chapter two

Angles of Regard: Psychology Meets Technology in the Perception and Interpretation of Nonliteral Imagery

Robert R. Hoffman, Arthur Markman, and Walter H. Carnahan

Contents

Keywords: *expertise, thermography, perceptual learning, color, cartography, display design, visualization, diagrammatic reasoning, concept formation, instructional design, mental models, automated classification, visualization software*

2.1 Introduction

Advances in remote sensing are often focused on technological innovations in imaging equipment and in computational techniques that enable image data to be interpreted automatically.[83] In this chapter, we discuss the importance of the human interpreter in processing remote sensing data. To this end, we will draw on literature from both remote sensing and cognitive psychology. We will illustrate the nexus of these areas with a demonstration experiment examining novice / expert differences in the interpretation of aerial thermograms.

2.1.1 What is remote sensing?

The prototypical remote sensing situation involves a sensor taking an image of a target a great distance away. The image is often formed from some energy source outside the range of sensitivity of the human visual system. For example, a satellite in geostationary orbit might generate an infrared image of the northeast coast of the U.S. As another example, a radio telescope in the Arizona desert might generate radio-frequency images of distant space. Such images pose a number of interesting interpretation problems. First, the perspective of the image is different from what people typically see. Aerial images typically yield overhead views of a region, whereas people typically view the world from a terrestrial perspective. Second, the scale of the images is different from what people typically encounter. Third, there are many ways to map intensities of the sensed quantity (e.g., infrared radiation) onto a visible image. These choices will affect the (visible-light) perceptual properties of the image. Finally, through evolution, the human visual system has adapted to allow us to deal with the visible world. Comparable evolved mechanisms do not exist for interpretation of nonliteral images.

Remote sensing scientists are well aware of such problems, and as a result the field of remote sensing has become a unified discipline devoted to the development of tools to facilitate interpretation by humans and to create tools for automated interpretation of images (see [68,120,175]).

To be clear at the outset, while the typical case of remote sensing involves a situation like that described above, research on remote sensing also involves cases that differ from the prototype. For example:

- *The sensor does not have to be far from the object of study.*

A natural resource manager may be tracking the migrations of radio-tagged caribou, and in this case the object of study is in direct contact with the sensor, but the radio signals are tracked via a geostationary satellite that is 25,000 miles in space. The manager views the data overlaid on an aerial perspective map, as if looking down at the Earth.

- *The sensor does not have to be far from the observer.*

A meteorologist predicting the weather using a ground-based radar may be a mere ten meters from the sensor (the radar antenna), but the view is of the atmosphere scores of kilometers distant from the sensor. An astrophysicist studying quasars using a ground-based radio telescope may be tens of meters from the sensor, but the object viewed may be billions of light years away. An aircraft pilot using a light amplification ("night vision") system is in direct contact with the sensor (i.e., special goggles).

- *The sensor does not always look downward, and the perspective is not always aerial.*

An aircraft pilot uses an infrared imaging system to look downward in a non-terrestrial perspective. Weather radar looks outward and upward from its antenna. The main products from weather radars involve a transformation in which data are viewed from an aerial perspective, but weather radar is also used to generate graph displays showing height in the atmosphere on the *y*-axis, time on the *x*-axis, with winds indicated by traditional "wind speed barb" symbols that use lines to indicate wind direction and speed. Such graphs do not depict aerial perspective.

What makes remote sensing "remote" is that either (1) the sensor is distant from the object of study by a scale equal to or greater than kilometers, or (2) the observer is distant from the sensor by a scale equal to or greater than kilometers, or both (1) and (2). Hence, remote sensing can be distinguished from "proximal sensing" and "micro-sensing" These terms designate such things as medical imaging, scanning tunneling microscopy, the bubble chamber of particle physics—any technology used to study things on micro scales. In these cases, the distance from the observer to the sensor is usually on the scale of meters and the distance from the sensor to the object of study is also on that or some smaller scale.

Although remote sensing evolved from aerial photography (visible-light), remote, proximal, and micro-sensing all involve the mediation of perception through a technology that detects some forms of information-containing energy that the human sensorium cannot detect—"nonliteral" imaging.* The human eye cannot see an electron, or an ultraviolet quasar, or the isothermal patterns indicative of plankton concentrations.

2.2 The "human factor" in remote sensing

Interpreting a remote sensing image is an open-ended task. The interpreter often does not know exactly what pattern is going to appear in an image. A project examining patterns of rainfall and drainage in a region might look for specific high-level properties of an area such as drainage basins. However, the specific perceptual manifestation of these high-level properties will differ from situation to situation, and thus, there is not a one-to-one mapping between particular properties of an image and the high level-properties it depicts. The job of the human interpreter is to make this link between the images and the high-level properties.

Remote sensing is an increasingly important technology in environmental and planetary science.[37,68,94] It is also becoming salient in the culture at large. Launches of new satellites (by both governmental agencies and commercial enterprises) with evermore advanced imaging capabilities have become almost routine. Satellite and radar images are routinely shown on televised weather forecasts, images from planetary probes adorn the covers of popular periodicals, and infrared imaging was daily fare on television news broadcasts during the 1991 Gulf War. But the wealth of new image data has created a problem—the "data analysis bottleneck." It would not be an overstatement to say that there are already enough satellite data in archives to keep the available image analysts busy for decades. Indeed, much data will probably decay before it can be analyzed in detail.† One design for the

*Despite vast differences in data types, the same human factors issues and the same practical issues (e.g., display design) apply equally well to micro-sensing as to remote sensing (e.g., expertise and display design in radiology), magnetic resonance imaging, X-ray diffraction crystallography, and so on.

†A hypothetical satellite of the type that initiated the field of remote sensing of the weather would transmit one visible light and one infrared photograph every half hour to fifteen minutes. Over ten years that would amount to between about 350,000 to 700,000 images. In orbit today are scores of remote sensing platforms of a great many types that collectively generate thousands, if not tens of thousands, of image data sets in a single day. Over the decades of satellite observation of the Earth, few image data sets have been preserved in image hardcopy form and countless digital tapes of remotely sensed data have decayed. As a matter of routine, some data are never preserved. For instance, the National Oceanic and Atmospheric Administration (NOAA) Weather Forecasting Offices only archive data that pertain to

Earth Observation Satellite system would have entailed the generation of over a terabyte of new multispectral image data per day.

Given the ever-increasing volume of remotely sensed data awaiting analysis, and given the continuing improvements in sensors and information processing systems, the emphasis in the field of remote sensing has been on widening the bottleneck through the use of automated data analysis approaches such as pattern recognition and algorithmic spectral analysis (cf. [83,105]). However, pattern recognition and automatic image processing techniques remain inadequate.[67] The human **must** be "in the loop," since the human, unlike the computer, can perceive and can form (and re-form) concepts. As Campbell put it in his text, *Introduction to Remote Sensing*:

> Although [computer analysis of pixel values] offers the benefits of simplicity and economy, it is not capable of exploiting information contained in relationships between each pixel and those that neighbor it. Human interpreters, for example, could derive very little information using the point-by-point approach.[24]

Even if most image analysis is ultimately conducted by machine, humans will continue to make the important interpretations, and make decisions based on the interpretations. They will continue to be the agent that creates the new algorithms that are used to process the data in the first place, and they will continue to do so in part by "fiddling" with their displays (e.g., the color coding schemes), and by reasoning around the interpretation anomalies that can arise as display schemes are applied to actual data sets (see [2]). It can be argued that the essential unity of remote sensing is expressed by the concept of "remote perceiving."[85,86] That is, the ultimate purpose of all the technology of sensors and information processing systems is to support the human interpretation of data to derive meaning and make decisions on the basis of the interpretation.

The importance of the human perceiver in the field of remote sensing is underscored in Campbell's text, in which nearly every chapter includes some allusion to the critical importance of the interpreter's knowledge and skill:

> . . . it is important to recognize that the role of the equipment is much more important, and much more subtle, than we may first appreciate. Display and analysis hardware determine in part how we perceive the data, and therefore how we use them. The equipment is, in effect, a filter through which we visualize data. We can never see them without this filter because there are so many data, and so many details, that we can never see them directly.

---

interesting or otherwise special cases of weather events. In other words, we are so overwhelmed that we are losing data at a considerable rate even while plans are being laid for generating terrabytes of new data.

> The human interpreter is good at distinguishing subtle differences in texture. . . . Direct recognition is the application of the interpreter's experience, skill, and judgment to associate the image patterns with informational classes. The process is essentially a qualitative, subjective analysis of the image using the elements of image interpretation as visual and logical clues . . . for image analysis, direct recognition must be a disciplined process, with a careful, systematic examination of the image.
>
> Today's interpreters of archived satellite imagery face some difficulties. Many of the original interpretations depended not only on the imagery itself but also on the skill and experience of the interpreters . . . many critical interpretations depended on the experience of the interpreters with long experience in analysis of images—experience that is not readily available to today's analysts.
>
> For manual interpretation, ancillary data have long been useful in the identification and delineation of features on aerial images. Such uses [are an] application of the interpreter's knowledge and experience. . . .
>
> The interpreter's task is not so much one of identifying separate objects as it is the accurate delineation of regions of relatively uniform composition and appearance. The goal should be to perform this mental generalization in a consistent, logical manner and to describe the procedures accurately. . . . The guiding principles should be the attainment of visual and logical clarity . . . [supporting information] includes not only the formal written material that accompanies the map [that is produced from remote sensing imagery] but also the wider realm of knowledge that [is used to] examine and evaluate the map.[24]

Statements of precisely these kinds can be found in every text on remote sensing, and all of them serve to highlight the importance of the interpreter's knowledge, skill, and experience in the interpretation of all forms of imagery (e.g., visible-light amplification, multispectral imagery, thermal imagery, radar, etc.) and all forms of processed imagery (e.g., ratio images, false-color images, multispectral chromaticity spaces, photoclinometric images, and so on). The interpreter's knowledge is, and must be, brought to bear on all aspects of image interpretation—the recognition of soil and rock types through the perception of tones, the discrimination of plant types in normalized vegetation indexed images, and so on.

In addition to the critical role of the human interpreter in routine procedures, important discoveries have been made in remote sensing (e.g., planetary science, archaeological applications, geomorphology, and so on), not through automated analysis, but through the human perception-interpretation process.[89,93] Here are a few of the countless stories that are suggestive of the cognitive foundation of remote sensing.

