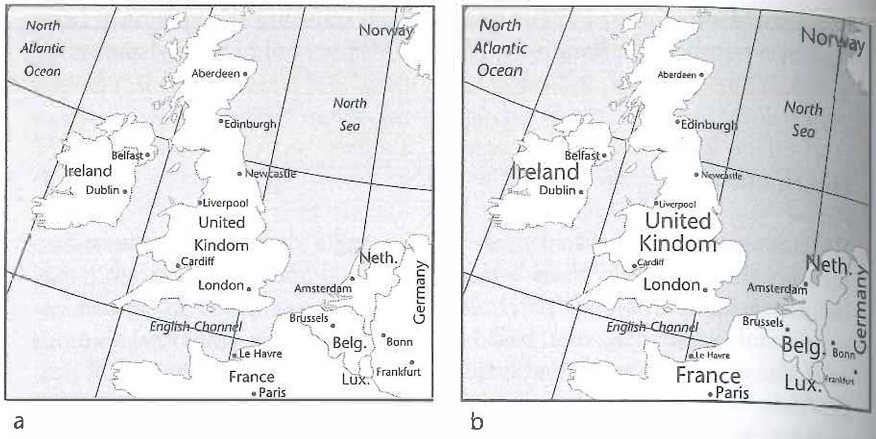

**FIGURE 3.64.** Based on Shortridge's results, the label sizes on map *b* will be distinguishable for about 75% of readers, while those on map *a* will not. How many sizes do you see?

was even larger. These results indicate that order can be perceived, even in cases for which discrimination is marginal or uncertain, and that expectations have an effect on discrimination—an indication that discrimination is not an entirely bottom-up preattentive process, as is often contended.

In their second follow-up experiment, Shortridge and Welch (1982) examined the issue of whether multiple feature differences between stimuli increase their discriminability. They measured discriminability of place labels distinguished on the basis of point size, type boldness, type case, and location-dot size, individually and in all possible combinations. They found that feature combinations increased discriminability up to three features, but a fourth had no impact. Interestingly, the discriminability of feature combinations was not predictable on the basis of their independent discriminability. Type size and boldness, for example, was a more discriminable combination than type size and case, in spite of the fact that case was considerably more discriminable by itself than boldness (which was the least discriminable feature). This lack of additivity suggests a holistic level of processing for these feature conjunctions.

### *Point Feature Discrimination*

Discriminability of text could be considered a special category of discrimination of point features on maps. A number of psychologists have looked at the interaction between point feature discriminability and visual

search for those point features. Quinlan and Humphreys (1987), for example, compared search tasks in which subjects searched for a single-feature target, two different single-feature targets, and a conjunction target combining the two features. Their evidence demonstrated that the rate of conjunction searches is influenced by discriminability of features, but the kind of search process used is not. Regardless of how discriminable the features, they found conjunction searches to proceed in serial fashion, while single-feature searches were executed in parallel when symbols were sufficiently different. Treisman (1988), however, offers some evidence that when two highly discriminable sets of distractors are present, attention can be directed to subgroups of items as a whole and, using inhibition of feature categories that cannot be the conjunction sought, can limit the search in such a way that a parallel process can be used.

There have been several studies that have dealt more directly with discriminability of positional symbols for maps. Most of these studies have measured "confusability" of symbols—a measure that does not separate discrimination from identification (assigning a label to symbols). An example of this kind of study is Johnson's (1983) empirical evaluation of the National Park Service point "symbol" set. He had subjects match symbols with labels (both with and without a legend present). A confusability index of sorts was devised based on the number of misidentifications for each symbol (Figure 3.65). Subjects were then presented with a visual search task in which the number of correct identifications in a limited time was determined. Those symbols that were highly confusable and were judged to be so because they look alike (rather than because they refer to similar things) also rated low on the visual search task (e.g., the

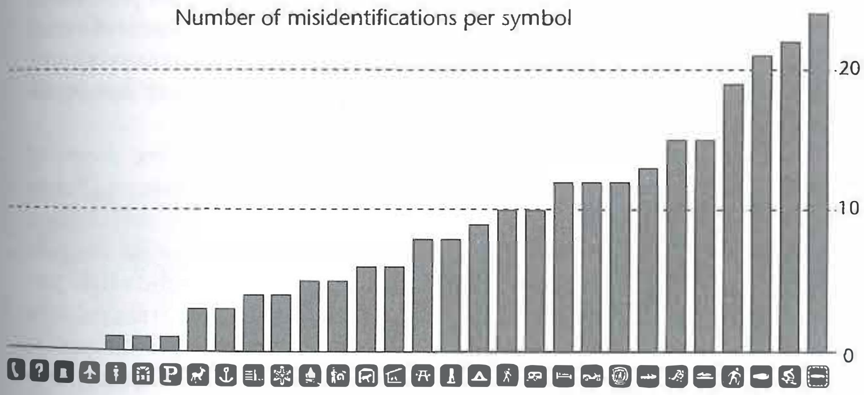

**FIGURE 3.65.** The part of the National Park Service symbol set that was tested, with number of misidentifications per symbol indicated. *Derived from Johnson (1983, Fig. 23, p. 112).*

lighthouse and the service station symbols). In a similar study with four alternative sets of symbols for tourist maps, Forrest and Castner (1985) found that iconic symbols took longer to locate than abstract symbols, but that fewer identification errors resulted. In addition, they found that the advantages of iconicity (for identification) and of simple abstract shapes (for visual search) could be combined by bounding iconic symbols with geometric frames (triangles, circles, and squares).

