Part of the information-processing approach to vision and visual cognition is based on the neurophysiological advances of the past decade. Researchers in these areas are gaining confidence concerning their eventual ability to explain how the brain works because of their success in explaining how individual and groups of cells respond to visual and other input. Some of the most recent computational theories of vision, including the widely cited theoretical work of the late David Marr (1982), has underpinnings in neurophysiology.

A second issue that seems to be regaining prominence in psychological explanations of human vision and visual cognition is the work of Gestalt psychologists. The highly reductionistic approach to explaining visual processes that assumed that all recognition of objects in visual scenes (whether symbols on maps or automobiles on a busy highway) begins with component features simply cannot provide the entire answer. Principles of perceptual organization must also be at work.

A third issue at the core of attempts to understand information processing at higher (i.e., cognitive) levels is the issue of knowledge structures. Two particular aspects of knowledge structures seem relevant to cartographic representation: mental categorization and knowledge schemata. Recent views of mental categorization diverge from the long-held belief that mental categories are relatively well-defined nonoverlapping units to a realization that categories are potentially overlapping, have fuzzy bounds, and are best modeled as groupings of related items around prototypes to which category members have varying degrees of similarity. These "categories" represent elements or nodes in knowledge schemata that define links and relations among these nodes.

Part I sketches the current information-processing approach to human vision and visual cognition, and pays particular attention to what it tells us concerning our potential to process information about the primary visual variables used on maps: location, size, value, hue, and the like. Emphasis is placed on linking map-reading tasks to what is known about how the eye–brain system works, on the role of Gestalt principles and other aspects of perceptual organization in map-reading tasks, and on the importance of mental categories and knowledge schemata to derivation of meaning from maps. Fundamental issues of an information-processing approach to visual cognition are presented in Chapter 2, and details relevant to maps and specific to different levels of processing are elaborated in the remaining chapters of this section.

CHAPTER TWO

An Information-Processing View of Vision and Visual Cognition

CARTOGRAPHIC IMPLICATIONS

To produce functional maps, we need to know something about what our visual–cognitive system is designed to do and what it is not designed to do, about the process by which vision and cognition allow us to derive meaning from visual scenes, and about the representations that are created at various stages of the process. The goal here is not to provide a comprehensive review of what is known about eye–brain functions and visual cognition nor about the cartographic research that has been linked to this knowledge (see Bruce and Green, 1990, and Pinker, 1984, for psychological overviews and Eastman, 1985a, and Medyckyj-Scott and Board, 1991, for cartographic perspectives). Rather, I will sketch the outline of a broad theoretical perspective on human processing of visual scenes and discuss how this theoretical approach can be applied to the process of extracting meaning from maps. This sketch is a selective one in which I draw on a limited set of current conceptions that I see as particularly applicable to the multilevel, multiperspective approach to cartographic representation advocated in this book.

Past efforts by cartographers to consider the implications of research in vision, visual perception, and spatial cognition for map symbolization and design have focused on details of specific low-level task abilities (e.g., discrimination, identification of shape, rank order, extracting figure from ground, etc.) without much concern for what vision and cognition are for, why vision and cognition work in the way they do, or the complex tasks to which maps are actually put. We have not been alone in taking this limiting approach. Much of psychology has pursued this path, and we have followed them.

Experimental psychologists have typically employed a reductionistic approach focused almost entirely on controlled, two-dimensional, laboratory test stimuli that are far removed from real-world experiences. These psychological studies have been quite influential in cartography, perhaps because they dealt with abstract symbols that seemed similar to those on maps. Recently, psychologists and cognitive scientists trying to understand vision and visual cognition have started to extend their focus from human processing of abstract stimuli to issues of how we see (and comprehend) the real world.

Gibson (1979), arguably, led the way in this regard with his *ecological* vision. His view, that the stimuli in the world "afford" meaning to what he considers a visually active but cognitively passive receptor, however, is not accepted by those who consider vision to be an information-processing system (a position that I adopt here). Gibson, in fact, argued strenuously against the idea that vision "processes" information. His conception was that vision "reacts to" information.

Contrary to Gibson's view, neurophysiological, neuropsychological, and psychophysical evidence provides strong support for the contention that vision is modular and involves a series of linked processes (see Chapter 3). In spite of the apparent incompatibility between Gibson's ecological optics and information-processing approaches, information-processing researchers (most notably Marr) have borrowed at least one of Gibson's principles: that to understand vision we need to consider what it is for in the real world. Following from this premise, Marr (1982) contended that several levels of theory development are needed and that computational theories may be most fundamental because they take into account what a system should do, and why, before trying to isolate how or with what mechanism.

There are two reasons why cartography should pay closer attention to research in psychology and cognitive science that is directed to perception and cognition of the real three-dimensional dynamic environment. The practical reason is that cartography is changing. Our "maps" have the potential to become more realistic by simulating 3-D and changing in real time. Eventually they may become embedded in virtual realities that allow the viewer to interact with them, much as the viewer interacts with real objects in the world. Perhaps more importantly, for the near term, attention to what vision and cognition are for in everyday experience may

be critical to understanding how vision and cognition work in any context. It is only logical to assume that fundamental structures have gradually evolved for recognition and identification of real-world objects and patterns. The evolutionarily recent development of abstract visual tools such as maps and graphics make it unlikely that special visual processes have evolved that allow us to read them. Understanding representations and processes used to grapple with the real world, then, is likely to take us farther toward understanding how vision and cognition react to stimuli that are as unnatural as maps than has trying to understand these stimuli in isolation.

MARR'S APPROACH TO VISION

David Marr's (1982) approach to vision has had a dramatic impact on understanding both vision and information-processing systems at various levels of analysis. This impact is due to Marr's clear delineation of the levels at which an information-processing task must be addressed if we are to understand it completely. He distinguished between three levels of understanding: the level of computational theory (at which we describe what a process must do and why, along with a logical strategy by which the process might be carried out), the level of representation and algorithms (dealing with how the theory might be implemented), and the level of processing device or hardware implementation (that considers how a particular representation might be implemented in the available device). Marr contends that some observable phenomena may find explanation at only one level, and thus that it is critical to consider the appropriate level of analysis whenever we evaluate evidence about how different processes function. Some failures of map symbols, for example, might be due to limits at the hardware level of neurophysiology, while others might be due to the representation forms or algorithms our brain applies to extracting meaning from a visual scene. Marr claims that the computational theory level is the most fundamental. If we recognize that, logically, our nervous system evolved to-meet certain needs, rather than that our perceptual processes evolved to make use of fixed predetermined neurological hardware, it becomes clear that understanding what vision is for is more important than understanding the neurophysiological mechanism by which it works. As a modifier to this de-emphasis of the hardware level, we should remember that the representations and algorithms used to implement the operations posited by computational theory evolved to meet survival needs of behavior in a complex three-dimensional world and, as a result, may not transfer well to the task of interpreting an abstract two-dimensional display. Hardware, in the form of our eye–brain system, can set limits for tasks (such as map reading) that vision has not evolved to accomplish.

Marr (1985, p. 103) defined vision as "the *process* of discovering from images what is present in the world, and where it is." Marr contends that to understand any information-processing system requires us not only to learn about the process, but to consider how information is represented. With vision it is, after all, a representation of the world formed on the retina that must be processed; and if Marr is correct, this retinal representation is transformed in a series of subsequent representations that lead from the two-dimensional retinal representation to a three-dimensional object-centered representation of the structure and organization of the viewed object or scene. From a cartographic perspective, Olson (1979) suggested a similar need to consider not only processes but "that which is processed" when she advocated giving attention to mental categories, mental organization, and mental constructs as things that change as a result of using a map.