- The signs [of auroras on Venus] were first noticed in 1982 when Larry Paxton, now with the Naval Research Lab, saw a "few puzzling patches of brightness" on the Ultraviolet Emissions Spectrometer about the Pioneer Venus Orbiter (UVS). . . . Paxton initially wondered if there might be something wrong with the image. A glance showed similar features on other images, so he next looked to see if there was perhaps a mistake in the mapping algorithm that had been used to make pictures from the spacecraft's data. There was such an error, it turned out, but correcting it just made the patches brighter (adapted from [49]).
- Though the National Oceanic and Atmospheric Administration (NOAA) satellites were designed with meteorologists in mind, the remote sensing community discovered that the sensors aboard the polar-orbiting satellites could be used to distinguish forest from non-forest. On one channel in the mid-infrared region, forests appear cooler and clearings warmer. This was very fortuitous (adapted from [10]).
- Analysts can make vegetation appear green or red, or even blue in the false-color composite image just by assigning different colors to different bandwidths of digital imagery. . . . The Landsat Thematic Mapper does give an analyst flexibility, but it also introduces a problem. . . . The analyst can be easily misled by the composite image . . . when swapping colors and bands if he does not understand the spectral characteristics of soil, rock, and vegetation. Typically, analysts create a standard false-color composite image by assigning bands 2, 3, and 4 to the blue, green, and red guns on the Cathode Ray Tube (CRT). [But] some analysts feel that the image "looks" better if they use band 7 instead of 2. As a result, they create a composite image that may cause wrong interpretation. . . . An interpreter can make any color associations he wishes, but he may have difficulty in interpreting the resulting composite image (adapted from [32]).
- Most of our current knowledge of galaxy morphology is based on the pioneering work of several dedicated observers who have classified and cataloged thousands of galaxies. . . . It is remarkable that . . . subjective classification labels for galaxies correlate well with physical properties such as color, dynamical properties (e.g., rotation curves and stellar velocity dispersions), and the mass of neutral hydrogen. Human classifiers will need all the help they can get to cope with the flood of data that is expected from efforts like the Sloan Digital Sky Survey, which by itself is expected to image more than a million galaxies. [To take] a first step toward finding an automated method of galaxy classification, we compiled a well-defined sample of galaxy images. . . . One of the galaxies got exactly the same classification by all six observers, but there was no such clear agreement on the other three galaxies. . . . Any classification depends on the color, size, and quality of the images used. . . . Our comparison indicates that although the [galaxy classification system] is convenient, the scatter

between observers is not negligible. Further work will focus on supervised [classification] to preserve human experience in multidimensional classification (paraphrased from [105], see also [60]).

- In a project involving the relocation of indigenous peoples in New Guinea, researchers presented false-color infrared Landsat images to a group of natives. "The most amazing thing that struck me every time was the ability of the local people to immediately and intuitively understand the geometry of Landsat images . . . at a glance they recognized their villages and other local features. They do not inquire where the pictures come from, nor do they have an inkling of the technology behind it all, but they understand the information and can use it immediately" (Wine Langeraar quoted in [53]).
- In the process of examining various bandwidths of Voyager data that came back from the moon Io, Linda Morabito of the Jet Propulsion Laboratory in Pasadena, CA was performing a contrast enhancement operation in order to resolve the faint background stars. Such an interpretation process is usually driven by its particular objectives, and only the objects of interest are studied. However, she observed an anomalous region of energy at the horizon of the moon. With a bit more manipulation of the contrast, it became clear that she had happened to catch a volcano in the act of erupting. The plumes would have been discovered eventually during systematic image processing and interpretation. However, this discovery of active volcanoes on another planetary body was very exciting (adapted from [77]).
- Beginning in 1982, a device called the Synthetic Aperture Radar was flown aboard Space Shuttle missions, with the purpose of mapping the surface of the earth. ". . . . scientists at the U.S. Geological Survey (USGS) saw some Shuttle Imaging Radar (SIR) pictures that changed their conception of the Sahara Desert's underlying structure. The Shuttle Imaging Radar System (SIR-A) sent back picture-like data revealing a vast network of valleys and smaller channels winding beneath the desert sands. In the Sahara's super-dry core region, the radar penetrated right through the sand to reveal gravel terraces and river banks surrounding an ancient drainage system." [169] Radars, generally, do not penetrate soils. But, unexpectedly, the SIR did penetrate the dry desert soils. Within a second or so of their very first glance at the pictures, the pictures spoke a thousand words to the USGS scientists. The initial discovery, and subsequent excavations (geological, hydrological, and archaeological) all made the headlines.

2.3 Psychological research related to remote sensing

As the previous section demonstrates, there are many ways in which the human contributes to the analysis of remote sensing images. There is a considerable amount of pertinent psychological research that has explored phenomena including: concept formation, graphical perception, expert/novice differences, and perceptual learning. We now summarize some of that research, organizing the discussion around aspects of perceptual learning, a concept that provides a unifying theme for this volume. Following this discussion, we present an experiment that illustrates perceptual learning in the context of remote sensing image interpretation.

2.3.1 Research on the interpretation of diagrammatic and pictorial displays

Because the human visual system is designed to recognize complex patterns, graphic displays can be developed to assist in the interpretation of complex data sets. Thus, psychological factors have become an important consideration in the display of multidimensional data (see [43,74,189]), including cartographic data[7,18,42,45,123,124,125,145,164,177] and meteorological data (e.g., satellite imagery, radar, etc.).[92] Topics in visualization and diagrammatic reasoning have become "hot" in the field of artificial intelligence (see [22]).

2.3.1.1 Diagrammatic reasoning

In concert with recent developments in scientific visualization, an extensive body of research has demonstrated that the solving of abstract problems (e.g., algebra, logic, combinatorics) and puzzles (e.g., the "Tower of Hanoi" problem) can be significantly facilitated if the problems are accompanied by diagrams that depict concepts, functions, and relations (e.g., [4,96,106,107,193]). Psychological research has shown that the comprehension of illustrations that accompany text depends on the type and difficulty level of the text, the level of expertise of the learner, the strategies the learner uses, and the relation between the type of material to be learned and the type of graphic (see [64,99,147,194,195]). Scientific material presented in textbooks can be more readily understood if accompanied by appropriate diagrams, e.g., showing clouds and the distribution of electric charges that cause lightning, diagrams showing the structure and operation of machines, electronic circuits, problems in mechanics involving pulleys and springs, etc. (see [4,108,126,139]), although textbooks often do not use graphics appropriately or effectively (see [112,197]).

An example is the work of educational psychologist Richard Mayer and colleagues.[130,131] They have conducted a number of experiments on how college-age students learn from illustrations in scientific texts. They have demonstrated that comprehension and recall of passages about common machines (e.g., the workings of a brake mechanism or a pump mechanism) and responses to a problem transfer task improve when the text is accompanied by labeled illustrations, especially for students who have had little household repair experience. Presenting the figure labels alone does not help, even though labels—for the most part—repeated information that was contained in the text. Students who saw the text with labeled illustrations showed better transfer performance (e.g., answers to what-if questions about the problem situation) and greater recall of explanation-related information. Furthermore, multiframe illustrations that showed the steps or sequences of events (e.g., what happens as a pump rod is pulled out) as well as those that labeled the parts of the mechanism were more instructive and beneficial to problem solving transfer than illustrations that just labeled the parts. In a study using animations,[130] those that were accompanied by verbal descriptions were more beneficial to problem solving transfer than when the verbal explanation was presented before the animation, even though recall of the verbal information was about the same in both conditions. Mayer's view is that labeled dynamical illustrations assist in the formation of useful mental models that combine principle-based understanding of causation with a dynamic, imagistic understanding.

2.3.1.2 Pictorial reasoning

Laboratory research on the effects of pictorial displays on performance in stimulus detection tasks, multicue judgment tasks, and memory tasks, has demonstrated that display formats can hinder or facilitate the perception and comprehension of information about spatial layouts, dynamic events, and causal processes, and the configural relations of systems that are defined by multiple variables or parameters (e.g.,[73,80]). Research has also shown that useful displays are often those that depict meanings pictorially, and dynamically mimic events and processes.[5,19,41,44,50,52,190]

Psychological studies of visual search in graphic displays have often utilized matrix-like arrays of symbols (e.g., letters, numbers, regular geometric shapes).[25,33,189,191] Psychological research using graphic displays has traditionally used highly impoverished artificial "maps" (e.g., a few simple contour lines).[34] In part, the artificiality of the stimuli is a consequence of the need for experimental control—to permit factorial manipulation of such variables as data density, the colors and sizes of areas or symbols, and target predictability. However, most of the results end up being of marginal applicability to real displays, such as maps, in which colors are variegated, shapes are convoluted, the data are multidimensional, etc.

Fortunately, some research has departed from the "artificial stimulus" tradition and investigated the perception and interpretation of remote sensing imagery.

2.3.2 Reasoning about remotely sensed imagery

While it has been argued that psychological processes are fundamental in the field of remote sensing,[86,89,92] psychological research is just beginning to examine the processes whereby experts and novices perceive and interpret remote sensing imagery (e.g.,[56,65,141,185]). While there is a tradition of research on "cognitive maps" and psychological factors in the interpretation and understanding of topographical maps and orienteering (cf. [20,51,59,81,165,180]), only recently has applied research shown that psychological factors of comprehension and memory are critical in the interpretation of meteorological charts[119] and aerial photos.[82]

> . . . there has been little work in understanding the human perception or cognition of remotely sensed imagery. Despite this lack of research, we often design automated measures of image cues, such as texture or pattern, with an implicit assumption that we are emulating visual cues or at least matching human performance. . . . But how do we know that such measures emulate visual cues of human processes without studying the human processes and performance?[82]

In his seminal studies on this, Hodgson[82,83] asked how much local information is needed for the human interpreter to classify a pixel. The automated analysis of pixel data to assign classification typically relies on small windows of 3×3 to 9×9 (for reasons having to do with computational efficiency, image resolution, etc.), but is such a window size adequate for the human interpreter? For smaller window sizes, neither trained nor untrained participants should perform well at a classification task, whereas for larger window sizes perhaps anyone could perform reasonably well.

In the experiment, black and white aerial photos covering the Denver metropolitan area were cropped to a range of window sizes (from 10×10 pixels up to 100×100 pixels), and the cropped images were presented to college students who had completed a course in aerial photointerpretation. Their task was to classify pixels into land-use categories (i.e., *"Is this area predominately residential, commercial, or transportation?"*). Since humans "synergistically use a number of image interpretation elements, such as tone, pattern, texture, shape, size and association"[82] in what cognitive psychologists refer to as "top-down" processing, it was expected that the optimum window size would be greater than that typically used in the automated analysis of imagery. For example, areas of commercial land use can involve large parking lots and buildings, requiring larger windows for correct classification. The research findings confirmed the notion that smaller window sizes "constrain adequate contextual analysis"[82] (i.e., top-down processing) upon which human perception depends:

> In normal circumstances, the photo interpreter may proceed from known to unknown parts of a photograph, thereby using information gained from one part of the scene in classifying other parts . . . [this] calls into question the notion of using small windows, regardless of the classification algorithm. . . . If the human cannot correctly classify small subimages, then should we expect an automated classifier to do so?[82]

Aesthetic judgments also come into play in the interpretation of remotely sensed imagery. Sociologist Michael Lynch[121,122] interviewed a number of expert astronomers about the "representation craft" of generating false-color displays based on multispectral data. The raw data from space probes and satellites can be portrayed using a variety of color palettes, some of which can be misleading. False-color multispectral satellite images that cover a range of visible and nonvisible bandwidths (e.g., infrared, radar, X-ray, etc.) can create special problems for display design and interpretation. A clear example is the photographs of the planets that NASA produced from Voyager data. Contrary to those brilliant and colorful depictions, most of the planets look rather like dirty tennis balls; they are not vivid. The "Great Red Spot" is not red; the moon Io does not look like a pizza (see [200]). For public relations purposes, astronomers orient explicitly to the aesthetic aspects of their images, and are unashamed to talk about the "pretty pictures" that adorn the hallways of their research facilities. For publications in popular outlets, experts sometimes change their figures to match the goals of their text and the figure captions, rather than writing captions to fit the figures. Even for technical publications, images are tailored to show the features that are being discussed (e.g., for an article about radio-emitting nebulae, an image may be recolorized to make the apparent emission pattern match a visible-light image).