### *Pattern Discrimination*

Discriminability of point features is probably most critical in situations where visual search is demanded. In contrast, discriminability of map patterns is probably most important when map readers are faced with general tasks related to pattern analysis (e.g., identification of homogeneous regions).

Based on Gestalt grouping principles, we would expect similar elements in close proximity to group and be seen as a whole. In order for patterns to be discriminable, then, grouping must occur for subsections of the visual scene. Patterns in which one feature (i.e., visual variable) is altered in an obvious way are usually quite discriminable, but patterns differing on a conjunction of features or rearrangement of subcomponents of elements composing them are not (Beck, 1966; Treisman, 1985). For patterns made up of coarse textures of the type used to depict qualitative data on maps, there seems to be a good understanding. Because of their probable importance in early vision (leading to Marr's primal sketch) considerable attention has been directed to how texture boundaries are identified and to algorithmic approaches to solving the same problem in computer vision. Malik and Perona (1990) compared the results of visual search measures of pattern element discriminability with a computational approach to texture discrimination and found nearly perfect correspondence (Figure 3.66).

Julesz (1975) has developed perhaps the most complete theory of texture discrimination. He posits three levels of discrimination (Figure 3.67). First is discrimination on the basis of darkness (or percent area inked). Second is discrimination of the characteristics of pattern arrangement. Both of these processes operate at the global level of the whole pattern. At a third level is discrimination of local geometry. Area patterns for maps, then, could be expected to be most discriminable if differences exist at all three levels. Global-level differences should allow quick parallel processes to be used in discrimination. Patterns that differ only at the local geometry level should be rather difficult to discriminate and will probably require serial processes in which patterns are closely examined one at a time.

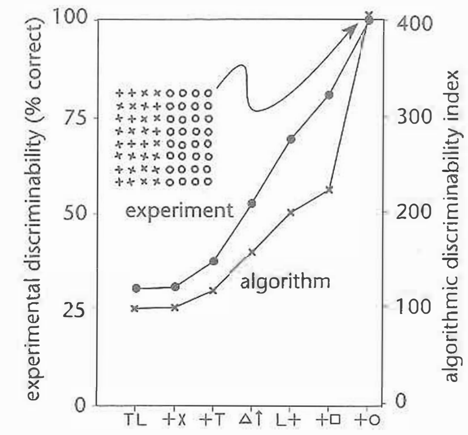

**FIGURE 3.66.** A comparison of texture discriminability determined by experiment and by computations. Texture pairs are those evaluated by both Kröse (1987) and Malik and Perona (1990), using experimental and computational procedures, respectively. The most discernible texture pair is shown as an inset on the graph and elements making up all texture pairs appear as labels on the X-axis. *Derived from Cleveland (1993, Table 1, p. 335).*

Julesz (1981) has formalized the mathematical description of his three levels of texture differentiation and demonstrated that the preattentive visual system is unable to process statistical information beyond the second order. It is possible, however, to create patterns that are discriminable even though they have identical first- and second-order statistics. Discrimination in these cases involves local conspicuous features that Julesz calls "textons" (elongated "blobs" of specific widths, orientations, and aspect ratios). This research, which Julesz has linked to Marr's

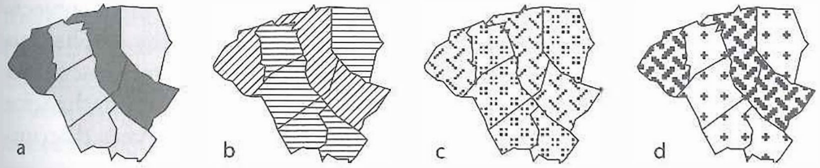

**FIGURE 3.67.** Discrimination of two map areas by value (a), by pattern arrangement (what cartographers would term orientation) (b), by local pattern geometry (c), and by the combination of all three (d).

primal sketch model, offers a sound basis from which patterns for use in interactive visualization might be devised (see Chapter 8 for discussion of one attempt to do just that).

Issues of texture discrimination have particular implications for design of area fills to be used in depicting qualitative information on maps. In this case, cartographic and perceptual logic suggests that patterns should avoid the use of differences in percent area inked because these will be seen as ordered. Without percent area inked, however, only pattern arrangement and local geometry are available.

### *Color Discrimination*

In relation to color discrimination, Luria et al. (1986) point out that while color hue discriminability is "truly astronomical," the discriminable number of colors drops rapidly as their number in the scene goes up. These authors cite results of 98% correct discrimination among 10 colors, dropping to 72% for 17 colors. In spite of the fact that color discriminability may be orders of magnitude more limited than simple same–different experiments have suggested, there is considerable evidence that for visual search tasks, symbols that differ from others by hue are much more discriminable than those differing by either shape or size (Williams, 1967). In addition, for symbol conjunctions of color and shape or color and size, color seems to act as the dominant cue (Eriksen, 1952). Williams (1967) has provided evidence from eye movement studies that for these conjunction searches, subjects fixate on targets of the specified color to determine whether they are the correct size or shape rather than fixating on targets of a specified size or shape to check their color. Quinlan and Humphreys (1987), in a visual search experiment involving conjunction targets, came to a similar conclusion.

In graphic applications, Lewandowsky and Spence (1989) have demonstrated that discrimination of different variables on a multivariate scatterplot is higher for point symbols of three different colors than for three different geometric shapes or three different letters. When subjects were asked to estimate the correlation of variable pairs on the graphs, this difference in discriminability resulted in novices having more accurate correlation estimates with scatterplots using color than experts had for scatterplots using three letters (that were not individually very discriminable).