Marr draws upon representational theories of the mind which posit that the mind has access to internal representations, that mental states are defined by what the internal representations specify, and that mental processes act on these representations. In the context of information processing, Marr (1985, p. 111) defines a representation as "a formal system for making explicit certain entities or types of information, together with a specification of how the system does this." As such, Marr's approach to representations used by the human visual system has much in common with the formal (i.e., semiotic) perspective on representation that I describe in Part II. As is true when a cartographer or other information designer selects a symbolic representation system to depict a specific set of data, the particular representations used by various stages of vision will "make certain information explicit at the expense of information that is pushed into the background and may be quite hard to recover" (Marr, 1985, p. 112). In discussing mental images of maps as one higher level representation form, for example, Peterson (1987, p. 38) points out that "mental images seem not to be copies of sensory impressions like 'pictures in the head' but rather they are intellectually processed and generalized representations, much like maps."[1]

Following from these philosophical perspectives, Marr (1982) developed a representational framework for vision in which he proposed three linked processing modules that transform an unstructured two-dimensional representation of the visual scene into a highly structured three-dimensional model representation (Figure 2.1). The nature of this framework was stimulated by evidence that mental representations of object shape are stored in a different place than representations of use and purpose, and that humans can grasp the shapes of things independently of

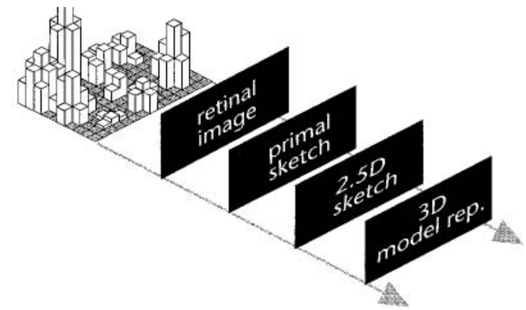

FIGURE 2.1. Marr's stages of vision. *Derived from Marr (1982).*

knowing what the things are named or what they are for (Marr, 1985). This led to a realization of what Marr (1985, p. 123) calls the "quintessential fact of human vision—that it tells about shape and space and spatial arrangement." This led, at the computational-theory level, to delineation of the primary purpose of vision as "building a description of the shapes and positions of things from images" (Marr, 1985, p. 123).

The retinal image provides the input for Marr's three-stage representational framework. This image consists of intensity values at points and has the limited purpose of representing these intensities. The first processing stage transforms the image information into what Marr and Nishihara (1978) termed the *primal sketch.* The primal sketch makes information about the retinal image explicit, particularly intensity changes across the image surface along with the geometry and organization of these changes. Marr contended that computation of *zero-crossings* (changes from light to dark) at various resolutions serve as the main input to the process of extracting a primal sketch from the visual scene (Figure 2.2). Primitive elements of the primal sketch are postulated to include perceptual units such as "blobs" and "edge segments." The primal sketch is envisioned as an array of cells that contain "symbols" indicating the presence of edges, bars, blobs, and so on, and their orientations—primarily features that are invariant over changes in overall illumination, contrast, and focus (Pinker, 1984) (Figure 2.3). As Pinker (1984) points out in a summary of Marr's theory, a crucial assumption (for theories of subsequent processing) is that the features symbolized in the primal sketch are extracted separately for various scales. This allows major features to be distinguished from details and leads to a hierarchical model for storage of shape categories in memory against which information from visual scenes is compared.

The next higher level of processing produces the *2½-D sketch,* a "representation of properties of the visible surfaces in a viewer-centered

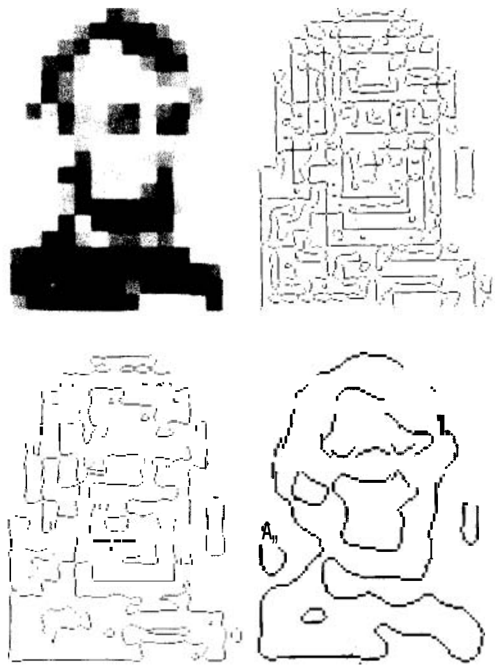

FIGURE 2.2. Zero-crossings from three spatial channels. It is the larger channel depiction (an image similar to one we obtain when squinting the eyes) that allows us to recognize Lincoln in L. D. Hammon's quantized image of Abraham Lincoln. *Reproduced from Marr* (1982, *Fig.* 2.23, *p. 74*). *From* Vision *by Marr. Copyright* 1982 *by W. H. Freeman and Company. Used with permission.*

coordinate system, such as surface orientation, distance from the viewer, and discontinuities in these qualities; surface reflectance and some coarse description of the prevailing illumination" (Figure 2.4) (Marr, 1985, p. 125). The 2½-D sketch is also conceived of as an array of cells with symbols indicating the various viewer-centered properties. Marr (1982) claims that this depiction has no input from top-down processing nor any global information about shape, only depths and orientations of local pieces of surface. The claim is for a completely precognitive process to this point (although Marr does not discount the possible role of top-down processes for directing attention to particular places in the visual field or for dealing with particularly ambiguous situations).

Finally, processing achieves what Marr (1985) terms the *3-D model representation* (Figure 2.5). This representation is posited to be an object-centered rather than a viewer-centered depiction of the three-dimensional structure and organization of the viewed shape, along with some level

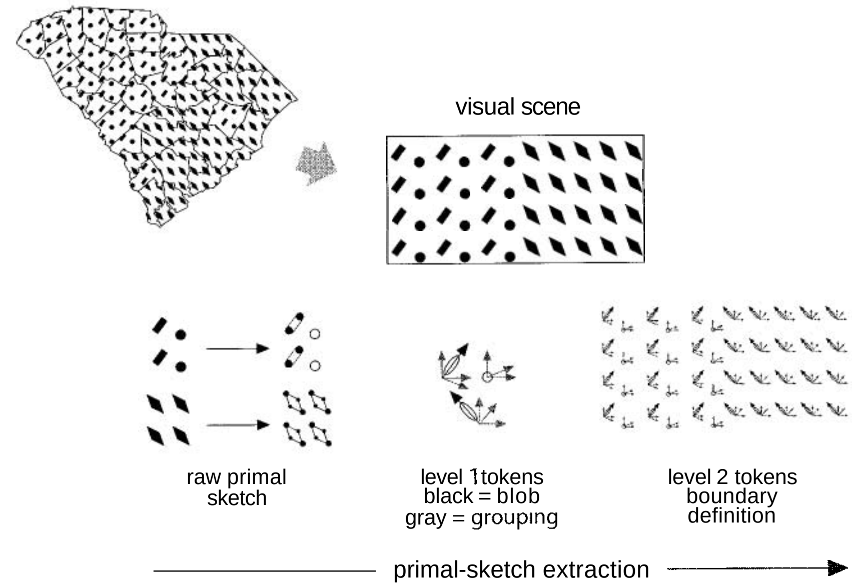

FIGURE 2.3. A depiction of the process of primal-sketch extraction from a display created by two area fills on a map. The raw primal sketch translates map marks to intensity changes in which edges and "blobs" (identified at various scales from zero-crossings) are isolated with their terminations indicated. These features are further processed to produce *place tokens* that are then combined through various perceptual grouping operations. One outcome of this grouping is the identification of boundaries between relatively homogeneous regions. *Derived from Marr* (1982, *Fig. 2.7, p. 53*).

of description of its surface properties. The 3-D model representation is considered to be a hierarchical model with each level consisting of a few axes to which volumetric shape primitives are assigned. The hierarchical model of shape descriptions separates information about arrangement of parts of a given size from the internal structure of those parts and information about shape of an individual from shape of general object classes (Pinker, 1984).[2] This modular system allows us to recognize objects (e.g., a bird or a building) as being in a particular class even if we can not recognize the individual (e.g., a duck or a church).