The primary or scientific goal of image processing is to support scientists' accurate perception and comprehension. As one expert astronomer put it, "through a complex series of adjustments and modifications of an image. . . . [a display should] enable researchers to *see* the physics."[121] To make this point, Lynch attempted to show that the aesthetic judgments are not so distinct from the scientific ones. Most of the expert astronomers insisted that their actual scientific work involves working with raw, quantitative data, perhaps depicted graphically but even then, using a more "boring" gray-scale rather than fancy colors. However, Lynch probed the experts about what it was that made some of their color graphic products especially pleasing:

> E (expert): This is one that starts out red and then goes through yellow and white. Here's some radio data thats multicolor. I like these maps that have sort of one color, they start out dim, then go brighter, then go white. I think those are prettier than these that have many different colors.
>
> I (interviewer): Is it strictly a matter of liking one color?
>
> E: The reasons are artistic rather than scientific.
>
> I: Other people told me they like a more uniform thing because it is not misleading.
>
> E: Yes, there's that too. You show this picture to someone and you say "This is what it looks like." Now, you can't really see it; it doesn't really look like anything. But this does seem to be more realistic.
>
> I: Choice of color would have some relation if it were an extension of the spectrum . . .

> E: Right. At one time I thought I wanted to make the X-ray purple and the radio red because it gives you an idea that this is a higher energy photon.[121]

The experts' "aesthetic" judgments clearly play a role, but do not do so in any arbitrary way—they are driven by their foundational scientific knowledge and their goals.

An awareness of the need for new approaches to visualization has led to a number of collaborative research and development projects. For instance, IBM's Lloyd A. Treinish (a space scientist by training) and Bernice E. Rogowitz (a psychologist by training), have been developing new graphic displays for the visualization of remotely sensed data (e.g., [154,155,162,182,183]). Software systems that are currently in use offer a plethora of mapping choices and tools, but provide little guidance in their application to specific tasks. Depending on the architecture and interface of the system, the user must either define the mapping(s) between data and pictures via the tools in such systems, or operate within a fixed set of predefined mappings, which may have little facility for customization or control. In both cases, the mappings that are available are driven by the structure of the data, not the task or the human's goal for the visualization. Hence, the design of visualizations can be time-consuming and often requires the assistance of computer graphics experts, who are not domain scientists.

Some of the IBM research was aimed at coping with large volumes of rapidly produced numerical weather forecasts as well as remotely sensed data as an aid in the presentation of information to the layman via broadcast media and the World Wide Web. To help address this problem, Treinish applied work done with Rogowitz on the creation of a rule-based advisory tool for the specification of appropriate color maps.[184] Their rule-based system assists the user by ensuring that data content is reflected in images and that perceptual artifacts are not erroneously interpreted as data features. The system also provides advice on representation depending on whether the goal of visualization is exploration or presentation.

The approach to visualization hinges on three premises:

- Traditional graphical representations of data will not suffice.
- Data structures have to be mapped onto perceptual structures.
- Rule-based systems can help users make good decisions about the visualization of data without requiring the user to become an expert in human vision (i.e., depth perception, color theory, etc.) or computer graphics (e.g., data structures, algorithms, etc.).

Instead of static or simple flip-book animations of two-dimensional contour maps, novel three-dimensional visualization strategies have been employed that preserve the fidelity of the original data as much as possible, and yet allow the user to define coordinate systems onto which data may be registered in space and time. Techniques support the registration of multiple data sets in geographic coordinates using cartographic warping of the data locations. Treinish et al. have found that the use of appropriately warped curvilinear grids can preserve the fidelity (e.g., missing data, original grid resolution) of the data—the transformations are both topologically and data-invariant. The selection of cartographic projections is dictated by the requirements of the user in the design of an appropriate visualization, as opposed to limitations of the system in use or the original characteristics of the data.

Examples of their innovative graphic products appear in Figures 2.1, 2.2, and 2.3. Figure 2.1 is a perspectival display illustrating the incorporation of multiple data types. Surface wind speed is overlaid on a surface chart (using a color palette of saturation shades of violet), and precipitation is indicated by a pastel (desaturated) "rainbow" palette. Also shown is a perspectival view of cloud structures based on radar data. Figure 2.2 is a perspectival display demonstrating a warping transformation. In this case, the column of ozone (the ozone "hole") is directly perceptible as a deformed surface. Figure 2.3 is another perspectival display illustrating how a single display

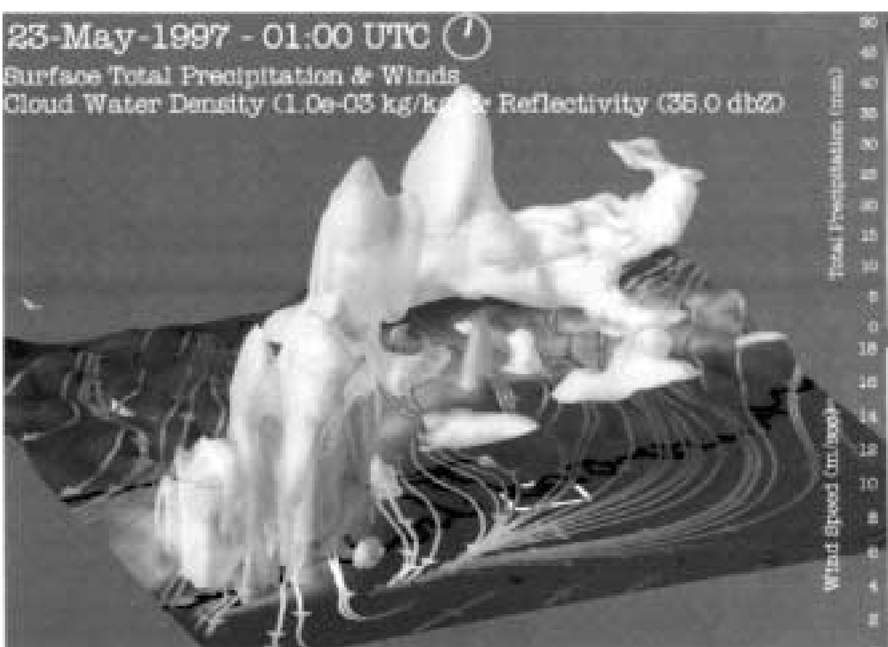

Figure 2.1 A perspectival display illustrating the incorporation of multiple data types. See color version of this figure in the color section following page 114. (From Treinish, L. A. and Rothfusz, L., *Proceedings of the Thirteenth International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography and Hydrology,* American Meteorological Society, Boston, 31–34, 1997. With permission.)

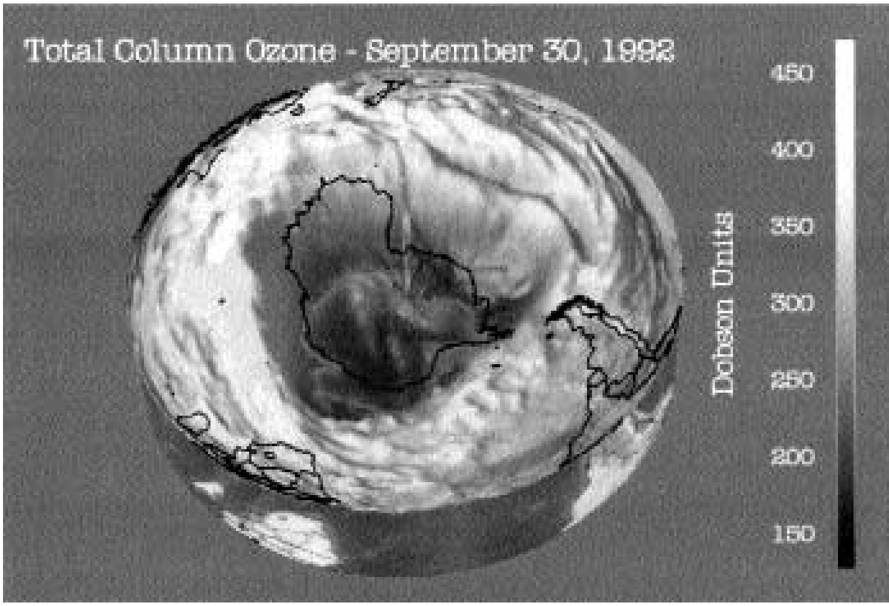

Figure 2.2 A perspectival display demonstrating a warping transformation. See color version of this figure in the color section following page 114. (From Treinish, L. A. and Rothfusz, L., in *Proceedings of the Thirteenth International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography and Hydrology*, American Meteorological Society, Boston, 31–34, 1997. With permission.)

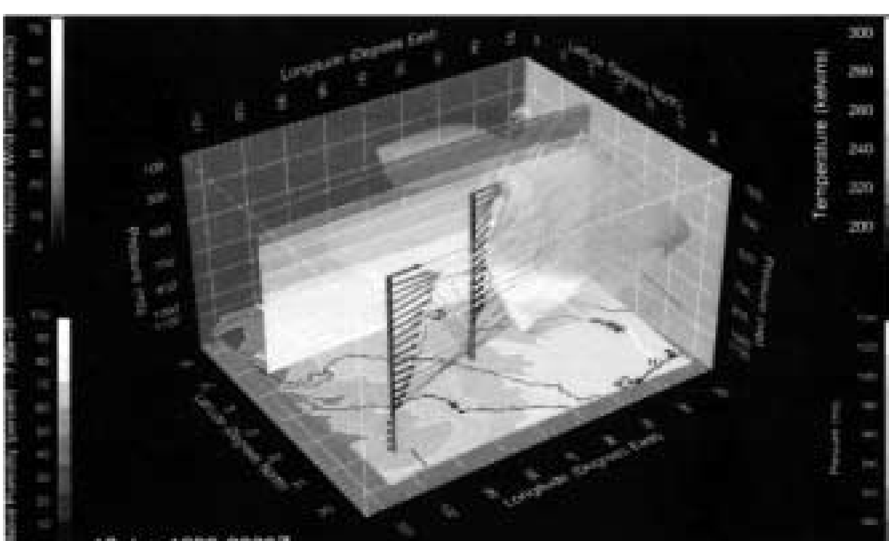

Figure 2.3 A perspectival display illustrating the incorporation of multiple data types. See color version of this figure in the color section following page 114. (Figure courtesy of L. Treinish, IBM.)

can portray multiple data types. This image portrays horizontal winds (using a color pallette of saturation shades of violet), relative humidity (using saturation shades of brown), surface temperature overlaid on the base map (using a two-tone palette of saturation shades of blue and green-blue), and air pressure (indicated in a semitransparent vertical plane using saturation shades of blue-violet and green). Also depicted are three-dimensional cloud structures. For all of these graphic products, the use of perspective, depth pseudoplanes, and animation permits the perceptual discrimination of the multiple variables (other images can be viewed at http://www.research.ibm.com/people/l/lloydt/ and http://www.research.ibm.com/weather).

2.3.3 *Research on concept formation*

Recognizing that a specific pattern is an instance of a high-level feature is a perceptual classification task. Perceptual classification involves a number of important subtasks. The relevant set of perceptual features must be extracted from the items. Part of this process involves actually learning which aspects of the stimulus should be construed as features.[168] This perceptual learning may occur during the process of classifying items, or it may involve people's theories about the domain.[196] Once the features have been extracted, it is necessary to determine the combinations of properties that distinguish one category from another.[144]

An important focus of current research in categorization is how the way people interact with a set of items influences what is learned about it.[128,158,159,160,161,196] "Humans are inherently comparing creatures, and whether instructed or not we simply jump in and compare when interacting with a novel set of materials."[23] Research has confirmed the intuition that categories can be learned incidentally through exposure as well as through an intentional process of analyzing features and relations (e.g., correlated attributes), even though patterns of attention differ for incidental vs. deliberate learning (for reviews, see [17,188]). However, traditional psychological research has focused on a directed classification task in which people see a series of examples (of simple, and often artificial stimuli composed of schematic faces, geometric patterns, etc.) and they are asked to classify them (e.g., [16,58,156]).