In relation to maps, Forrest and Castner (1985) cite an unpublished study by DeBrailes confirming the dominance of color in discrimination of point symbols for visual search tasks on maps. Although Forrest and Castner, along with other cartographers, argue that varying hue of point symbols will be a particularly good idea on maps where visual search is expected (e.g., travel maps, navigation charts, etc.), none have considered the fact that there is a pronounced male–female difference in color acuity, and that for both males and females that acuity drops with age.[13]

Neurophysiological evidence suggests that discrimination on the basis of color or luminance contrast is a lower level visual process than is estimation of luminance. Shapley et al. (1990) argue that early vision computes contrast (not reflectance as Land and McCann, 1971, had predicted). Experimentation with cats (a species whose early visual processes are considered quite similar to those of humans) has indicated that "the response of retinal, lateral geniculate and some primary cortical neurons is proportional to contrast over a low-to-medium contrast range, and then may saturate at high contrast" (Shapley et al., 1990, p. 435). This emphasis of early vision on contrast rather than reflectance helps explain phenomenon such as color constancy (that we see a color as the same under various lighting conditions), simultaneous contrast (that perception of a color or shade will change due to the background it is on), and assimilation (the additive effect on brightness of an object produced by the brightness of its background). The emphasis on contrast over reflectance may also relate to the important role contrast appears to have in segregation of figure from ground and the equivocal results that have been obtained when attempts are made to determine whether light or dark areas are most likely to be seen as figures (see discussion of figure–ground above).

### *Motion Discrimination*

As we would predict based on the idea of indispensable variables discussed earlier in this chapter, human vision is very sensitive to motion (for which both location and time are changing). In relation to motion, the Gestalt principles suggest that the common fate of objects moving together will allow a viewer to visually group those objects and discriminate them from their background. Just as we can visually group objects moving together, however, evidence suggests that humans are very sensitive to constancy of spatial distance between moving edges (Mowafy et al., 1990). Their results indicate that discrimination between coherent and uncorrelated motion can be achieved with similar levels of accuracy to the ability for detecting any motion at all (i.e., changes of a few seconds of arc). Mowafy et al. (1990, p. 591) contend that "processing relative movements in the environment is a fundamental characteristic of human motion perception." Therefore the evidence for this ability to discriminate coherent motion has good evolutionary support. Ability to detect motion and discriminate between coherent and noncoherent motion has obvious implications for the design of animated maps. For example, on a map depicting flows, we might expect viewers to be attracted to even small deviations (in speed or direction) of a single moving arrow from the flow of a group. No cartographic research has been directed to this or related issues.

## Judging Order

According to Shapley et al. (1990), one of several fundamental "facts" of perception is that if objects or patterns can be discriminated, we can usually also assign an order. Another "fact" (that they admit they have little scientific evidence for) is that there are "natural continua" along which discriminations are easy (e.g., larger–smaller, left–right), and that we can expect discrimination to be hard along non-natural continua (e.g., alphabetical order).[14]

For maps and other graphics, Bertin (1967/1983) contends that humans find inherent order in spatial location, size, color value, and texture. DiBiase et al. (1992) point out that time is inherently ordered as well, and that for animation temporal order provides one of three dynamic variables, all of which are intuitively ordered (the other two being duration and rate of change).[15] In *Some Truth with Maps* (MacEachren, 1994a), I suggest that color saturation and focus are ordered graphic variables that Bertin omitted from consideration and that color hue and orientation are marginally ordered. Few of these contentions, however, have been examined experimentally.

Most of the attention to whether various graphic variables used on maps are intuitively ordered has been directed to color hue, color saturation, and color value for area fills. The cartographic goal of the research has been to determine color sequences appropriate for choropleth and other quantitative maps. Experimental tasks used in this research have not allowed preattentive processes to be segregated from attentive ones. Whether the process of seeing the order of various hue-value sets is a logical–cognitive one or a purely visual one, therefore, cannot be determined from the evidence. Results do support the contention that color value and color saturation are ordered and that hue is not (Cuff, 1973; Gilmartin, 1988). In contrast to Cuff's empirical results, the phenomenon of advance and retreat is expected to cause red to appear to be located on a visual plane in front of blue. Travis (1990) offers three physiological explanations for this effect: (1) that because of chromatic aberration the eye's lens causes short wavelengths to have a shorter focus than long wavelengths, (2) that the visual and optical axes of the eyes do not coincide, and (3) that the apparent brightness of light depends on the point of entry through the pupil. He contends that using saturated red and blue will make objects literally "stand out" from the display.

A recent study, by students of mine, seems to favor the advance and retreat hypothesis and contradict Cuff's findings (Bemis and Bates, 1989). For hypothetical temperature maps, subjects were found to consistently see an order in shaded isotherm maps using a bipolar range of colors (with blues at one end and reds at the other). They proposed an interesting explanation for the contradiction between their results and Cuff's. Since 1973 when Cuff collected his data, a blue–red range has become much more common for temperature maps (e.g., on television news and in many newspapers). Bemis and Bates (1989) contend that the logic of the order has been learned, an explanation that is intuitively appealing.[16] A further test of this hypothesis would be to assess the relative order seen in value versus spectral series using reaction-time methods. If an identification of order for color value is a preattentive process while identification of order in a color hue sequence is a cognitive process, reaction times to judge "higher" values should be faster than reaction times to judge "higher" hues, and the presentation time threshold at which order can first be judged should be much less for value than for hue.