The modularity of processing suggested by Marr has interesting implications for the interpretation of map symbols. For example, when we are visually scanning a National Park Service map looking for a camping site, this modular processing may allow us to sort out the set of winter sports symbols from the set of camping symbols before specific recognition of any symbol takes place. To locate a campsite symbol, we do not

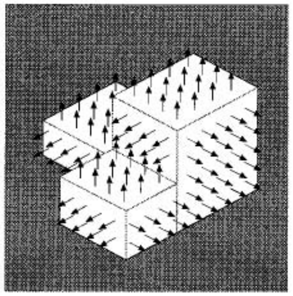

FIGURE 2.4. A depiction of Marr's concept of a 2%-D sketch from a section of the vrism mav similar to that in Figure 2.1. Emphasis is on orientation of surfaces (shown as dashed lines between object surfaces and solid lines between object and background). Derived from Marr (1982, Fig. 3.12, p. 129).

have to understand that a particular symbol means "ski bobbing" to rule it out as a candidate for a closer examination. Ratajski's (1971) system of standardized signs for economic maps seems to fit well with the hypothesis that visual processing is hierarchically organized (although he designed his map symbol system more than a decade before Marr presented the hypothesis) (Figure 2.6).

Marr's approach to shape derivation from visual scenes has had considerable impact on the field of machine vision where it has proved to be

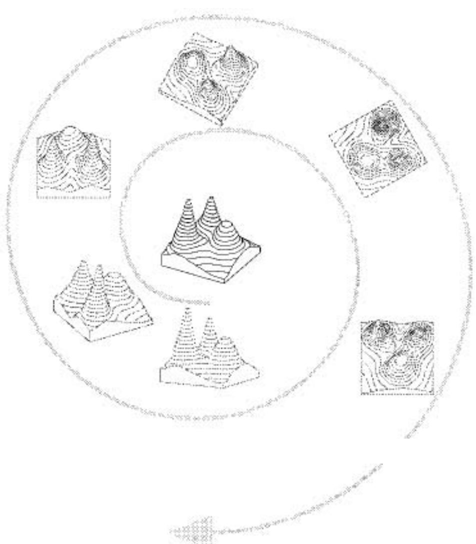

FIGURE 2.5. My concept of Marr's (1982) 3-D model representation. The representation is a complex structural model from which multiple viewpoints can be constructed. The figure symbolizes a mental "flyby" of the 3-D object representation. Not illustrated is Marr's conception of multiple hierachically linked structures that deal with the 3-D representation in parallel at various scales.

more successful than past iterative empirical approaches that tried to engineer a solution with no underlying theory upon which to base that solution. It also has proved superior to highly reductionistic approaches that searched for general principles by severely restricting the scope of problems (e.g., to a world of white toy blocks). Marr achieved a greater measure of success by starting at the level of computational theory. At this level he had to begin by deciding what vision was for before trying to determine how it worked. Marr's impact is due to the interaction of his three-tiered theoretical approach (of computational theory, representation–algorithm, and hardware implementation) to treatment of vision as an information-processing problem, and emphasis on the forms of representation acted upon at various stages of that process.

VISUAL COGNITION

Processing of Visual Stimuli

Visual cognition encompasses issues of how cognitive processes interact with vision to enable us to interpret the world and our apparent ability to mentally manipulate visual information in the form of images. To consider visual cognition and its potential implications for cartographic representation, we look to Steven Pinker. Pinker (1984) suggests that one important implication of Marr's work was that it convincingly illustrated different processes at work and different representations worked on at various stages of vision and visual cognition. As a result, there is a need for theories to distinguish processes of early vision (that are probably dominated by bottom-up processing) from those at higher levels of processing (that use output from early vision in combination with existing knowledge or knowledge structures). That Marr's 3-D model representation is modular and hierarchical is a key feature that allows us to link what is known about early (precognitive) visual processes to cognitive processes for recognition of shape (as Marr attempted) and to higher level visual problems such as graph comprehension (as Pinker went on to do). Marr's overall approach, with its focus on extracting shape and pattern from visual scenes in which there are no predetermined symbol–referent relationships, seems particularly applicable to the image end of the image-graphic continuum that I will describe in Part II. Pinker (1990) has extended Marr's basic structure to the other extreme (i.e., graphics) in developing a theoretical approach to graph comprehension.

Pinker's (1990) theory of graph comprehension addresses several fundamental issues that are also likely to underlay map comprehension. The theory will, therefore, be presented here in some detail as a basis

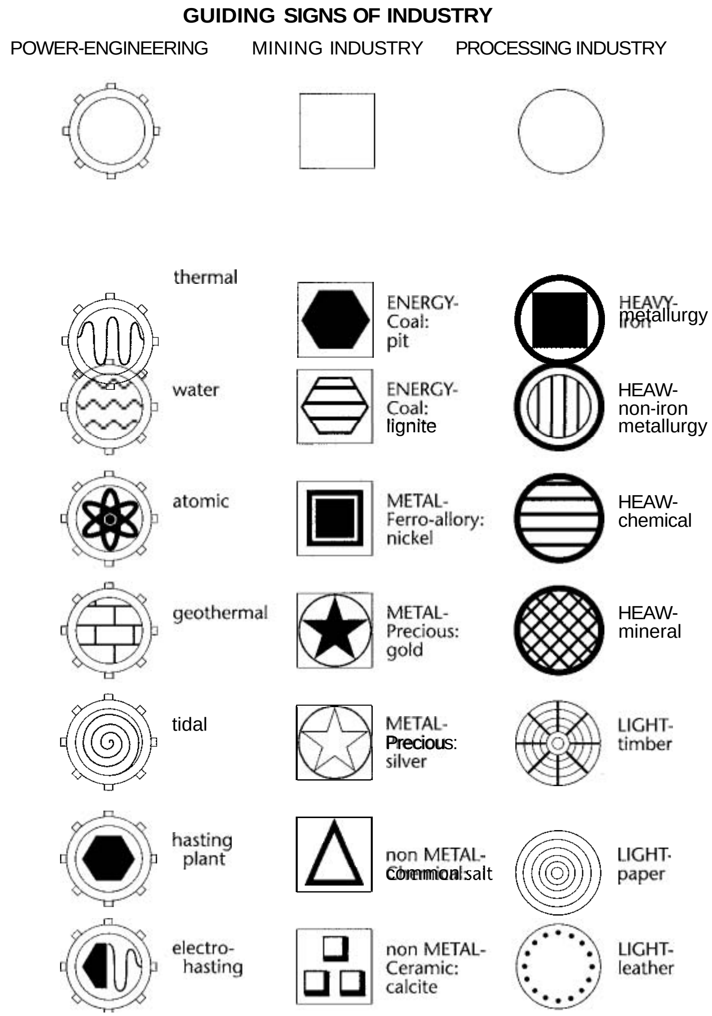

FIGURE 2.6. In preparing a standardized set of symbols for economic maps, Ratajski clearly relied on hierarchical grouping of visual features to assist readers in sorting without having to direct attention to (and waste effort on) the minor details that distinguish symbols in each group. These symbols represent only a sample of the 155 symbols that Ratajski designed for the three economic categories cited here. *Deriwed from Ratajski* (1971, *Figs.* 14–18 *and* 21, *pp.* 153–155 *and* 157).