Tasks such as interpreting remote sensing images do not typically involve this kind of trial and error learning. Instead, the high-level categories of interest in an image interpretation setting are learned in the context of identifying items in the images, using the images to solve other problems, and communicating the results to others. The categories are learned as a byproduct of this other kind of reasoning. This distinction is important, because categories learned in laboratory tasks tend to involve the formation of rules and the memorization of exceptions.[144] In contrast, categories formed incidentally to performing a primary task are less rule-governed, and tend to emphasize those features that were important for solving the primary task.[127,158,159,160,161,198,199]

2.3.4 Research on expert/novice differences

On the basis of recent psychological research on expertise,[31,55,87,111] including expert / novice differences in map interpretation and the use of cartographic visualization tools,[133] one would expect major differences between the perceptual skills of the expert and the initiate image analyst. Such differences, once specified, could have implications for instructional design, image or display design, workstation design, and the development of decision aids or expert systems. Training issues are critical since it takes many years to become an expert image analyst.

Kibbe and Stiff conducted a study of novice / trainee interpretation of aerial photographs.[100] Participants were shown aerial photographs overlaid by line drawings that outlined some of the salient features in the pictures (e.g., hill contours, roadways, etc.). Some of the line drawings were sparse, some were more detailed. Some were properly aligned with the photo that they superimposed and some were misaligned, and the participants' task was speeded recognition of (mis)alignment. Line drawings that indicated the longer features were responded to more rapidly (by about 1 s, with 10% greater accuracy) than line drawings with short lines, but only when the line drawings were properly aligned. In one condition, the participants were initially exposed to the stimuli in a task in which they themselves attempted to produce outline drawings for each of the photographs. This active engagement in the interpretation task had a positive effect in their subsequent recognition reaction time performance for the properly aligned drawings, even though the line drawings used in the task were not their own.

Moray and Richards attempted to demonstrate the expert / novice difference in a study of memory for radar-like displays.[135] A group of college-age students, a group of Royal Air Force (RAF) fighter controllers (with some experience with radar), and a group of RAF aerospace systems engineers (with more radar experience) were presented with a series of simple radar-like displays, basically showing a circle with one or more "targets" indicated by a small letter "X" somewhere within the circle. After seeing each stimulus for about 10 s (about equal to 1 radar sweep), they had to try and reproduce the targets by redrawing the stimulus, either immediately or following a delay (of up to 30 s). Both accuracy and confidence tended to decline as the delay period was increased. However, there appeared no statistically significant differences between the two expert groups in overall recall accuracy, although there was a trend for the more highly experienced of the experts to do slightly better at the task. The effect of expertise was more manifest for the more highly experienced participants in one respect—for displays involving three (as opposed to one or two) targets, the participants reported that they were able to remember the pattern, leading to better performance and increased confidence. In general, ratings of confidence were not a very good indicator of recall accuracy, a finding common in the literature on memory and expertise.[140]

As the Kibbe and Stiff[100] and Moray and Richards[135] studies suggest, quite a lot of the modern research on expertise has involved domains in which the interpretation of nonliteral imagery is a primary task. Hence, this work pertains to the topic of perceptual learning.

2.3.5 Research on perceptual learning

By tradition in psychology, the processes of perception and learning are for the most part treated separately (e.g., separate chapters in the standard textbooks). But learning is not just the accumulation or storage of propositional facts and static knowledge structures. Across development, there is a change in the way information is acquired. At first, attention is passively "captured" by salient or easily isolated stimulus features. With experience and practice, new distinctive features are discovered, permitting the active strategic search for critical information.[15,75,76,98,137] One gains an ability to rapidly detect and discriminate, and not just isolated features or cues. Rather, one learns to attend to the invariant patterns or cue configurations that specify distinctive properties, patterns which were not previously noticed or were discerned only with difficulty.[30,54,57,70,143]

Perceptual learning, operationally defined, is the increase in the ability to discriminate and extract information as a result of practice and experience, as assessed in terms of an increased specificity of the responses to stimuli.[54,70] Over the course of development, perceptual skills change, permitting the rapid search, discrimination, recognition, and comprehension of complex informational patterns, provided that the responses are mapped consistently to specific elements or patterns in the visual scene (or display).[26,36,62,78,142,148,150,172]

As one acquires a perceptual skill involving the classification of complex patterns, performance can come to reflect the underlying rules, yet the human may be unable to articulate the relevant dimensions of difference, that is, the knowledge becomes "tacit."[1,13,14,72,103,111,113–115,143,149–152,172,187,188] The explanations people give for what they are doing may be justifications rather than descriptions of the rules they are following. Furthermore, performance at the classification of complex patterns can improve without the learner developing a deep conceptual understanding of the significances in the patterns.[48,63,146,167] Rich discriminations can be made without conscious analysis, verbal labeling, or even awareness of isolated features. Indeed, at high levels of skill, complex stimuli can be perceived at a rate that precludes verbal labeling, perhaps illustrated best in the research on the learning of Morse code.[70]*

A great deal of psychological research on perceptual learning has focused on child development (e.g., depth perception, locomotion, and object perception). However, in her classic book on perceptual learning and development, Eleanor Gibson pointed to the role of perceptual learning in such expert skills as sonar interpretation, X-ray interpretation, identifying aircraft by their silhouettes, and infrared aerial photointerpretation.[70] Skill automatization has been observed in a number of studies of expertise.[27,38,39,40,44,69,170] As one expert personnel manager observed, "When I was younger I used to carefully consider all the aspects, get lots of opinions, and agonize over decisions. But now, I can generally just recognize what to do."[171] Indeed, the automatization of skill is believed to characterize the shift from novice to expert.[46,71,101,102,110]

For instance, Myles-Worsley et al. had expert radiologists and medical students observe and then attempt to recognize a series of chest X-rays.[137] Reaction time data showed that experts allocated their attention more efficiently, focusing on abnormal features that distinguished the images. Similarly, Norman et al. demonstrated the effect of perceptual learning in a study of student, intern, and resident dermatologists, in which the participants were shown 100 slides for diagnosis.[143] The researchers measured reaction time and error rates in a subsequent recognition task. The reaction time data showed that experts engaged in a rapid perceptual process that does not rely on independent analyses of separate cues. In other words, reaction times and errors were not predictable on the basis of the simple presence or absence of isolated features that are typical of each diagnostic category.

Research has shown that experts in domains as diverse as architecture, nursing, and electronic circuit design can indeed "see" and evaluate things

---

*Lewicki, Hill, and Bizot[115] conducted an experiment in which participants were run in thousands of trials, each trial consisting of a matrix of four numbers. One of the numbers was the "target," and its position would differ in the matrices across blocks of 7 trials. The task was speeded reaction-time—the participants pressed a key to indicate the position of the target on each trial. A complex rule was used to determine the sequence of target movements (e.g., if the target appeared in quadrants 3, 2, 4, and 1 in trials 1, 3, 4, and 6, then it would appear in quadrant 1 on trial 7). Over several thousand trials, participants' reaction times and error rates for trial seven decreased significantly. Yet, participants were not able to verbalize the underlying rule or even give it a coherent characterization. Interestingly, when taken out of the speeded reaction time task and asked to predict the position of the target on trial 7, they performed no better than chance. Such results obtained even when participants were offered a considerable award ($100) if they could figure out the rule, and even when participants were college faculty who knew that the experiment involved studying the nonconscious processes of perceptual learning.

that novices cannot.[98,111] When novice livestock judges confront the judgment task they can miss seeing important features of livestock that experts readily detect[171]; the eye movements of radiologists while they scan X-ray films are quite different from those of novices—the experts can selectively search for abnormalities; expert cardiologists directly comprehend cardiovascular biomechanical events as perceived through a stethoscope.[98] In many domains of expert decision making, such as fire fighting, power plant operation, jurisprudence, and design engineering, experts often make decisions through rapid recognition of causal factors and goals, rather than through any explicit process of generating and evaluating solutions.[101]

In an example from remote sensing, experts at the identification of rocks and landforms seem to directly recognize different types, even though they were originally taught the diagnostic features explicitly. Hoffman studied terrain analysts with the U.S. Army Corps of Engineers, experts at interpreting aerial photographs for engineering purposes.[84] The experts had behind them decades of experience at analyzing aerial photographs to assess soils, bedrock, vegetation patterns, drainage patterns, and the like. In one experiment, experts were asked to think aloud about their perceptions as they viewed aerial photographs:

> Based on what I see in this photo, this is a semiarid climate. Over here you see there's a gently rolling, irregular plain. The gullies down here are V- and U-shaped. Lots of lowlands . . . rounded contours. But over here are some escarpments. The gradients are fairly uniform, so it's homogeneous rock. About midway down the slopes is a relatively thick tonal band, so there's a thick bed there, pretty thick, and at the base it's flared and a bit scalloped. This terrain is flat shale, with loess topsoil.[93]

Exactly how did the expert know that the climate is semiarid? Exactly how did the expert know that the shale is flat? Which of the expert's statements were premises about perceived features, and which were conclusions? What does it mean to say that a plain is "irregular"? Many things were being perceived, but these were stated in terms of the higher-level concepts of terrain analysis (e.g., "escarpment"), not in terms of specific perceptible cues. The informational cues were certainly there in the photos, but they could only be seen by the novice when they were pointed out and explained.

In another experiment, experts were presented a photograph for inspection, but were only allowed two minutes of viewing time—ordinarily the full systematic terrain analysis process can take hours. After the viewing period, the experts were allowed two minutes in which to report everything that could be remembered about the photo. In one particular trial, an expert began his retrospection by asserting that any personnel sent to the depicted area would need to be prepared for certain types of bacterial infection. The experimenter's response was, "You can see bacteria in a pond taken from 40,000 feet?" Here was the reasoning sequence that the expert recounted:

> The photo covered an area of tropical climate. The vegetation was mature and uniform, so the contours to the top of the tree canopy could be taken as a reflection of the contour of the underlying soil, and since the soil layer would be relatively thin, the contour to the tree canopy reflected the underlying bedrock, which appeared to be tilted interbedded limestone. The bedrock also determined the pattern to the streams and ponds, and there appeared one pond that did not have a major distributary running away from it. Given the climate, the vegetation (tropical legumes) and the stagnant water, the presence of bacteria was a sure bet.[93]

This appears to be a long chain of inferences, dependent on a great deal of conceptual knowledge. Yet, in the actual experimental trial, the expert's judgment was very rapid, the sort of judgment that one might be inclined to call "direct" or "immediate"—more of a perceptual thing than a linear, conscious, and deliberative process.