Although several studies have found value ranges to be judged as ordered (with dark values usually seen as the high end of the scale), this ordering is not perfect even for simple maps. McGranaghan (1989), for example, had subjects judge which of two states on a map of the western United States had the higher data value. While the darker of the states was selected as "more" in a majority of cases, only 30% of the subjects consistently saw darker as "more." Most of the inconsistent ordering occurred when the map's background was gray or black rather than white, with the black background resulting in about twice as many intransitivities as the white and the gray resulting in about four times as many.

## Judging Relative Magnitude

Early perceptual research in cartography was devoted almost exclusively to attempts at deriving functional relationships between physical magnitude of different aspects of map symbols and psychological magnitude. McCleary (1970) reviewed this research and suggested that the one generalization that seemed to apply across the board was that map readers underestimate differences between map symbols. The precise functional relationship seemed, at first, to depend primarily upon the particular stimuli being tested and the questions asked (e.g., Olson, 1976). It gradually became clear that there is also substantial individual (subject) variation in magnitude judgments (McCleary, 1975; Griffin, 1985) and that map context created further problems (Gilmartin, 1981).

Based on the extensive testing Flannery did in 1956 and repeated in 1972, Robinson et al, (1984) and other authors of cartographic texts adopted the guideline of adjusting scaling on graduated circle maps to account for underestimation of differences. Others (e.g., Cox, 1976) have suggested that we might be better off simply providing more anchors (in the form of legend circles). The among-experiment and among-subject variations together with context effects have made many practitioners suspicious of the empirically derived guidelines for perceptual scaling and today it is doubtful whether many cartographers actually use them.

In relation to gray tones for quantitative maps, there seems to be a bit more consensus. Kimerling (1985) was able to demonstrate a correspondence among what were apparently divergent results and showed that usable gray scales could be devised. The two most significant issues he considers are the interaction between area fill texture and perception of value and the interaction between judgment task and value perception. In terms of texture, Kimerling found that the finer the texture, the more curvilinear the relationship between perceived and actual gray tone. The implications of this finding are that a different set of gray tones is required for maximum discriminability if a map is produced on a laser printer (with dots spaced at about 60 lines per inch) versus on a film recorder (with dots at 100 or 120 lines per inch) (Figure 3.68). Judgment was also found to be dependent upon the visual task, with a different actual–perceived gray tone function for judgment of percent black versus a partitioning task or tasks leading to a set of maximally discriminable gray tones.[17]

## PERCEIVING DEPTH FROM A TWO-DIMENSIONAL SCENE

Closely related to concepts of judging order and magnitude (as well as to the visual levels discussed above) is the simulation of depth in two-dimensional displays.[18] Vision is designed to deal with a three-dimensional world. Interpreting depth in a visual scene is a complex process that appears to be facilitated by a large number of interdependent cues. For good evolutionary reasons, vision does not require all depth cues to be present in order to interpret features of a scene as being at varied distance from the observer. This makes it possible to trick vision into interpreting a map or other display as three-dimensional by combining some of these cues in appropriate ways. How vision interprets depth cues is relevant to cartography because cartographers are often faced with the problem of simulating three-dimensional information on a two-dimensional display (when depicting terrain, but also for more abstract multivariate information).

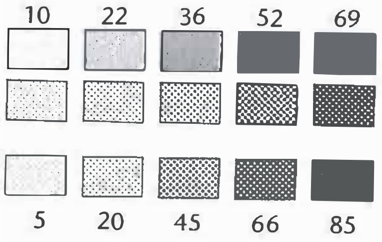

**FIGURE 3.68.** A gray scale designed for maximum between-category contrast with production on a 100 lines/inch image setter (top) compared to the same grays produced at 45 lines/inch resolution (typical of laser printers) (middle), and to grays adjusted in color value to achieve maximum contrast at the coarser resolution (bottom).

### A Taxonomy of Depth Cues

Kraak (1988) provides a taxonomy of cues for depth perception and a review of the cartographic literature relevant to each. His taxonomy distinguishes between "physiological" and "psychological" depth cues. Some authors have called the latter "pictorial." Since this latter term puts emphasis on characteristics of the display rather than of the cognitive processing of that display, it is adopted here. The physiological depth cues have to do with the physical processes of vision as it reacts to the real three-dimensional environment. Pictorial cues, in contrast, are those related to the object's structure and the way that structure organizes visual input. In the context of computer graphics, Wanger et al., (1992) provide a list of depth cues similar to those cited by Kraak. Each of these sources includes pictorial cues (or subcategories of cues) omitted by the other and disagree on whether motion parallax should be considered a physiological or a pictorial cue (with Kraak opting for the former, and Wanger et al. for the latter). If we look to the art literature, we find additional depth cues not included in either the cartographic or the computer graphic taxonomies (along with some differences in terms for cues in common) (Metzger, 1992). A composite of these sources results in the following taxonomies of depth cues that may be relevant to maps:

Physiological

*Accommodation*: A change in thickness of the eye's lens as it focuses on an object.

*Convergence*: The difference in angle of gaze by the two eyes focused on the same object.

*Retinal disparity*: The difference in image (visual array) derived by each eye (which has a slightly different point of view).

Pictorial

*Perspective*: Kraak (1988) subdivides perspective into four components, and we will follow this subdivision here.

- *Oblique projection*: Representation of a scene from a viewpoint that is not an elevation (profile) or plan view (overhead) suggests a three-dimensional solid, thus depth.
- *Linear perspective*: Lines that are parallel in reality seem to converge with distance (e.g., a pair of railroad tracks).
- *Retinal image size*: Objects appear smaller the farther away they are.
- *Texture gradient*: Texture appears to decrease with distance.