from which to build a similar approach to map understanding. An interesting link to cartography is already included in Pinker's theory. This link is through his reliance on Bertin's (196711983) delineation of the tasks that the graph reader must accomplish if successful extraction of information from a graph is to occur. These tasks are (1) to identify the conceptual or real-world referents that the graph is conveying information about, (2) to identify the relevant dimensions of variation in the graphic variables and determine which visual dimensions correspond to which conceptual variable or scale, and (3) to use levels of each visual dimension to draw conclusions about particular levels of each conceptual scale. According to Pinker, these tasks imply that the reader must mentally represent the physical dimensions of marks on the graph. He calls this mental representation a *visual description.* In addition, the reader must draw upon what Pinker labels a *graph schema* to determine how physical dimensions are to be mapped onto mathematical scales.

Pinker (1990) relies on Marr's distinction between precognitive and cognitive aspects of vision to postulate that a *visual array* (equivalent to Marr's primal and/or 2%-D sketch or to iconic memory as described by Phillips, 1974) serves as the input to visual cognition which transforms this array into the *visual description* (analogous to short-term visual store [STVS] as described by Phillips, 1974, 1983, or the visuospatial scratch pad [VSSP] as proposed by Baddeley and Hitch, 1974).[3] This visual description is conceived of as a structural description of the information in the visual array. This structural description involves *variables* that "stand for" perceived entities or objects and *predicates* that specify relationships among, or attributes of, entities. Pinker develops a formal graphical notation system for depicting these visual descriptions, but for our purposes in this chapter the outline of his theory and its implications for map understanding can be appreciated without these details. I will return to the issue of visual descriptions and fill out the details in Chapter 4.

A key issue with Pinker's visual descriptions is that many different ones are possible as representations of a specific visual array. Part of the task of the theory, then, is to predict which visual description is most likely. Pinker (1990, p. 78) sets out to do this "based on knowledge of how the human visual system works." Specifically, Pinker details four factors intended to explain why certain visual descriptions and not others result from particular visual arrays.

The first factor, borrowed from Kubovy (1981), is that certain visual variables are considered *indispensable attributes,* attributes having a dominant perceptual status. These indispensable attributes are *space* (i.e., location) and *time.* Pinker ignores time because his theory is directed to static graphs. For dynamic maps and graphs, however, the fact that time has been demonstrated to be an indispensable attribute is critical. It tells us that change in positions or attributes over time should attract particular attention and serve as a perceptual organizer that is much stronger than hue, value, texture, shape, and so on. Agreeing with Kubovy, Pinker (1990) argues that spatial location will be the dominant factor (over the other visual variables—discounting time) in establishing *perceptual units* (or what the fundamental pieces of a visual scene are). In the example shown in Figure 2.7, there are clearly three basic units. If pattern variables dominated spatial location, however, there could be only two units (the left and middle bar vs. the right and left bar). In addition, Kubovy is cited as arguing that the indispensable attributes have stronger configural properties than other variables (Figure 2.8), are more discriminable, and are less subject to nonlinear functional relations between actual and perceived differences. Nonlinear functions between perceived and actual gray tones, for example, have been the focus of considerable cartographic attention. Kimerling (1985) has demonstrated that the relationship is increasingly curvilinear as the texture used to produce the gray fill in map areas decreases (Figure 2.9). Leonard and Buttenfield (1989) present results that suggest a complex interaction between texture and value when gray tones on maps are created by area fills coarse enough for texture to be noticed.

In addition to the above aspects of indispensable attributes, it is argued that attention is more *selective* to indispensable variables than to other visual variables. This means, for example, that on a map we are more able to attend to a particular location (regardless of symbol shape) than to a particular shape (regardless of location). Selective attention is related to Bertin's ideas about "selectivity" of graphic variables. Bertin (1967/1983, p. 67) seems to accept the primacy of location when he declares that "selective perception is utilized in obtaining an answer to the question: 'Where is a given category!'" Bertin's *associativity* principle, however, suggests that some other graphic variables have an ability to influence the primacy of location. Associativity, in his view, is the ability to visually group all variations on a particular (nonlocational) graphic variable. Shape is associative in this sense, while size is not (Figure 2.10). For the indispensable variable of location to take precedence, it may be necessary for other variation to be associative. Bertin does suggest that the best selection is achieved by relying only on location and 'juxtaposing separate images on the plane" (i.e., using something equivalent to Tufte's [1990] small multiples) (Figure 2.11).

Principles associated with indispensability of space and time "place

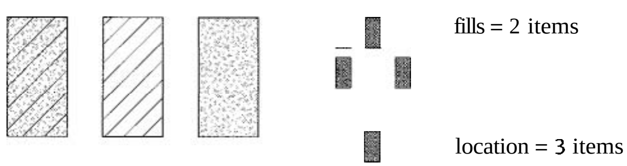

FIGURE 2.7. Location versus pattern as perceptual organizers. Readers will see three rather than two perceptual units. In fact, even after the two nonspatial units are identified (the pair with lines and the pair with random dots), it is difficult to "see" two units here. *Derived from Pinker (1990, Fig. 4.5a, p. 80).*

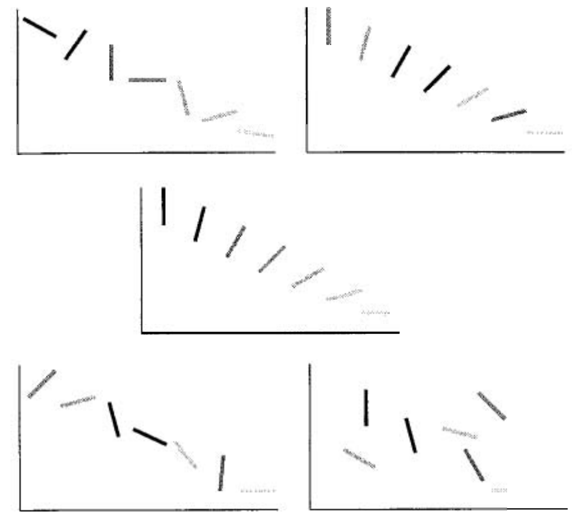

FIGURE 2.8. Configural properties of location, value, and orientation. If both location and value (upper left) or location and orientation (upper right) are ordered, the individual and joint relations are readily apparent. When all three are ordered, the individual and joint relationships are also easily extracted. It is, however, considerably easier to notice ordered location when *both* value and orientation are random (lower left) than ordered value together with ordered orientation when location is random (lower right).

constraints on the parts of an array that variables may stand for, on how numerical variables represent physical continua, and on how predicates are encoded or verified with respect to the visual array" (Pinker, 1990, p. 83). A second factor thought to govern "how atomic perceptual units will be integrated into a coherent percept" is the Gestalt laws of grouping (Pinker, 1990, p. 83). It seems clear that Gestalt laws of grouping play a