*2.4 A demonstration experiment: expert/novice differences in the interpretation of aerial thermograms**

To illustrate the human factor in remote sensing, consider the false-color aerial thermogram shown in Figure 2.4. This image covers an area of about four city blocks in a small town in the midwest U.S. The depiction is of mid- and far-infrared radiometric data, displayed in terms of a "rainbow code" that is standard in this application (cf. [138]), i.e., the assessment of patterns of energy conservation.[2] The colors depict temperatures falling in the range of -5 to 7°C, a range that includes most of the wintertime variation in temperatures of surfaces (glass, brick, siding, heat vents, etc.), and other sources such as trees, small bodies of water, forested areas, etc. The high extreme temperature is coded as white (which did not fit into the 8×10-in. photo frame); temperatures at and below the low extreme are coded as black. Temperatures between are assigned to colors (about 0.7°C per color) based on common

*Walter Carnahan (Indiana State University) provided the aerial thermograms used in this study. The authors would like to thank Scott Lissner (Longwood College), Robin Akerstrom, Andrea Bubka, Stan Grabon, Garry Zaslow (Adelphi University) for helping in the collection and analysis of the data. Finally, thanks to Mark Detweiler (Pennsylvania State University) for his comments on a draft version of this material.

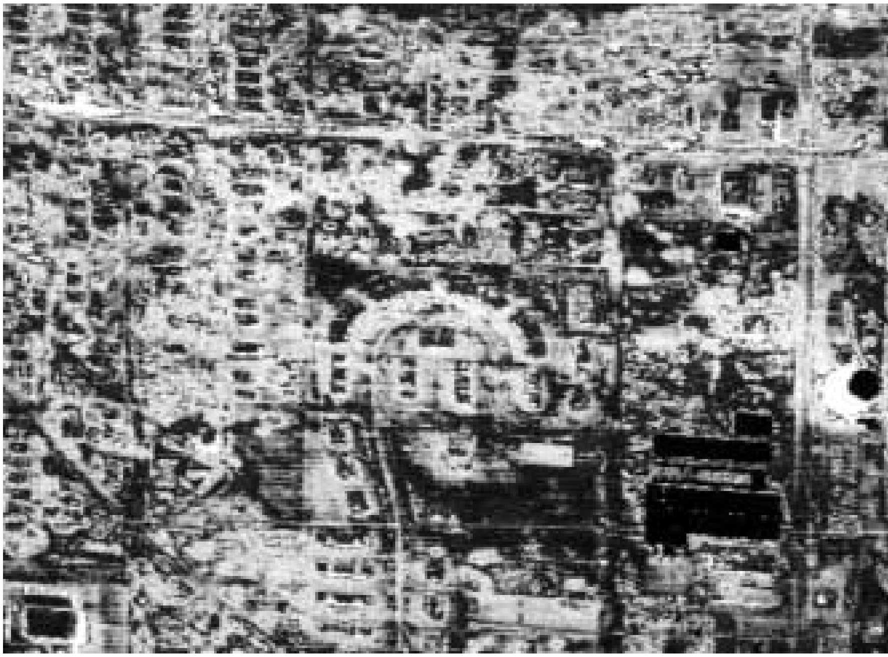

Figure 2.4 An aerial thermogram of a suburban area. See color version of this figure in the color section following page 114. (Indiana State University Thermography Project. With permission.)

associations: “Cold” (blue hues), “Cool” (green hues), “Warm” (red and orange hues), and “Warmer” (yellow hues).

While perhaps reasonable *a priori,* remote sensing scientists have learned that such color codes can generate interpretation anomalies when applied to actual data sets.[2,176] In the case of thermograms, flat water surfaces and metal roofs can act as infrared mirrors, reflecting the sky temperature and appearing anomalously cold (right-hand corner of Figure 2.4). As a second example, a salient color boundary (green-blue adjacent to the red) straddles a critical temperature—buildings with poor insulation get coded as red to yellow, houses with good insulation get coded as green. By design, this makes it easy for the expert to discriminate buildings in need of better insulation. Experience with students, however, suggested that the naive viewer is inclined to interpret the “green blobs” as trees. Coincidentally, trees are relatively warm in winter thermography and appear as “yellow blobs,” which initiates seem inclined to interpret as houses with poor insulation.

The motivation for the studies we describe stems from the pedagogical strategy taken in most textbooks on remote sensing. The introductions of texts and manuals (e.g., [35,117,173,174,192]) invariably focus on technical information about electromagnetism and sensors. Some include one or a few examples of multispectral imagery (e.g.,[37,175]), but the reader's attention is directed only at the feature that is the focus of the application (land use, hydrology, geology, etc.). Rarely is an explanation given of how a classification was determined, or of how a false-color scheme was established. There is no guidance to interpretation; there is no encouragement to reason. We designed this study to include conditions that would emulate these aspects of the textbook pedagogical strategy.

It is known that the format and focus of instructions can shape the learner's strategies in concept formation[179] and in the learning of both simple puzzle tasks and more complex tasks such as mathematical problem solving and computer programming.[109,118,153] However, explicit introductory teaching alone does not necessarily promote skill acquisition, whether in applied contexts such as electronics troubleshooting[163] or in academic contexts such as mathematics puzzle solving.[132] In a study in which novices (college students) were taught about the process of plastic extrusion,[104] one group was first taught about the general domain concepts (mixing, injecting, etc.) and the multiple relations between the concepts, and was subsequently taught about the lower-level details (e.g., the operation of the machine screw as a function of injection pressure). Another group was taught in the opposite order (lower-level details followed by general concepts). A test of what the students had learned involved asking them to say whether domain concepts presented pairwise were or were not related. The results showed that the group that was first taught about the general level domain concepts performed better than the group that was first taught at the detailed level. For the acquisition of knowledge structures in complex domains, attention needs to be given to instruction on both the hierarchical relations (general to particular) among domain concepts, and the multiple relations between domain concepts.

Furthermore, there are important training issues involving the relation of instruction and hands-on practice vs. tests of knowledge acquired through the study of manuals. At an extreme called the "dissociation" effect, after practice at a complex task (e.g., the operation of an industrial process via a computer-driven graphical interface), people can sometimes provide a verbal description of the task while their performance remains poor. Conversely, in some conditions performance can be good and yet the person may be unable to give an adequate verbal explanation of what it is that he or she is doing in the task.[6,21,79,129,167]

What these findings from applied psychological research suggest for the pedagogical strategy used in remote sensing is that instruction should not proceed simply by diving into the gory details of electromagnetics. The hypothesis which we examined was that even for a skill that is highly reliant on perception and perceptual learning, initial verbal instructions can nevertheless make some positive difference. But this hypothesis must be qualified by "cognitive load" considerations (see [181]). In the learning of task procedures (e.g., physics problem solving), instructions can be detrimental if they require the learner to integrate multiple sources of information (e.g., diagrams, equations, worked examples, etc.) before the learner has developed an understanding of domain concepts, i.e., memory schemas that can be used to lessen the burden on working memory.[188] One goal of the present research was to see if a similar finding might obtain for the interpretation of remote sensing images—basic information about electromagnetism must be integrated with the task of picture perception. We hypothesized that providing elaborate *a priori* instructions that delve into details about thermograms might not promote identification performance, and might actually suppress the overall rate of responding in an identification task.

2.4.1 Experiment 1

The purposes of Experiment 1 were to assess expert performance at the interpretation task and conduct a reference analysis of the materials. One of the authors (R.R.H) served as the expert, having spent four years analyzing scores of aerial thermograms, visible light aerial photographs, and radar images, and hundreds of visible and infrared meteorological satellite images. He had received his training in image interpretation at the U.S. Army Corps of Engineers and the U.S. Air Force Geophysics Laboratory.

On each of nine trials, the expert was presented a thermogram and generated identifications for a period of five minutes. A sheet of celluloid was laid over each thermogram, and using a felt-tip pen the expert outlined the identified features and objects while providing a verbal description which was recorded on audiotape.

For four stimuli depicting rural regions, the expert identified an average of 27 things and features. Within an average of about three minutes the expert asserted that he had identified all of the interpretable features. For the five stimuli that depicted suburban regions, an average of 107 things were identified, but there were yet more things that could have been identified when three of the trials ended.

In most cases, confirmation of the expert's identifications (e.g., identifying a form as a house with poor insulation, identifying a road, a stream, etc.) was totally unnecessary. A few identifications had to be confirmed either by reference to training manuals or to ground truth. The few remaining uncertain identifications, all apparent thermal anomalies, were explained in light of on-going research (e.g., [28]).

According to the expert's identifications, over 60 qualitatively different things could be identified in the thermograms. The reference menu appears in Table 2.1.

2.4.2 Experiment 2

2.4.2.1 Design and procedure

The main purpose of Experiment 2 was to investigate performance at the exact opposite end of the skill continuum relative to the expert. The participants were 48 Adelphi University undergraduates. In a post-experimental ques-

Table 2.1 Categorization for Correct Responses in the Identification Task.

Transportation
Dirt trailpath, tractor path[5], sidewalk, dirt road, two-lane street[4], four-lane street[4], boulevard (divided), expressway, expressway ramp, expressway overpass, bridge, automobile, truck, railroad car, railroad track[4], parking lot[4]
Topography
Tree[4], shrub, forest, hedge rows[5], grass field[4,5], park, hill, hillock[5], puddle[5], pond, stream, river, sand or silt river or pond bank, tilled farm field[5], till rows[5], fallow farm field[5], irrigation gully[5], crop row[5], drainage catch basins
Residential buildings
House[4], attached garage[4], unattached garage[4], porch[4], patio, trailer home[5], trailer park
Commercial buildings
Small industry[4], large industry, small commercial[4], large commercial, piping, water tank[4]
Roof and building structure
Partially insulated roof[4], poorly insulated roof[4], uninsulated roof[4], standing water on roofs[4], superstructure, composite materials[4], metal roof[4], eave, flue[4], vent, chimney[4], water damage
Thermal shine sources
Windows[4], doors, building height asymmetries[4], walls[4], roof pitch, reflecting surfaces[4], solar-load (from sidewalks, patios, etc.)
Heat phenomena
Heat shadows (e.g., from cars that had been parked during the day), compass heading (asymmetrically heated hill or hillock slopes)[5], wind direction, furnaces, heat plants, auto or truck engines, pipes[4], heat leaks, remnant heat (e.g., from car engines pausing at road intersections)[4]

Superscripts denote the figures (either 2.4 or 2.5) in which a feature can be seen.

tionnaire, only a few indicated a passing familiarity with satellite images, (i.e., from TV weather forecasts), and perhaps a vague memory of having seen one or two satellite images, typically in a high school or junior high school course in general earth science.

Aerial thermograms covering suburban areas (see Figure 2.4) manifest the regularities one might expect of aerial views of streets and houses. On the other hand, rural areas have asymmetrical ponds, forests, and farms, and meandering streams and paths. An example appears in Figure 2.5. As a consequence of physical variations (terrain texture and slope, plowing, soil moisture, etc.) and their effects on thermal emissivity, even rectangular farm

Figure 2.5 An aerial thermogram of a rural area. See color version of this figure in the color section following page 114. (Indiana State University Thermography Project. With permission.)

fields can present an asymmetrical panoply of colors and forms. We hypothesized that the uninformed viewer would be more likely to perceive the rural stimuli as abstract art rather than as aerial views.

The experiment also probed for an effect of instructions. There were four instructional conditions. In Instructional Condition 0, participants were told only that they "would be shown some pictures." In Condition 1, participants were informed that they would be shown aerial perspective pictures. In Condition 2, participants were told the pictures were aerial perspectives and they were also given a brief (350 word) explanation of thermography (i.e., that it is the detection of wavelengths outside the visible spectrum), with emphasis on the fact that the pictures do not depict the world as it appears to the human eye. The color coding scheme was not explained. In Condition 3, participants were given a fuller (500 word) description of thermography, including a discussion of the color coding scheme, and a discussion of the ways heat can manifest itself in thermograms (e.g., reflection vs. transmission of heat, thermal shine, heat leaking through poorly insulated roofs, heat coming from chimneys, etc.).