*Motion*: Movement (actual or simulated) of the observer's point of observation produces changes in the relative retinal displacement of objects at different distances. Successive presentation of static images in which objects are displaced relative to one another can (particularly in the presence of other cues) also result in a sensation of depth.

*Interposition*: Using Gestalt principles of good continuation, vision will assume that whole objects juxtaposed with what appear to be part objects are really whole objects blocking our view of other whole objects farther away.

*Shadow*: A cue to obstruction or overlap, indicating that one object is blocking light from falling on another object.

*Shading*: Illumination gradient can indicate the shape and orientation of a surface.

*Color*:

- *Chromostereopsis* (also called color stereoptic effect or, more commonly, advance-and-retreat): The differences in wavelength of colors are thought to result in apparent differences in distance (with reds appearing closer than blues at the same true distance).
- *Aerial perspective*: With distance colors become less distinct (less saturated) and lighter (higher value), often with a bluish tint due to atmospheric scattering.

*Detail*: With distance detail becomes less visible and edges become blurred.

*Reference frame*: In order to judge relative size, vision must match retinal size to some frame of reference—apparent distance will therefore vary with what an object is compared to.

Not surprisingly, the bulk of cartographic attention to depth perception and how specific cues might prompt this perception is related to terrain mapping. Terrain is three-dimensional and cartographers have struggled with collapsing those three dimensions onto a two-dimensional page since the earliest maps were made. Although contour lines involve no depth cues, virtually all other methods of depicting relief rely on one or more of the cues listed above. Simulation of three dimensions on maps can be grouped into techniques that involve physiological cues, that rely on perspective, that use static nonperspective pictorial cues, and that include motion. For motion to cue depth, the user must assume a perspective view (but linear perspective is not essential). Since the possibility of motion as a depth cue requires a dynamic display, further discussion of these cues will be postponed until Chapter 8 in the context of geographic visualization environments (which, as they will be defined here, are dynamic).

## Applying Depth Cues to Maps

### *Physiological Approaches*

Computer technology has facilitated production of displays that make direct use of binocular parallax as the primary depth cue. Such displays consist of pairs of representations, usually perspective views, that depict the mapped area from slightly different points of view (simulating the different points of view resulting from the spacing of our eyes). Seeing depth in stereo pair maps usually requires that the observer's head does not change position while viewing, and/or that special glasses be worn. One technique, referred to as anaglyph plots, uses opponent colors of red and green to produce two overlapping views. When an observer wears glasses having one red and one green lens (if she has normal color vision) the two views will be separated with one seen by each eye. This technique was used for maps at least as early as 1970 in the Surface II package that could generate anaglyph fishnet maps.

### *Perspective Approaches*

Included here are the four perspective cues of oblique projection, linear perspective, retinal size, and texture gradient. These cues are typically manipulated together on perspective view maps, with oblique projection common to all. Different representational techniques can put uneven emphasis on the remaining three perspective cues. The well-known fishnet plot (Figure 3.69), for example, emphasizes texture gradient. In contrast, layered contours ( Figure 3.70) and block diagrams emphasize linear

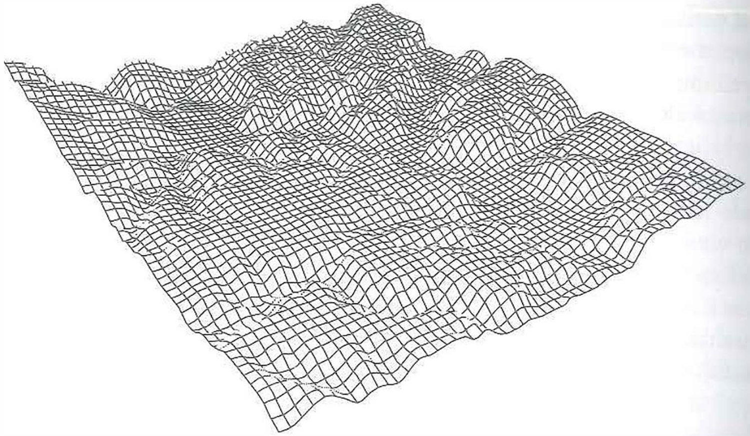

FIGURE 3.69. A typical fishnet plot depicting the terrain around Johnstown, Pennsylvania.

perspective and size disparity, and solid modeling (Figure 3.71) emphasizes linear perspective with shading and shadow as additional (nonperspective) depth cues. All of the methods mentioned make use of interposition as an additional cue (e.g., fishnet plots are rarely generated without hidden line removal). I have uncovered no empirical comparisons among the various styles of perspective map, but some attention has been given to perception of fishnet plots.

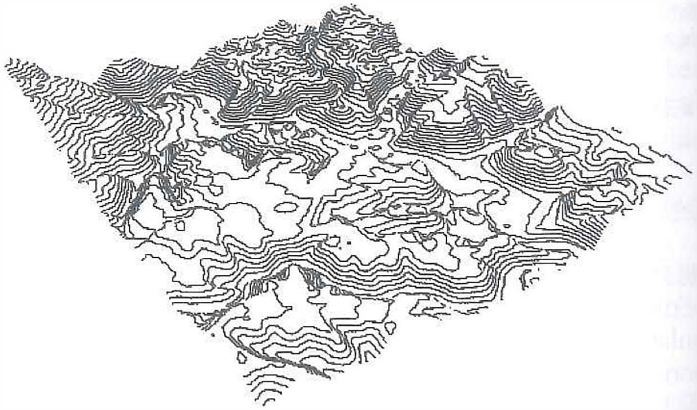

FIGURE 3.70. Layered contours applied to the same region as shown in Figure 3.69.