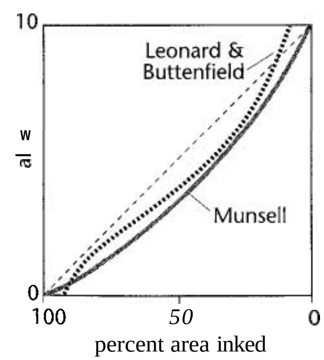

FIGURE 2.9. Kimerling's translation of the Munsell gray scale into percent area inked form (based on 133 lines/inch map output) allows comparison with Leonard and Buttenfield's gray scale for laser printer maps using approximately 50 lines/inch patterns). *Derived from Kimerling (1985, Fig. 4, p. 137) and Leonard and Buttenfield (1989, Table 2, p. 100).*

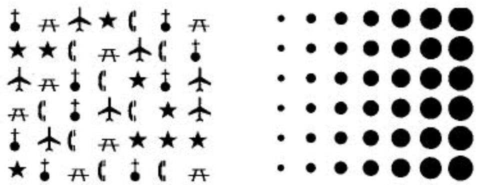

FIGURE 2.10. Shapes can be visually seen as a whole (left), but symbols of different size are seen primarily as different (right). *After Bertin* (196711983, *Figs. 3 and 8*, *p.* 65). *Adapted by permission of the University of Wisconsin Press.*

large part in sorting out groups and relationships in maps as well as graphics. McCleary (1981), for example, relies heavily on Gestalt principles in his approach to designing effective graphic and cartographic presentations. In spite of general agreement about the relevance of Gestalt principles to maps and graphics, however, Pinker identifies an issue that has made application of the principles problematic in both a theoretical and an applied context. The limiting factor is that little or no work has been done concerning the relative strengths of these principles (e.g., is it more important for map symbols to be proximate to one another, similar in color, arrayed in linear-order, etc. and which of these factors dominates if they conflict?).[4]

A third factor thought to influence which visual description is likely in response to a specific visual array is the mental representation of magnitude. Although magnitudes are probably represented in the visual array as continuous values, there is evidence that higher level representations tend to be encoded as discrete values on an ordinal scale with only about

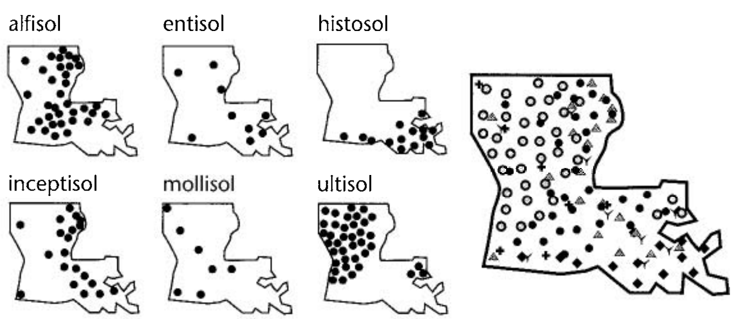

FIGURE 2.11. The selectivity of location is part of what makes Tufte's (1990) idea of small multiples work. With the small multiples shown here, a quick impression can be obtained of the correspondence (or lack of it) between pairs and groups of soils. A single map with various shaped point symbols, however, can be extremely confusing (even at 200% size). *Derived from maps of the entire United States by Gersmehl* (1977, *Figs.* 1, 2, 5, *and* 7–9, *pp.* 423 *and* 425–427).

seven steps. Pinker (1990) also notes that the likelihood of a value being encoded at all will depend on its context. Extreme values, for example, are more likely to be perceptually encoded.

A final factor related to the extraction of a visual description from the visual array is the coordinate systems used. Pinker (1990) borrows from Marr and Nishihara's (1978) theory to postulate that (1) polar or rectangular coordinates are usually used to represent shape and position, and that (2) different elements of the scene are represented in separate local coordinates that are centered on larger objects that are part of a Gestalt group.

Along with the issue of which particular visual description will be generated to represent a visual array, Pinker distinguishes between *default* and *elaborated* visual descriptions. Default visual descriptions are thought to result exclusively from bottom-up processes. These visual descriptions will be small due to constraints on short-term visual store, which allows between four and nine nodes (i.e., elements) to be kept active at one time. In addition, the contents of default visual descriptions will depend on default encoding likelihoods of predicates (i.e., relations). Although any form of relation that a person can conceive of can be brought to bear on the visual array, the default visual description will contain only those relations that are "just noticed" without conscious thought. The relationships that are just noticed will vary from person to person and are subject to frequency of use (or practice), which partially explains "expert" versus "novice" abilities at recognizing patterns. This accords with cartographic evidence that training can improve performance on what seem to be precognitive tasks (Olson, 1979; Castner, 1983) and with the fact that nonexperts are often trained to search for patterns on photographs generated in particle physics laboratories (Judson, 1987). As Schneider and Shiffrin (1977) point out, when someone repeatedly assigns particular visual patterns to specific categories, recognizing these patterns becomes automatic (i.e., preconscious). An elaborated visual description will start from the default description, but will make-use of top-down processes to add elements and predicates to that initial description.

Pinker's model of graph comprehension depends on what he calls a *graph schema* to mediate between knowledge representations (or memory) and the visual description derived. These graph schemata are structures for organizing information about variables and relationships in graphs. Graph schemata, according to Pinker, are responsible for (1) specifying how information in visual descriptions is translated into a *conceptual message,* and (2) how queries that involve interrogation of the visual description or that stimulate further interpretation of the visual array are translated into a process that creates or retrieves the required information. Pinker (1990, p. 95) specifically describes a graph schema as "a memory representation embodying knowledge in some domain, consisting of a description which contains 'slots' or parameters for as yet unknown information."

Pinker's graph schemata seem to follow the general idea of cognitive schemata suggested by Neisser (1976, p. 55) as "plans for finding out about objects and events, for obtaining more information to fill in the format." Pinker, however, differs with Neisser in suggesting that precognitive processes involved in creating the visual array-determine which schema will be applied, rather than considering schemata as only having a top-down function to direct "exploratory movements of the head and eyes." (Neisser, 1976, p. 55). Pinker's viewpoint seems to agree with that of Antes and Mann (1984) who suggest that the schema applied is often dependent upon how the information is presented to us and at what scale, and Eastman (1985b) who, in relation to maps, demonstrated that map design variables could influence grouping or chunking of information extracted from maps.

Graph schemata can specify what must be true of a particular graph type as well as how it might vary from exemplar to exemplar. Schemata as Pinker describes them can exist at different levels of detail, from those that allow us to distinguish between a graph and other kinds of information display to those that allow us to recognize a particular graph type and interpret the information presented within it (e.g., a line vs. a bar graph and the associated implications of a discrete vs. a continuous function).

Putting the concept of multiple levels of representation (i.e., visual arrays, default and elaborated visual descriptions, and cognitive graph schemata) together with a conception of the kinds of interaction possible between them, Pinker (1990) presents a graphic model of the process of the information-processing stages of "graph comprehension." The model (Figure 2.12) presents graph comprehension as a highly interactive process that makes use of both bottom-up precognitive processes and top-down cognitive processes to gradually build an overall understanding of the graph.