We also probed for an effect of providing an example during instruction. Easily understood "start-up" examples[153] can have a positive effect on simple and complex problem solving,[29,158,179,201] and on the acquisition of perceptual discrimination skill.[9] Practice effects can be most noticeable for participants who are given an example during instruction, especially if the example is prototypical of the test problems.[109] However, in much of the available research on this topic, it is possible for the example to *be* the instructions (i.e., *"Here's an example of the type of problem you will be seeing…"*), forcing participants to deduce the underlying scheme of the problem type (e.g., the names of mammals ordered by size). The materials and task which we were utilizing made it possible to provide examples but not explicitly engage the participants in an analytical interpretation of them. This would permit a test of the "start-up" hypothesis and would also simulate the pedagogical strategy in textbooks on remote sensing.

The final independent variable we investigated was the role of feedback, or more precisely, the way behavior might change during feedback. In the feedback conditions, during the first two trials the participants were told when a response was correct (e.g., *"Yes, that is a warm pond,"* or *"Yes, that is a street."*), and for each incorrect response the experimenter provided the correct identification (e.g., *"No, that's not a butterfly, it's an area of grassland,"* or *"No the green blob isn't a tree, it's a house"*). It is known that this type of corrective feedback can have a positive effect on the acquisition of problem solving skill.[118] Corrective feedback also can have a positive effect on category learning, concept formation, and multicue judgment, for both artificial stimuli (e.g., geometrical patterns) and for more realistic stimuli[3,61,95] including light amplification scenes.[141] (Conversely, false or delayed corrective feedback can disrupt the transition from chance-level exemplar discrimination to concept discovery and performance well above chance; see [116,129].)

Our expectation was that corrective feedback would lead to a performance improvement apart from any effect of instruction level. Especially for the suburban images, it should be easy after some feedback to generate dozens of correct identifications merely by indicating, for example, *"This is a house, this is another house, this is another house,"* and so on. Furthermore, we predicted that the effect of feedback would be modulated by stimulus type. That is, feedback with suburban stimuli might lead to performance improvement, but if stimulus type is then switched to rural, some of the gain from the prior feedback might be lost.

Participants were run individually in a small quiet room. They were shown a series of four images for five minutes per image. (The color photos were mounted in protective folders that masked the color palette and digital reference information.) During each trial, their task was to say everything they could about what they thought they saw. Their statements were audio recorded and the experimenter took notes about the image location of the things the participants identified.

2.4.2.2 Data analysis method

In addition to a response in which the participant showed an awareness of aerial perspective (e.g., *"This looks like the view from an airplane."*), the menu from the expert (see Table 2.1) was used to define what would count as a correct response.

A response could also be incorrect in a variety of ways. Based on pilot research, responses to be counted as incorrect could be unambiguously classified as: (1) mere repetition (correctly or incorrectly) of information that had been given in the instructions (e.g., *"The purple areas are colder."*), (2) default responses (e.g., *"All I see is a bunch of colors."*), (3) form responses (e.g., *"This white shape looks like a rabbit's head," "This area looks like a blob of green algae."*), (4) menu-coincident form responses (e.g., *"This white area has the shape of a tree."*), (5) incorrect identifications (*"This green blob looks like a tree."*), (6) incorrect thermal interpretations (e.g., *"The yellow area is where there is no shade."*), and (7) incorrect extra-menu responses. This last category was for features or objects that could appear in aerial thermograms, but which happened to not be present in the thermogram (e.g., *"This could be an airport—there's a row of lights,"* or *"This area is probably rocks."*).

We expected our participants would spend a considerable amount of time just inspecting the stimuli, and produce perhaps 20 responses at most. For Instructional Condition 0 and the rural stimuli, we expected more errors, more form responses, and even a failure on the part of some participants to realize that the stimuli were in aerial perspective. For Instructional Condition 3, we expected more correct and incorrect thermal interpretations than at the other instructional levels.

The sums for each participant on each trial of the number of correct and incorrect responses were determined. Analyses of variance were computed on total number of responses, number of correct responses, number of error responses, and proportion of correct responses. To determine the proportions, the number of correct responses was divided by the total number of responses for each participant on each trial.*

2.4.2.3 Results

For Instruction Level 0 there were proportionately more error (form and default) responses, and a trend for the number of errors to increase over trials. Participants in Instructional Condition 0 made many more errors (frequency range of 7–12) than participants in any of the other conditions, especially for the rural stimuli. Form and default responses predominated (frequencies on the order of 7–12 per participant) for all but one participant for whom (correct) menu responses (frequency = 14) and (incorrect) extra-menu responses (13) predominated. Participants in both Instructional Conditions 0 and 1 shown the rural stimuli made no thermal identifications—either correct or incorrect.

Given the newness of the task and materials to the participants, we expected that the pattern for the total number of responses would largely

*Further details on the experiments and the statistical analyses can be provided upon request. Although in this discussion we highlight a number of the trends in the results, we draw conclusions concerning only those trends that achieved statistical significance.

reflect the total number of errors, which it did. Thus, we focused on the proportion correct score. Only in Instructional Conditions 1, 2 and 3 did performance for some participants reach or surpass a proportion correct of 0.50, and even then most often for the suburban stimuli and in conditions in which feedback was provided. In Instructional Condition 1, there appeared fewer form responses than in Condition 0 (frequency range of 0–8 per participant), more menu responses (range of 3–20), and more extra-menu responses (range of 3–12). This trend continued across Instructional Conditions 2 and 3, where a few participants yielded no form responses. Finally, in Instructional Condition 3 there appeared a number of correct thermal identifications (typically just a few per participant but the obtained range was 0–20), and only a few incorrect thermal identifications.

We speculated that presentation of a "raw" example during instruction would have little overall effect on subsequent performance. This is a weak hypothesis, but it can be turned on its head. It is known from research on instructional design that the effectiveness of a start-up example can depend on whether there is a match or mismatch of the example and the first test problem.[29,118] We reasoned that if there could be any effect of an example it should be revealed by a comparison of two particular conditions, one in which the example type corresponded with the type of the test stimuli, and one in which it did not. Furthermore, if the suburban stimuli are more interpretable, the presentation of a suburban example during the presentation of the instructions should facilitate performance on the subsequent test trials. Conversely, the participants shown a rural example might be placed at a disadvantage.

Comparing two conditions—suburban example vs. rural example presented during the instructions, with *all* of the test stimuli being suburban—there was a general increase in proportion correct across instructional levels but only for the participants shown the suburban example, with performance uniformly surpassing proportion correct of 0.5 at Instruction Level 3. This is illustrated in Figure 2.6. Although an example might invite an interpretive process, it can also perplex it or leave it unaffected, depending on the example type and the instruction level. This is interesting because it is usually assumed that a text or manual should provide representative examples early on.

Figure 2.6 also illustrates the finding that across the instructional conditions, the proportion of correct responses tended to increase (and the number of errors tended to decrease). The greatest contrast in proportion correct, as well as in totals and errors, often resided in the comparison between Instructional Conditions 0 and 1 vs. Conditions 2 and 3. The conclusion from this seems clear—being informed that the pictures were in aerial perspective seemed to both constrain performance (fewer errors and lower total number of responses) and guide it (more menu and extra-menu identification responses). The present results show that at the earliest stages of learning a skill that is perceptually laden, basic instruction involving

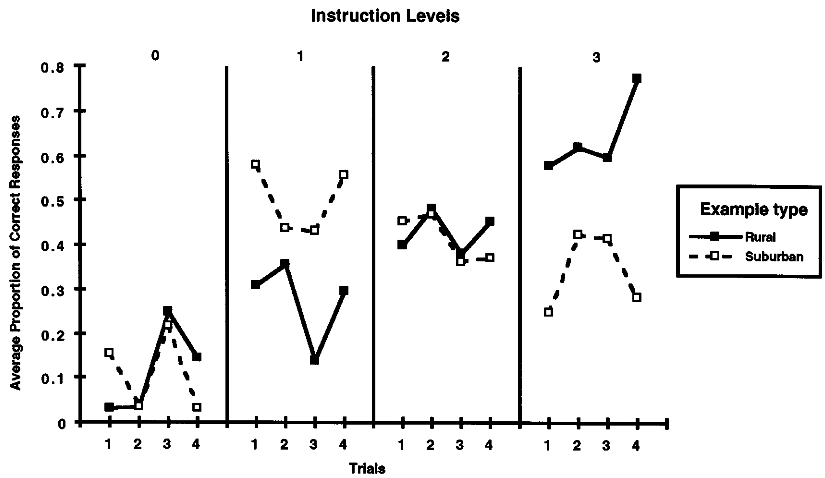

Figure 2.6 Results showing average proportion of correct responses as a function of example type, instruction level, and trials.

even a single critical concept can have a considerable and positive effect on performance.

On the other hand, the results for Instructional Conditions 1, 2, and 3 also suggest that initial instructional information can actually inhibit overall responding. Higher levels of instruction combined with the less interpretable stimuli (i.e., the rural images) seem to have inhibited overall performance, whereas at lower instruction levels the less interpretable stimuli generated the most errors. Although in Instructional Condition 3 there appeared more thermal identifications, there was a clear decrease in total number of responses and the numbers of (correct) menu identifications and (incorrect) extra-menu identifications.

We had hypothesized that the more elaborate instructions could induce a cognitive load, that is, overwhelm participants with "front-end" verbal information. While this could explain the findings, the effect we obtained could also be due to simple inhibition. That is, participants might have been less inclined to speak up having been given elaborate instructions. This interpretation is supported by the finding that across instructional conditions there was a steady increase in the number of default responses—in which the participants expressed frustration or said, *"I don't see anything."* In Instructional Condition 3, every participant made at least one such response.

Further research could attempt to isolate the cognitive load factor from response inhibition. This would be important because it is never desirable for initial instructional information to inhibit responding—that is precisely the time when the instructor needs to know that the learner is thinking.

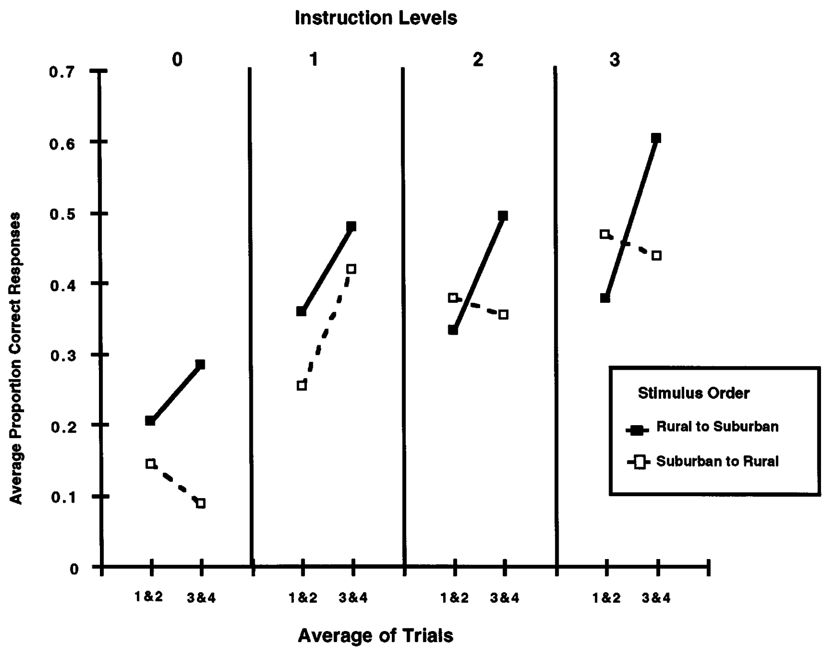

Figure 2.7 Results showing average proportion of correct responses averaging trials 1 and 2 and trials 3 and 4, as a function of stimulus order and instruction level.