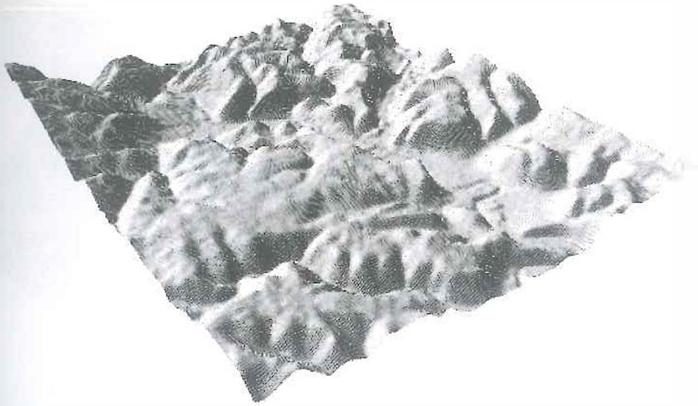

FIGURE 3.71. Solid rendering of the region from Figure 3.69.

That fishnet plots, with their strong texture gradient, do work was convincingly demonstrated by Rowles (1978). She found that subjects were able to judge relative height quite accurately, even when the point of view for the perspective was as high as 75° (nearly overhead) or as low as 15° (Figure 3.72). The view from 15°, however, results in considerable occlusion of map sections, something that probably helps to cue depth but can make the map much less useful (unless it can be dynamically oriented to allow hidden locations to be uncovered).

### *Nonperspective Approaches*

Whether or not fishnet and other perspective view maps are effective, they all suffer from two problems. No matter what point of view is taken, there will be some hidden features and (if linear perspective is used) scale will change across the map. To avoid these issues, considerable attention has been given to use of nonperspective depth cues with the goal of an effective plan-view relief representation that suggests depth. Most cartographic attempts to create the illusion of depth without perspective use shading and/or color.

With shading, there is a long history of manual techniques using pencil, airbrush, and other tools. The procedures for what is termed "plastic relief"[19] borrow from principles of light and shadow in art and psychological principles of depth perception, but to be effective must also incorporate considerable knowledge of geomorphic structure of the terrain

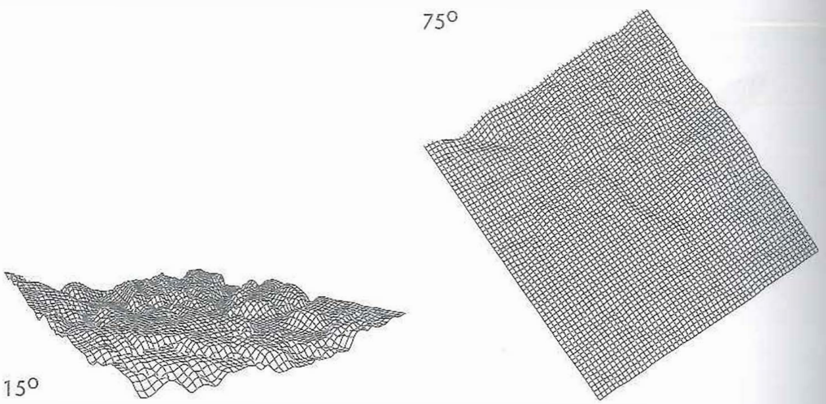

**FIGURE 3.72.** The Johnstown fishnet terrain map shown at elevations of 15° and 75°, viewpoints for which Rowles (1978) found no significant decrease in ability to estimate height. In Rowles's examples, however, relative relief was much greater.

being represented. Imhof (1965/1982) provides the most comprehensive account of these methods. One of the things that has been learned (primarily from long experience and years of marginal success at computerizing the process rather than from empirical research) is that perception is sensitive to what might be called the texture of shading as well as to its value. Humans can immediately recognize the difference between a perfect match of shading with slope–aspect values and shading that looks real. Not only do real surfaces not reflect light as perfectly as a virtual computer surface can, the real environment has complex interactions of direct with reflected light that our visual system has evolved to expect. For terrain shading to look real, it must incorporate at least some of the subtle variation from perfect reflectance that occurs in the real environment. Many theories have been proffered for the ideal reflectance model, but little empirical research has been done to determine their relative merits. In spite of the lack of empirical research, plastic shading has developed to the point in cartography that it has been successfully modeled with computer software (Figure 3.73). Perhaps the most impressive result thus far is Pike and Thelin's (1989) digital relief map of the United States, described by Lewis (1992) as a "cartographic masterpiece."

One issue that all disciplines interested in shading as a depth cue seem to agree upon is that the simulated light source needs to be from above the scene, and above-left is usually cited as best. This phenomena seems to be based upon a schema (or expectation) that light in the environment is from above. When applied to art (e.g., in the representation of a vase of flowers on a table or a figure in repose) this light-from-above rule is quite logical. On a map, the rule results in light from the northwest, a direction that is at odds with reality in the northern hemisphere. In spite of the physical impossibility of the scene, humans consistently treat terrain shading on maps in the same way that they treat shading on a painting. This reaction is so strong that a map produced with terrain illuminated from the south will appear inverted, with the hills looking like valleys and the valleys like hills.