Pinker's model has some parallel's in the cartographic literature, but for the most part cartographic efforts in this direction have been less formalized and only limited attention has been given to information-processing models. This is perhaps because of the surface similarity to information theory models that served as the basis for the communication paradigm (and our disillusionment with that perspective). One of the first information-processing-like concepts suggested in relation to maps was Olson's (1976) hierarchical model of map-readmg tasks. This model divided tasks into those involving (1) comparison of individual symbol characteristics (e.g., shape or size), (2) assessment of symbol group characteristics (e.g., pattern), and (3) use of the map as a decision-making or

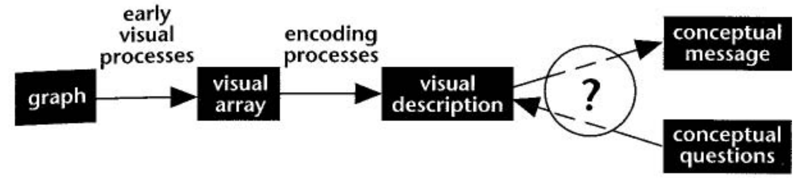

FIGURE 2.12. Pinker's model of visual information processing associated with graph comprehension. The circle with a question mark represents the iterative queries possible between the visual description and prior knowledge. *After Pinker (1990, Fig. 4.14, p. 94). Adapted by permission of Ablex Publishing Corporation.*

knowledge-building device. Olson's approach was largely intuitive and was not linked directly to information-processing theories. Although the model helped suggest some differences to anticipate concerning the relative impact of design research and user training at various task levels, the levels delineated do not match current theories of visual cognition.

At the precognitive level, Dobson (1979b) and Shortridge (1982) have offered information-processing approaches to low-level perception of map symbols. Dobson's goals were applied ones and his work evolved toward a call for a human factors approach to "map engineering." He argued that this approach should use an information-processing perspective as a base from which to design experiments geared toward solving specific map design problems. Shortridge, in contrast, pointed out some cartographically relevant issues that information-processing work in psychology uncovered (e.g., the issue of integral and separable dimensions of multidimensional symbols), but did not share Dobson's optimism for the promise of cartographic research linked directly to information-processing models.

Eastman (1985a), working at a more conceptual level, provides a clear sketch of the information-processing perspective and its potential applicability to user research in cartography. He also deals with the issue of linking a semiotic approach to map symbolization with an information-processing view of map reading. In his presentation of system and process models for map reading, Eastman is careful to point out the fundamental difference between an information theory approach to map communication that measures bits of information transmitted (with the goal being communication of the most bits) and attention to information processing (with the goal being to understand how people actively see and conceptualize about map information). For cartographers to facilitate visual and cognitive information processing, we must understand both the system that does the processing and the processes themselves.

As found in Pinker's model of graph comprehension, Eastman's (1985a, p. 97) discussion of information processing as it might be applied to cartography incorporates the concept of schemata as "a cognitive structure that can be broken down to produce the essential aspects of a typical example of a particular class of objects (called a 'prototype')." He goes on to characterize schemata as "networks of concepts or entities interconnected by a set of relations . . . within a hierarchical framework" (Eastman, 1985a, p. 97). Basing his view primarily on that of the psychologist Palmer (1975, 1977), Eastman contends that entities exist at two levels, the global and the local. Global properties are said to pertain to "chunks as a whole," while local properties are derived from component parts. Chunks and schemata are said to differ only in degree of elaboration, an argument that schemata are hierarchical. This matches to some extent with Pinker's conception of general and specific schemata with the difference being that Eastman allows schemata to exist for parts of a problem context while Pinker's schemata apply to the entire problem (with only the level of detail varying).

As Eastman notes, there is considerable evidence from the study of chess expertise that mental structures stored in memory can facilitate the organization and interpretation of a visual scene. In particular, research by Chase and Simon (1973) has demonstrated that experts organize information into larger chunks, thus enabling them to assess a particular arrangement of the chess board more quickly than novices—but only if that arrangement represents a likely stage in a chess game. Novices, who do not have a well-developed schema for possible arrangements, must process the visual scene at a more local level.

A major difference between Eastman (1985a) and Pinker (1990) is the weight given to top-down processing. Eastman favors the perspective of Navon (1977) who argues that top-down global processes act first to control more local processing of a visual scene. Eastman cites eye movement research in support of this view. Although eye movements when viewing a map have been found to be individualistic, orderly recurrent patterns are common for individuals (Eastman and Castner, 1983). Eastman (1985a) contends that this result can be interpreted as personalized schemata exerting control over how an individual sees a visual scene. He suggests a cyclical process in which a schema directs exploration, which samples the visual scene, and the results of this sampling have the potential to modify the schema (Figure 2.13). While some eye movement evidence does suggest that top-down processes may be dominant in steering attention across the visual field, other results can be interpreted as evidence for automatic precognitive processes that are reacting to the visual array without any conscious control. A recent analysis of differences in viewing strategy for a range of symbolization types seems to support this latter view (Morita, 1991) (Figure 2.14).

Whether or not the eye movement research provides evidence for

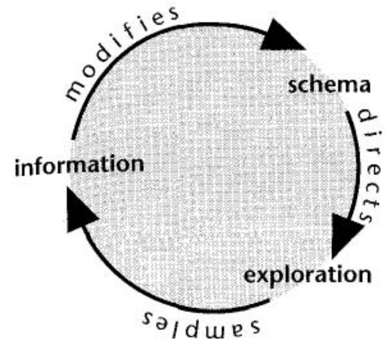

FIGURE 2.13. Eastman's interpretation of Neisser's concept of a schema applied to visual exploration of a map or other graphic display. *After Eastman (1985a, Fig. 3, p. 99). Adapted by permission of The* Cartographic Journal.

top-down processes as a control on attention, this evidence says nothing about the role of top-down processes in translating the visual array into Marr's 2½-D sketch or Pinker's visual description. As with most complex human processes, what is most likely is that both top-down cognitive processes and bottom-up precognitive visual processes complement each other, with each taking precedence some of the time. It may be that behavior in the environment requires more reliance on bottom-up processes as a first sort to ensure that we obtain some information fast enough to do us some good (e.g., when driving a car our eye movements are likely to be attracted to movement with no thought processes involved telling us that we should attend to movement more than to color). In the case of information graphics, however, the problem context for vision is considerably restricted and it is logical that our visual–cognitive processing system can take advantage of this to make better use of expectations in directing where we look or what features we attend to.

Eastman's conjunction of information-processing and semiotic approaches is complemented within the psychological literature by Kosslyn's (1989) efforts toward "understanding graphs and charts." He emphasizes a mix of visual information processing and semiotic principles in the development of acceptability guidelines for graphics. In Kosslyn's information-processing model, a series of factors are identified that influence processing at different stages (and might impair or facilitate graph interpretation). Between the "perceptual image" (Pinker's visual array and Marr's primal and 2½-D sketches) and short-term memory (Pinker's visual description), issues of discriminability (a hardware-level issue), distortion (a hardware and representation issue), organization (a representation issue), and priorities (a representation issue) come into play (Figure 2.15). Short-term memory itself will have capacity limitations and there will be issues of kind of encoding between short-term memory and long-term memory. Knowledge in long-term memory (accessed via Pinker's schemata) acts on the representation in short-term memory, either ac-