The type of test stimulus did affect performance, with the stimulus difference defined here as one involving perceptual salience and interpretive significance (suburban vs. rural). For some of the conditions we ran, the first two test stimuli were rural and the second two were suburban, or vice versa. While performance generally improved across instructional conditions, the switching of stimulus type could be disruptive. Specifically, switching from suburban to rural did not significantly change performance, but switching from rural to suburban resulted in a consistent albeit modest improvement in performance. Overlaid upon this is the performance improvement across instruction levels. These results are shown in Figure 2.7.

Looking across the conditions we ran, knowing that the stimuli are in aerial perspective (Instructional Conditions 1, 2, and 3) seems to have helped, but for the rural stimuli—hypothetically less interpretable—performance showed little or no improvement across instruction levels (when no feedback was provided). Although some participants in Instructional Conditions 2 and 3 made some correct thermal identifications, participants at Instructional Levels 0 and 1 shown the rural stimuli made no thermal identifications—either correct or incorrect. In Instructional Conditions 1, 2, and 3, all of which mentioned that the pictures were in aerial perspective, the preponderance of erroneous form responses seen at Instructional Level 0 gave way to a preponderance of menu responses (both correct menu responses and incorrect

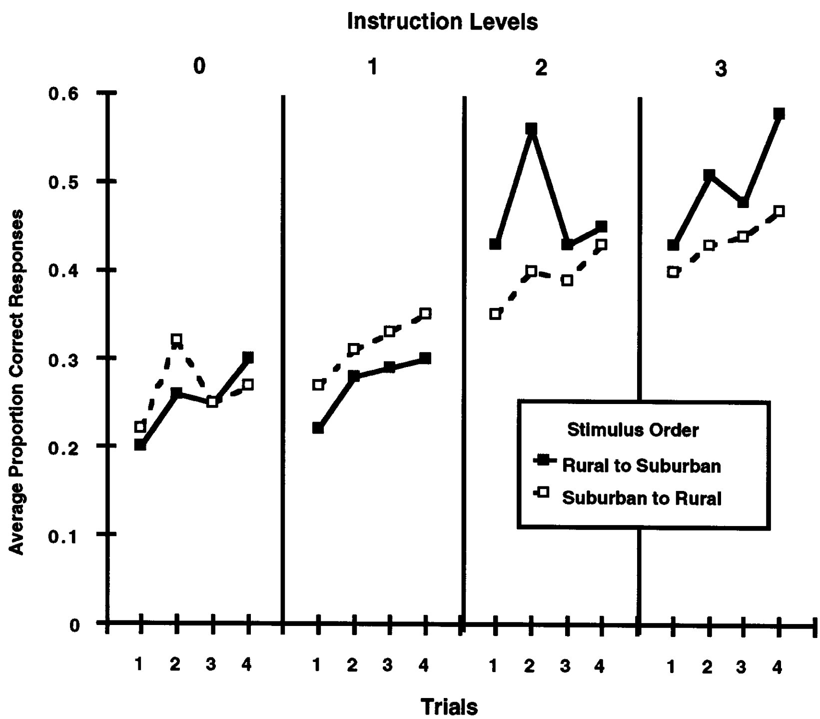

Figure 2.8 Results showing average proportion correct as a function of stimulus order, instruction level, and trials, for the conditions in which participants received feedback.

extra-menu responses). Knowing that the stimuli depict aerial perspective helped, but even with that knowledge, performance in terms of proportion correct did not consistently improve across instruction levels. In other words, when participants were confronted with the actual interpretation task, some stimuli were easier to interpret incorrectly.

Our expectation was that corrective feedback would lead to a performance improvement apart from the effect of instruction level, and also result in a clear effect of trials. Statistical testing confirmed these expectations (see footnote on page 31). Results for the feedback conditions are presented in Figure 2.8. Apparent in Figure 2.8 is the improvement in performance comparing instruction levels, and also the general improvement across trials within each of the instructional conditions.* While feedback had the effect of

*Notice also the "breakpoints" in Instructional Conditions 0, 2 and 3 between trials 2 and 3 in the rural-to-suburban stimulus order conditions, reflecting the effect of switching from rural-to-suburban stimuli. The effect of the switching of stimulus type seems to have been occasional and modest compared to the effects of trials and instruction level in the conditions in which feedback was provided.

increasing the total number of responses (means of 14.32 vs. 7.31), it also increased the number of errors (participant means of 9.41 vs. 4.98). In the conditions in which no feedback was provided, the number of errors decreased only slightly from Instruction Levels 0 to 1, whereas in the feedback conditions, number of errors more clearly decreased across instruction levels.

2.4.3 Implications for education in remote sensing

The study reported here was designed to merely demonstrate the viability of psychological research into the human factors of remote sensing image interpretation. The findings are entirely preliminary, and are based on only a sampling of the important variables. In addition, our small sample size in terms of the number of participants in our experimental conditions allowed individual differences to percolate through and cloud the interpretation of some of the statistical interactions we obtained. Nevertheless, the results dovetail with other research on the effects of examples, instructions, and feedback (e.g., [29,118]).

For example, the results suggest that illustrative introductory examples should be used in texts in a particular way:

- Introductory exercises should be engaged *before* the student is burdened with the details of electromagnetism, sensor systems, optics, thermodynamics, etc. The examples should follow only a very brief explanation of the electromagnetic spectrum.
- The introductory examples should be in sufficient number to illustrate the diversity and variety of remote sensing imagery. Some research on learning has shown that variability as well as prototypicality of practice items can enhance transfer (e.g., [11,12]). The presentation of a range of examples—covering the critical attributes, salient feature contrasts, degrees of difficulty, etc.—is especially conducive of initial concept learning.[8,9,116,179]
- Every example should be accompanied by specific, step-wise guidance allowing the student to conduct an interpretation and engage his or her perceptual and reasoning processes.
- That interpretation task should *not* be the mere “pointing out” of some particular salient feature (e.g., *”In this aerial photograph you can see the fracture line indicating faulted bedrock.”*). Rather, the task should be more engaging (e.g., *”Can you explain the hot swath across the playing field?”*).
- Some introductory exercises should be conducted with an instructor who would provide immediate, explanatory, elaborative feedback. Those exercises should be conducted so as to encourage and facilitate the initiate’s verbalization of his or her reasoning, without fear of error or criticism.

A number of modern theories of perceptual learning and categorization emphasize the importance of concept coherence and the perception of family resemblances or prototypes (e.g., [136,157]). An important question that stems from such theories is: what is the nature of the information presented by remote sensing images?[82,93] In the suburban thermograms there seems to be sufficient information to support the perception of aerial perspective and the identification of certain features (e.g., streets, houses, trees, etc.). Not all of the responses in Instructional Condition 0 were form responses—some responses showed that the participants were able to discover the aerial perspective on their own. But expert performance shows that there is also information sufficient to indicate such things as compass orientation, effects of solar heating, the internal structure of roofs, underground steam pipes, and so on. The instructions for Level 3 involved elaborating the notion of thermography, and so we expected that at Level 3 we would obtain more thermal identification responses than we did. The paucity of *incorrect* thermal identifications suggests that once the color coding is explained, some of the meaning in the thermograms become discernable—the color coding scheme is somewhat effective in informing about thermal properties.

2.5 Conclusions and prospects

We submit that there are two important take-home messages from this exploration of the psychological "angle of regard" on remote sensing.

1. A number of psychological processes are, and will remain, critical to the enterprise of remote sensing. Foremost among these are:
   - the acquisition of a rich and highly organized knowledge base that supports the formation of mental models, and
   - the acquisition of perceptual skills.

The achievement of expertise relies on fundamental cognitive processes such as concept formation, but it also relies upon judgmental processes, hypothetical reasoning, aesthetic analysis, feedback, and so on. The last two decades have witnessed a significant migration of psychological research away from the traditional artificial laboratory into applied domains (see [90,91]) including the study of expertise (see [31,55,88]), and we look forward to much more research on the psychological foundations of remote sensing.

2. Psychological research in collaboration with remote sensing scientists can lead to new advances and approaches.

   These would include new approaches to instructional design and training, in order to accelerate the progression from trainee to expert, and new methods of visualization and display design—new "cognitive technologies"— to support and amplify the performance of the remote sensing scientist. For all forms of nonliteral imagery, we are hopeful that further specification of the information that is available, through psychological research, would permit the design of displays and instructional and feedback conditions that would facilitate the perceptual learning and interpretation processes and also contribute to efforts at automation.

References

1. Anderson, J. R., Skill acquisition: compilation of weak-method problem solutions, *Psychol. Rev.*, 94, 192–210, 1987.
2. Artis, D. A. and Carnahan, W. H., Survey of emissivity variability in thermography of urban areas, *Remote Sensing Environ.*, 12, 313–329, 1982.
3. Balzer, W. K., Doherty, M. E., and O'Connor, R., Effects of cognitive feedback on performance, *Psychol. Bull.,* 106, 410–433, 1989.
4. Bauer, M. I. and Johnson-Laird, P. N., How diagrams can improve reasoning, *Psychol. Sci.,* 4, 372–378, 1993.
5. Bennett, K. B. and Flach, J. M., Graphical displays: implications for divided attention, focused attention, and problem solving, *Hum. Factors,* 34, 513–533, 1992.
6. Berry, D. C. and Broadbent, D. E., On the relationship between task performance and associated verbalizable knowledge, *Q. J. Exp. Psychol.,* 36A, 209–231, 1984.
7. Bertin, J., *The Semiology of Graphics,* University of Wisconsin Press, Madison, 1983.
8. Biederman, I., Human image understanding: recent research and a theory, *Comput. Vision, Graphics, Image Proc.,* 32, 29–73, 1987.
9. Biederman, I. and Shiffrar, M. M., Sexing day-old chicks: a case study and expert systems analysis of a difficult perceptual-learning task, *J. Exp. Psychol. Learning, Mem. Cognit.,* 13, 640–645, 1987.
10. Booth, W., Monitoring the fate of forests from space, *Science,* 243, 1428–1429, 1989.
11. Bransford, J. D., Franks, J. J., Morris, C. D., and Stein, B. S., Some general constraints on learning and memory research, in *Levels of Processing in Human Memory,* Cermak, L. S., and Craik, F. I. M., Eds., Wiley, New York, 1979, 331–354.
12. Bransford, J. D., Franks, J. J., Vye, N. J., and Sherwood, R. D., New approaches to instruction: because wisdom can't be told, in *Similarity and Analogical Reasoning,* Vosniadou, S., and Ortony, A., Eds., Cambridge, Cambridge University Press, 1989, 470–497.
13. Broadbent, D. E. and Aston, B., Human control of a simulated economic system, *Ergonomics,* 21, 1043–1053, 1978.
14. Brooks, L., Nonanalytic concept formation and memory for instances, in *Cognition and Categorization,* Rosch, R., and Lloyd, B., Eds., Erlbaum, Hillsdale, NJ, 1978, 169–211.
15. Bruner, J. S., *Going Beyond the Information Given,* Harvard University Press, Cambridge, MA, 1957.
16. Bruner, J. S., Goodnow, J. J., and Austin, A. A., *A Study of Thinking,* Wiley, New York, 1956.
17. Burns, B. C., Ed., *Percepts, Concepts and Categories,* Elsevier Science Publishers, Amsterdam, 1992.
18. Buttenfield, B. P. and MacKaness, W. A., Visualization, in *Geographical Information Systems: Principles and Applications, Vol. 1,* Maguire, D. J., Goodchild, M. F., and Rhind, D. W., Eds., John Wiley, New York, 1991, 427–4430.