**FIGURE 3.73.** The Johnstown terrain map produced with computer-generated plastic shading (using ArcInfo).

In an effort to represent terrain aspect information clearly while also creating effective relief shading, Moellering and Kimerling (1990) developed a unique color-rendering process that has subsequently been labeled MKS-ASPECT™ (Moellering, 1993). They started with the assumption that aspect is a nominal (qualitative) phenomenon for which color hue differences provide a suitable representation.[20] They set out to devise a color-matching system that would allow observers to visually separate terrain regions with different aspects while also providing appropriate depth cues leading to interpretation as a three-dimensional surface. The system relies heavily on OPT (described above in the discussion of eye and brain). As noted above, OPT predicts four unique hues from which all others are derived. It also predicts that certain hue combinations are not possible: those across the diagonals of the square color space (red–green or blue–yellow). The four unique hues are considered to be the maximally discriminable hues (when at maximum saturation and medium lightness). One guideline that Moellering and Kimerling arrive at from OPT is that aspect should be grouped into four, eight, sixteen, and so on, classes using the four unique hues or these four plus their first order combinations, second order combinations, and the like. They argue that the resulting hues (for eight or more classes) should be seen as a circular progression of related colors.

Moellering and Kimerling (1990) had the primary goal of depicting aspect classes clearly. Initially they matched the four unique hues with cardinal directions. Although a discriminable map was obtained, the resulting representation prompted a number of inversions of features (e.g., ridges seen as valleys). Their technique (unlike true relief shading) does not take into account a light source or reflectance due to that light source. The impression of relief obtained is due entirely to slope aspect. Moellering and Kimerling were able to achieve a reasonable impression of depth by rotating the unique colors so that yellow (the highest value color) was aligned with the standard light source azimuth (315° or northwest), and the value of all other hues was adjusted to match the deviation of azimuth from northwest.[21] It is claimed that the MKS-ASPECT™ system eliminates one of the most severe problems with standard gray tone relief shading: that the visual interpretation of the scene will be highly dependent upon the exact angle of illumination for the hypothetical light source (Moellering, 1993). By not relying on color value alone, identification of ridge lines or valleys is not as dependent upon how their alignment matches with that of the illumination. No empirical test of Moellering's claims has yet been undertaken.

Recentely Brewer (1993) has developed an alternative color scheme for mapping slope and aspect in conjunction. This scheme uses a hue range to represent aspect, with yellow as the anchor hue aligned with northwest. Other hues were selected so that a value progression was achieved in each direction from yellow, and each of the eight distinct aspect categories would have a sufficient saturation range for three saturation steps (plus unsaturated gray) to be discernable. Slope categories were depicted with these saturation steps; the higher the saturation, the steeper the slope. Its main advantage over Moellering and Kimerling's MKS-ASPECT™ system is that Brewer's color scheme results in a much more effective depiction of the terrain form, while still providing easily interpreted aspect information and adding three categories of slope.

An alternative use of color hue as a depth cue in terrain representation is found in Eyton's (1990) application of color chromostereopsis. As noted in the introduction to Part I, the idea of chromostereopsis (or more commonly advance-and-retreat) can be traced cartographically at least to Karl Peucker in 1898. The theory seems to have found at least partial support in research spanning the intervening decades (e.g., Eyton cites German publications on the topic as early as 1868 as well as Luckiesh, 1918; Kishto, 1965; etc.). A variety of explanations for the processes involved have been offered (see discussion of judging order above). It is uncertain, however, whether any standard layer tinting used on maps actually produces the effect (because there seem to have been no empirical cartographic tests). One problem, identified by Eyton (1990), in applying chromostereopsis to most paper maps is that the halftone processes of four-color lithographic printing will interfere with the effect. This interference is due to the fact that color appearance on lithographically printed maps results from the combination of overprinted inks and visual combination of adjacent dots.

Eyton (1990) experimented with several methods of producing the chromostereopic effect. He achieved limited success when a set of spectrally ordered colors were used on a layer tint map in which contours were created by adjacency of different colors. When he added black contours, the result (viewed as a color transparency) was said to have a "quite apparent" effect. Only an informal evaluation is offered, however, with 21 of 23 students in a cartography class claiming to see the effect. A map with black contours at double the contour interval was found to produce a weaker effect, leading Eyton to conclude that contour interval controlled the degree of depth seen. To explain the impact of the black contours, Eyton (1990, p. 23) argues that "the contour lines helped to create a rounding of the terrain form. Without contour lines the colors floated in planes; with the contour lines the display took on the appearance of a plastic surface with smooth, rounded features." An even more dramatic effect is cited for a continuous tone version of the map (in place of layer tinting). On this map, contour intervals of 200, 100, and 50 feet were compared, with the 50-foot interval producing the most dramatic effect. Eyton suggests that an explanation for the impact of contour lines on the perception of depth might be found in the fact that contour spacing is a cue to steepness of slope. This contour-enhancing effect remains to be empirically evaluated.

According to Eyton, the main problems involved in successful printing of maps using the chromostereopic effect are that standard printing changes the relative brightness of various hues (and brightness or color value interacts with the effect) and printed colors (particularly when inks are overprint or dithered) lack spectral purity. A solution proposed is to use fluorescent inks at full saturation with no overprinting. Fluorescent inks give the appearance of reflecting more light than is incident on the page. Again, Eyton provides anecdotal evidence, indicating that few students saw a fluorescent ink map as having depth, but the same map with black contours appeared to be three-dimensional to almost all the students. A final variation on the maps was obtained by adding hill shading (in gray) to the fluorescent ink maps. The added cue seemed to aid the perception of depth, but again the effect was strongest when contours were included as well.