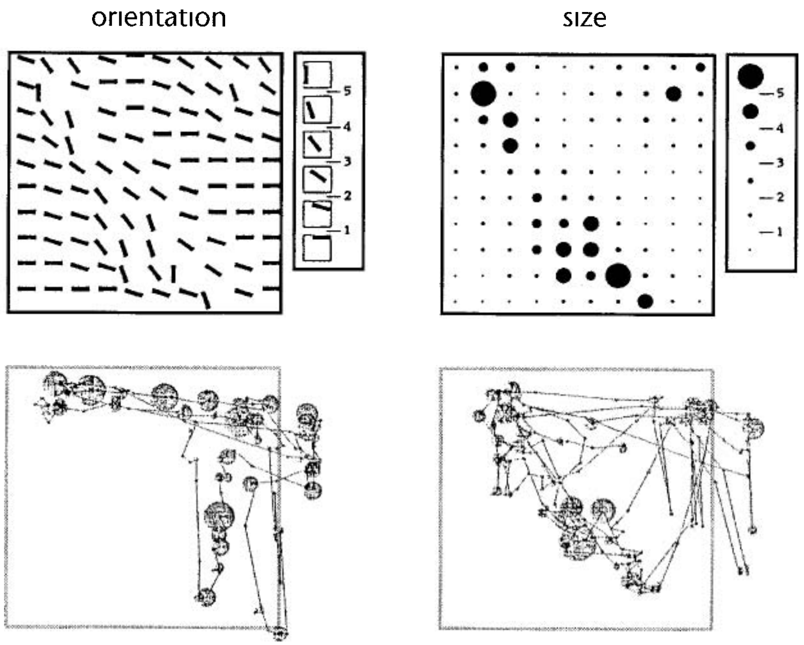

FIGURE 2.14. The influence of symbolization type on visual scanning of mapped patterns. The two schematic maps depict indentical data using two of Bertin's graphic variables: orientation and size. *Reproduced from Morita (1991, Figs. 3 and 4, pp. 4 and 8). Reprinted by permission of the author.*

cepting what is encountered or prompting reorganization or directed search of the visual scene (i.e., reference back to the visual array). This process will obviously be an iterative one.

Eastman's (1985a) view that we should consider maps as facilitators rather than as communicators of information is complementary with Pinker's (1990) view that displays leading to clear perceptual organization will be most effective. In both cases, it is anticipated that characteristics of the display will influence how easily existing schemata are brought to bear on the problem of deriving information through viewing the display. These perspectives fit with my own current ideas about geographic visualization (considered in detail in Part III). Geographic visualization (GVIS), however, with its high level of interaction between viewer and map display and its emphasis on searching for unknown patterns versus interpreting a predetermined message, demands emphasis on somewhat different information-processing tasks. In response to this need, John Ganter and I developed an information-processing model of map-based visualization as a pattern-matching process (MacEachren and Ganter, 1990) (Figure 2.16).

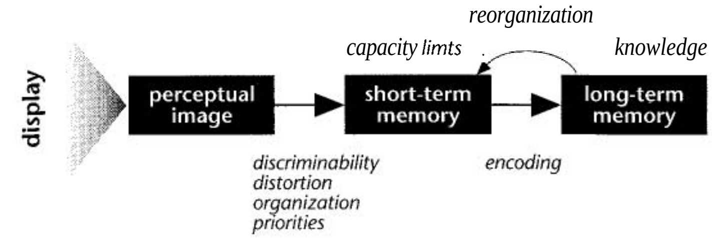

FIGURE 2.15. Kosslyn's information-processing model of graph perception. *After Kosslyn (1989, Fig. 2, p. 188). Adapted by permission of John Wiley & Sons, Ltd., from* Applied Cognitive Psychology.

At the time we developed this pattern-matching approach, we were only vaguely aware of Marr's information-processing model of human vision and we had not encountered Pinker's work on graph comprehension (which was published in the same year). Our model evolved from an overall concern for the impact of scientific visualization on cartography, study of the scientific creativity and scientific visualization literature, and from the perspective on history and philosophy of science presented by Howard Margolis (1987). The fact that two rather similar models evolved from quite different research perspectives indicates a strength for the basic tenet of both models—that *cognitive "schemata" exert some level of control on how we see evidence and that with repeated use, these schemata become ingrained to the point that noticing patterns of particular types becomes automatic, or precognitive.*

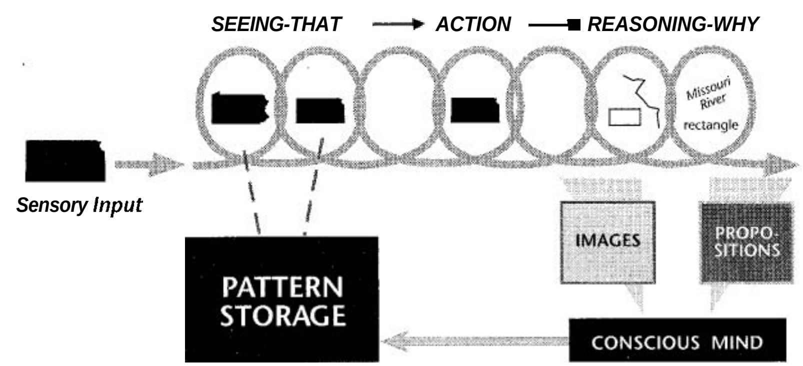

FIGURE 2.16. The pattern identification model of cartographic visualization. *After MacEachren and Ganter (1990, Fig. 2, p. 70). (The initial graphic depiction of this model was developed by John Ganter and appeared in Ganter and MacEachren, 1989.) Adapted from* Cartographica *by permission of University of Toronto Press, Inc. Copyright 1990 by University of Toronto Press, Inc.*

Processing of Imagery

One aspect of the Ganter–MacEachren model that differs from the discussion of graph comprehension by Pinker (1990) or the application of information-processing theory to cartography by Eastman (1985a) is the explicit attention to imagery. Pinker (1984), however, has given considerable attention to imagery in his previous work. His overview of visual cognition presented visual information processing and visual imagery as complementary parts of visual cognition. Imagery, in this context, is seen as a process for interacting with spatial aspects of long-term memory representations. Both Kosslyn (1980) and Pinker (1984) suggest that imagery may use some of the same processes as vision, and that it is analogous to visual processing of stimuli that are present. Finke (1980) suggests that perception uses one set of processes at a neurological level for feature analysis and grouping, another set at a higher level of analysis related to object shape, size, orientation, and the like, and further processes that apply general knowledge and cognitive skills. He goes on to contend that some imagery phenomena reflect the operation of similar (if not identical) midlevel as well as high-level processes. In Pinker's graph-comprehension model (although he does not make this explicit point himself) images may be something equivalent to a visual description and, as such, may be amenable to the same kind of message extraction and interrogation that Pinker postulates for visual descriptions derived from visual arrays.

Within cartography, the concept of mental images is intuitively appealing. All cartographers think visually. It is therefore difficult for a cartographer to conceive of imagery as being an epiphenomenon with no real use, as has been suggested by some anti-image protagonists (see Pinker, 1984, for discussion of the debate). Considerable evidence has been compiled that images of maps can be formed, stored in memory (in some probably nonanalog representation format), extracted from memory, and used in ways similar to a physical map (see, e.g., Peterson, 1985, 1987; Lloyd and Steinke, 1985; Lloyd, 1988; and Kosslyn, 1980). Some of this evidence is reviewed in Chapter 4 where the issue of how information obtained from maps might be mentally represented is considered in detail. For the present, we will accept the existence of images and consider where they fit in an information-processing account of vision and visual cognition.

Peterson (1987) suggests that both maps and images are products of spatial thinking and are dependent upon the process of arranging objects in space. In addition, he contends that mental images, like maps, are not copies of sensory impressions, but intellectually processed and generalized representations. This commonality between cartographic maps and spatial images led him to propose a GIs-like model of how image and propositional information might be drawn upon in spatial problem solving. Peterson suggests that as in a GIS (where data, equations, and mathematical models provide nonvisual information and maps, graphs, and remotely sensed scenes provide visual information) our mental information-processing system has access to both propositional and image information. When we are faced with an interpretation or decision, Peterson suggests that a competitive process (as hypothesized by dual coding theory) is used. The system accesses both kinds of information and the answer to the query will be made using whichever information source can be retrieved most quickly (Figure 2.17). One of Peterson's main points is that the structure or design of maps will influence the extent to which information obtained from them will lead to cognitive representations accessible as images (or, I would add, stimulate generation of images from existing knowledge). Since the kinds of questions and problems for which map information might be useful are spatial, it seems logical that images will be suited to dealing with them. As a result, Peterson argues that good map designs should be evaluated, in part, by the extent to which they prompt memory representations that are accessible as images.

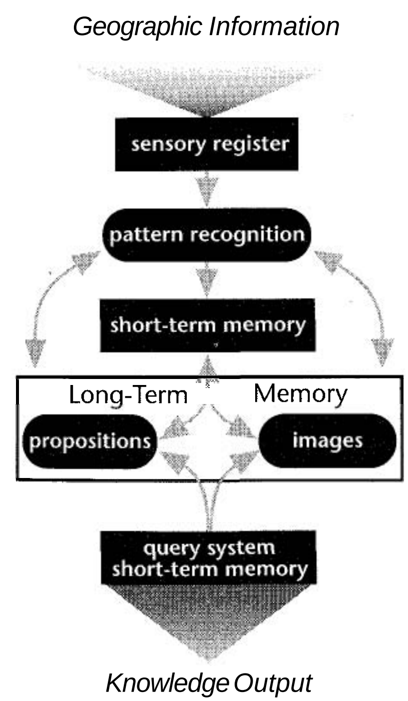

FIGURE 2.17. Peterson's human information system model. After Peterson (1987, Fig. 5, p. 40). Adapted by permission of The Cartographic Journal.

Pinker (1984) specifically considers the issue of what images are good for, if we agree that they do exist. He cites four possibilities. Images might serve as a global coordinate system to which information could be related. Similarly, we might use images as a way to compute spatial information (e.g., distances and directions) that we did not encode directly. A third possibility is that images are utilized in some object and pattern recognition tasks, particularly when we are anticipating one of a small number of possibilities. Finally, Pinker suggests that images might provide the best medium for solving certain abstract problems. He contends that images are "representations in a medium with certain fixed properties, and can be subjected to transformations such as rotation and scaling" (Pinker, 1984, p. 57). He suggests that abstract problems could be tackled by "translating their entities into imagined objects, transforming them using available image transformations, detecting the resulting spatial relations and properties, and translating those relations and properties back to the problem domain" (Pinker, 1984, p. 56).

This last potential use for images is exemplified by their purported role in scientific discovery. Visual imagery has been credited with playing a significant part in various scientific discoveries (Hadamard, 1945; Shepard, 1978). Faraday and Maxwell appear to have used imagery in their initial conceptions of electric and magnetic fields; Einstein in a variety of contexts emphasized his own tendency to use visual over verbal thinking; and Kekule's introspective accounts suggest that visual images led to his discovery of the structure of benzene (Shepard, 1978). The image-transformation processes Pinker suggests seem particularly applicable to theorizing in mathematics and physics. Scientific visualization researchers appear to have directed their attention in these fields to facilitating these transformations by making abstract images explicit.[5]

Peterson (1987) has not been alone in suggesting that concrete visual representations can serve to prompt the mental imagery needed for understanding problem situations and creative thinking about those situations (e.g., Beveridge and Parkins, 1987; Larkin and Simon, 1987; MacEachren and Ganter, 1990; MacEachren et al., 1992). Larkin and Simon (1987) emphasize the perceptual inferences that can be made from diagrams and suggest, by calling the information "zero cost," that these inferences are precognitive. Arnheim (1985) offers similar ideas under the heading of "intuition." To Arnheim, intuition is a component of visual thinking that operates like a "gift from nowhere." Both mental images and explicit representations appear to share the advantage of using Kubovy's indispensable variable of space. Both facilitate spatial grouping of information and inferences based upon spatial position. Arnheim's gift from nowhere and Larkin and Simons's zero-cost information, then, may result from the ease with which visual cognition (operating on visual descriptions of an actual scene or on visual images generated with a similar structure) deals with space.

CONCLUSION

The above discussion outlines a view of vision and visual cognition that is achieving growing acceptance in the cognitive psychology and cognitive science communities. Although not all researchers in these fields even agree about the fundamental issue of whether vision and/or cognition can be considered an information-processing system, this perspective is probably the dominant one and makes intuitive sense as one of several avenues to understanding maps as representational devices.

From a cartographic perspective, looking at vision and visual cognition as information-processing activities leads to a focus on a series of representation of what is seen on a map's surface. Representations of different kinds result from each stage of processing. At the lowest level, the representation that results is thought to be a visual array that contains "place tokens" or abstract "symbols" that stand for fundamental components (i.e., edges, objects, discontinuities, etc.) derived from the visual scene (i.e., the map display) and basic information about the components and their relationships (e.g., size, depth in the visual field, brightness, etc.). Higher level processing is judged to result in a visual description of the visual array from the previous stage. Knowledge schemata then mediate between this visual description and knowledge in long-term memory to interrogate the visual description, cause its modification, and eventually derive some level of meaning from the map display. The meaning derived can result in modification to existing knowledge schemata or creation of altogether new schemata, yet another representation of the map display. A key issue that must be emphasized here is that I am not hypothesizing an information-"transmission" system (with its concomitant emphasis on ratios of signal to noise) but a modular system in which information is "created" and re-created by a series of interpretive processes.

Applying these concepts to maps, I contend that the structure of visual descriptions derived from viewing maps will be based upon both general and specific map schemata (the latter resulting from expert knowledge or interpretation of legend information). A key factor in map schemata and legend understanding will be the basic human facility for categorizing the world. That this facility probably does not usually result in precisely delineated categories has significant implications for map understanding.

The two main levels of processing and their associated representations form the focus for the remaining two chapters of Part I. At issue will be the processes by which each new level of representation is generated, the implications of these processes for map symbolization and design, and the corresponding implications of symbolization and design decisions for the success at which reasonable cognitive representations are achieved.

NOTES

1. We may even find that the intense efforts of the past decade by cartographers to understand map generalization closely parallel some of the work by information-processing researchers to understand visual cognition. The cartographic principles developed may inform theories of mental image formation (at least in relation to maps and other abstract visual scenes) and work by cartographers and others on visual cognition may suggest some new approaches to those trying to develop a more unified theory of map generalization (rather than the fragmented element-by-element approaches characteristic of most work thus far).

2. A variety of theories have subsequently been proposed to explain how parts are categorized and identified; see, for example, Biederman (1987) and Hoffman and Richards (1984).

3. The VSSP is proposed as a temporary storage location in which (a shelf on which) visual information can be briefly stored until needed by the "central executive." The VSSP complements an "articulatory loop" where phonetic material can be stored (see Baddeley, 1988, for details).

4. For an overview of one cartographic attempt to evaluate the relative impact of specific Gestalt grouping variables, see Chapter 3.

5. See Part III for detailed discussion of computer-assisted visualization in a geographic context.

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