A final depth-cue technique for static plan view maps worth noting is the "Tanaka method." Tanaka (1932) made use of shadow rather than shading (or color) to produce a sensation of depth in contour maps. The basis of the technique is to treat contours as if they represent a three-dimensional "layer-cake" model of terrain (which in a perspective view would result in a layered contour depiction of the sort discussed in Crawford and Marks, 1973). Tanaka's technique simulated the appearance of a layered contour map by putting white and black contours on a gray background. Contours toward the light source are in white with those away from the light source in black (Figure 3.74). Width of contours "varies

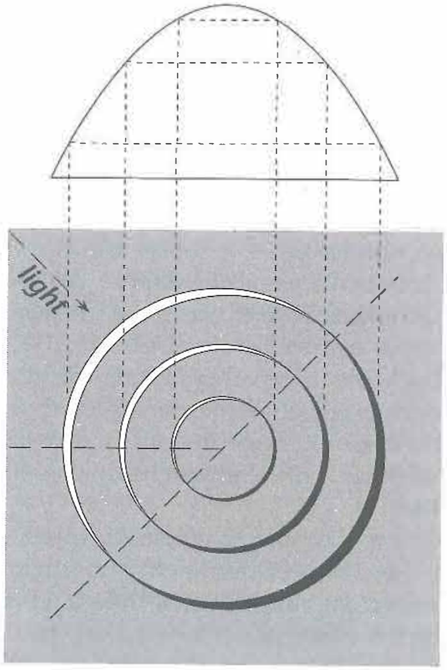

**FIGURE 3.74.** Representation of Tanaka's layer-cake method of terrain depiction. *After Japan Cartographers Association (1980, Fig. 4, p. 162).*

with the cosine of the angle θ between the horizontal direction of the incident ray and the normal to the contour at the point under consideration" (Japan Cartographers Association, 1980, pp. 162–163). Although the ability of Tanaka's method to provide a 3-D appearance is clear from examining a color version of a map using the Tanaka method (Japan Cartographers Association, 1980, f. 162), no empirical evaluation of the method exists nor any empirically derived guidelines on appropriate maximum widths for the variable contours.

## SUMMARY

The goal of this chapter has been to provide an overview of a range of issues relevant to visual processing of maps. Perspectives from neurophysiology, psychology, cognitive science, human factors engineering, and cartography are woven together in an effort to build an understanding of how maps are seen that can serve as a framework for research on and guidelines for map symbolization and design. It is only by understanding what vision is for and its limits that we can hope to comprehend the complex process involved in "seeing" a map.

Vision has been treated as a complex information-processing system that generates a succession of "representations." At the lowest level are representations of the visual scene on the retina of the eye. These are processed by our neurological hardware through a series of stages leading toward an organization of input into a coherent description of the visual scene in a form that can be interrogated by higher level cognitive processes.

After briefly reviewing the neurophysiological hardware issues and the limits that they place on map displays, the bulk of the chapter emphasized perceptual organization and perceptual categorization and judgment—two areas that have received considerable attention in the cartographic literature of the past four decades. In terms of perceptual organization of map information, the topics of perceptual grouping, attention, visual search, and figure–ground are emphasized. Psychological, cartographic, and other research on these topics is considered in relation to Bertin's contentions about the fundamental graphic variables available for creating map symbols. Although a great deal of cartographic research of the 1960s and 1970s dealt with issues of magnitude estimation, it is now clear that other aspects of perception are more relevant to map design. In the section on perceptual categorization and judgment, the emphasis, therefore, was placed on what we know about discrimination of symbols and patterns and about the propensity of the visual system to distinguish differences in kind and differences in order, two topics that seem particularly relevant to design of interactive visualization tools. Finally, the chapter concludes with a section devoted to the simulation of three dimensions on flat two-dimensional maps. This is an area of cartography with a long history, but one in which the integration of psychological principles and cartographic practice has been minimally addressed. Again, the topic of simulating the third dimension has become more important than ever in the context of visualization.

The perceptual emphasis in this chapter sets the stage for discussion of how knowledge is linked to perceptual input in the interpretation and use of maps. We pick up this thread in the next chapter, where the emphasis is on cognitive processes of mental categorization and spatial knowledge representation. These topics are critical to design of interactive visualization tools intended to facilitate visual thinking as well as to the formalization of spatial knowledge that will be required by expert systems for map generalization, symbolization, and design.

## NOTES

1. A fixation is a brief focus on a small section of the visual field where the item at the center of the fixation is "seen" by the foveal area of the retina and the cells connected to it.

2. The sections below are ordered because language requires order. The processes, however, are interdependent and are as likely to occur simultaneously as in the order presented or any other order. Some detection differences, for example, are necessary for grouping and attention. On the other hand, some grouping is required for objects to be isolated for discrimination and some attention to particular locations is needed to note differences between these locations.

3. Although Slocum did not specifically mention Gestalt psychology, he provides standard examples of the laws of similarity, proximity, and good continuation and directs the reader's attention to Arnheim (1974) and Woodworth (1938) as sources.

4. Eastman never cites the graphic organization variables he considered as Gestalt principles. To put his study in the context of the present discussion, however, I have made this link to Gestalt theory explicit in my review of Eastman's research.

5. This finding supports OPT, which would predict that wavelength averaging can happen for red–yellow and green–blue but cannot happen for red–green and blue–yellow mixtures. These mixtures are not possible (according to OPT) because there is no neurological mechanism for mixing across the diagonals of opponent color space.

6. The term "divided attention" is potentially misleading. Pomerantz and Schwaitzberg (1975) use it as an antonym for selective attention. Attention is not really divided (in time) between parts, as the term might imply, but is directed to the whole created by relationship of the parts. Thus "holistic" attention might be a more appropriate term.

7. Each of these hypotheses has implications for the development of visual-