19. Buttigieg, M. A. and Sanderson, P. M., Emergent features in visual displays for two types of failure detection tasks, *Hum. Factors,* 33, 631–651, 1991.
20. Byrne, R. W., Memory for urban geography, *Q. J. Exp. Psychol.,* 31, 147–154, 1979.
21. Camerer, C. F. and Johnson, E. J., The process-performance paradox in expert judgment: how can experts know so much and predict so badly? in *Toward a General Theory of Expertise: Prospects and Limits,* Ericsson, K. A. and Smith, J., Eds., Cambridge University Press, New York, 1991, 195–217.
22. Chandrasekaran, B., Narayanan, N. H., and Iwasaki, Y., Reasoning with diagrammatic representations: a report on the Spring symposium, *The AI Magazine* (1993, Summer) 49–56.
23. Callaghan, T. C., Commentary, in *Percepts, Concepts and Categories,* Burns, B. C., Ed., Elsevier Science Publishers, Amsterdam, 1992, 492–493.
24. Campbell, J. B., *Introduction to Remote Sensing,* Guilford, New York, 1996.
25. Card, S. K., Visual search of computer command menus, paper presented at the Tenth Int. Symp. on Attention and Performance, Venlo, Netherlands, July 1982.
26. Carlson, R. S., Sullivan, M. A., and Schneider, W., Practice and working memory effects in building procedural skill, *J. Exp. Psychol. Learning, Mem., Cognit.,* 15, 517–526, 1989.
27. Carmody, D. P., Kundel, H. L., and Toto, L. C., Comparison scans while reading chest images: taught but not practiced, *Investigative Radiol.,* 19, 462–466, 1989.
28. Carnahan, W. H. and Larson, R. C., An analysis of an urban heat sink, *Remote Sensing Environ.,* 33, 65–71, 1990.
29. Catrambone, R. and Wachman, R. M., The interaction of principles and examples in instruction, in *Proc. 14th Annu. Meet. Cognit. Sci. Soc.,* Erlbaum, Mahwah, NJ, 1992, 749–754.
30. Chase, W. G. and Simon, H. A., Perception in chess, *Cognit. Psychol.,* 5, 55–81, 1973.
31. Chi, M. T. H., Glaser, R., and Farr, M. J., Eds., *The Nature of Expertise,* Erlbaum, Mahwah, NJ, 1988.
32. Childs, C. M., Interpreting composite imagery requires training, *Tech-Tran,* Newsletter of the Engineer Topographic Laboratories, U.S. Army Corps of Engineers, 13, Fort Belvoir, VA, 3, 1988.
33. Christ, R. E. and Corso, G. M., The effects of extended practice on the evaluation of visual display codes, *Hum. Factors,* 25, 71–84, 1983.
34. Christner, C. A. and Ray, H. W., An evaluation of the effect of selected combinations of target and background coding on map reading performance, *Hum. Factors,* 3, 131–146, 1961.
35. Clark, J. D., Ed., The GOES User's Guide, National Environmental Satellite Data and Information Service, National Oceanic and Atmospheric Administration, Washington, D.C., 1983.
36. Cleermans, A. and McClelland, J. L., Learning the structure of event sequences, *J. Exp. Psychol. Gen.,* 120, 235–253, 1991.
37. Colwell, R. N., Ed., *Manual of Remote Sensing,* American Society for Photogrammetry and Remote Sensing, Falls Church, VA, 1983.
38. Crandall, B. W., A comparative study of think out loud and criticial decision knowledge elicitation methods, in Special Issue on Knowledge Acquisition, *SIGART Newsletter,* 108, Special Interest Group on Artificial Intelligence, Association for Computing Machinery, New York, 1989, April, 144–146.

39. Crandall, B. and Calderwood, R., Clinical assessment skills of experienced neonatal intensive care nurses, Report, Klein Associates, Yellow Springs OH, 1989.
40. Crandall, B. and Getchell-Reiter, K., Critical Decision Method: a technique for eliciting concrete assessment indicators from the intuition of NICU nurses, *Adv. Nursing Sci.*, 16, 42–51, 1993.
41. Crossman, E. R. F. W., Cooke, J. E., and Beishon, R. J., Visual attention and the sampling of displayed information in process control, in *The Human Operator in Process Control*, Edwards, E. and Lees, F. P., Eds., Taylor & Francis, London, 1974.
42. Curran, P. J., Remote sensing methodologies and geography, *Int. J. Remote Sensing*, 8, 1255–1275, 1987.
43. Davies, D., Bathurst, D., and Bathurst, R., *The Telling Image: The Changing Balance between Pictures and Words in a Technological Age*, Oxford University Press, Oxford, 1990.
44. De Keyser, V., How can computer-based visual displays aid operators? *Int. J. Man-Machine Sys.*, 27, 471–478, 1987.
45. DiBiase, D., MacEachren, A. M., Krygier, J. B., and Reeves, C., Animation and the role of map design in scientific visualization, *Cartography and Cartographic Inf. Sys.*, 19, 201–214; 256–266, 1992.
46. Dreyfus, H. L. and Dreyfus, S. E., *Mind over Machine: The Power of Human Intuition and Expertise in the Era of the Computer*, Free Press, New York, 1986.
47. Durrett, H. J., Ed., *Color and the Computer*, Academic Press, New York, 1987.
48. Durso, F. T., Cooke, N. M., Breen, T. J., and Schvaneveldt, R. W., Is consistent mapping necessary for high speed search? *J. Exp. Psychol. Learning, Mem. Cognit.*, 13, 232–229, 1987.
49. Eberhart, J., The night skies of Venus: another kind of aurora? *Sci. News*, 130, 364–365, 1988.
50. Edlund, C. and Lewis, M., Comparing ecologically constrained and conventional displays in control of a simple steam plant, in *Proc. Hum. Factors and Ergonomics Soc. 38th Annu. Meet.*, The Human Factors and Ergonomics Society, 1994, 486–490.
51. Eley, M. G., Determining the shapes of land surfaces from topographical maps, *Ergonomics*, 31, 355–376, 1988.
52. Endsley, M. R., Toward a theory of situation awareness in dynamic systems, *Hum. Factors*, 37, 32–64, 1995.
53. EOSAT, Mapping New Guinea using LANDSAT and dugout canoes, *LANDSAT Data Users Notes*, 5, 3–4, 1990.
54. Epstein, W., *Varieties of Perceptual Learning*, McGraw Hill, New York, 1967.
55. Ericsson, K. A. and Smith, J., Eds., *Toward a General Theory of Expertise: Prospects and Limits*, Cambridge University Press, New York, 1991.
56. Essock, E. A., Sinai, M. J., McCarley, J. S., Krebs, W. K., and DeFord, J. K., Perceptual ability with real-world nighttime scenes: image-intensified, infrared, and fused color imagery, *Hum. Factors*, 41, 438–452, 1999.
57. Estes, W. K., Of models and men, *Am. Psychol.*, 12, 609–617, 1957.
58. Estes, W. K., *Classification and Cognition*, Oxford University Press, Oxford, 1994.
59. Evans, G. W. and Pezdek, K., Cognitive mapping: knowledge of real-world distance and location information. *J. Exp. Psychol. Hum. Learning Mem.*, 6, 13–24, 1980.

60. Fayyad, U. M., Djorgovski, S. G., and Weir, N., From digitized images to online catalogs: data mining a sky survey, *AI Mag.* (1996, Summer) 51–66.
61. Fischhoff, B. and Slovic, P., A little learning. . . Confidence in multicue judgment tasks, in *Attention and Performance VIII,* Nickerson, R. S., Ed., Erlbaum, Mahwah, NJ, 1980, 770–801.
62. Fisk, A. D., High performance cognitive skill acquisition: perceptual / rule learning, in *Proc. 31st Annu. Meet. Hum. Factors Soc.,* Human Factors Society, Santa Monica, CA, 1987, 652–656.
63. Flach, J. M. and Grunzke, P. M., Automatic processing through the back door, in *Proc. 31st Annu. Meet. Hum. Factors Soc.,* Human Factors Society, Santa Monica, CA, 1987, 1054–1056.
64. Fleming, M. L., Designing pictorial / verbal instruction: some speculative extensions from research to practice, in *The Psychology of Illustration, Vol. 2: Instructional Issues,* Houghton, H. A. and Willows, D. M., Eds., Springer Verlag, New York, 1987, 136–157.
65. Foyle, D. C. and Kaiser, M. K., Pilot distance estimation with unaided vision, night-vision goggles and infrared imagery, *Digest of Technical Papers of the International Meeting of the Society for Information Display,* Society for Information Display, Playa del Rey, CA, 1991, 314–317.
66. Friedhoff, B. W., *Visualization: The Second Computer Revolution,* W. H. Freeman, San Francisco, 1991.
67. Friedl, M. A., Estes, J. E., and Star, J. L., Advanced information extraction tools in remote sensing for earth science applications: AI and GIS, *AI Appl. Nat. Resour. Manage.*, 2, 17–30, 1988.
68. Fussell, J., Rundquist, D., and Harrington, J. A., On defining remote sensing, *Photogrammetric Eng. Remote Sensing,* 52, 1057–1511, 1986.
69. Getty, D. J., Pickett, R. M., D'Orsi, C. J., and Swets, J. A., Enhanced interpretation of diagnostic images, *Invest. Radiol.,* 23, 240–252, 1988.
70. Gibson, E. J., *Principles of Perceptual Learning and Development,* Prentice-Hall, Englewood Cliffs, NJ, 1969.
71. Glaser, R., Thoughts on expertise, in *Cognitive Functioning and Social Structure over the Life Course,* Schooler, C. and Schaie, W., Eds., Ablex, Norwood, NJ, 1987, 81–94.
72. Graeser, A. C. and Clark, L. F., *Structures and Procedures for Implicit Knowledge,* Ablex, Norwood, NJ, 1985.
73. Greeno, J. G., Situations, mental models, and generative knowledge, in *Complex Information Processing: The Impact of Herbert A. Simon,* Klahr, D. and Kotovsky, K., Eds., Erlbaum, Hillsdale, NJ, 1989, 285–318.
74. Hall, S., *Mapping the Next Millennium*, Random House, New York, 1991.
75. Hammond, K. R., Clinical inference in nursing, II: a psychologist's viewpoint, *Nursing Res.,* 15, 27–38, 1966.
76. Hammond, K. R., Naturalistic decision making from a Brunswikian viewpoint: its past, present, and future, in *Decision Making in Action: Models and Methods,* Klein, G., Orasanu, J., Calderwood, R. and C. E. Zsambok, C. E., Eds. Ablex, Norwood, NJ, 1993, 205–227.
77. Hansen, C., Personal communication, Jet Propulsion Laboratory, Pasadena, CA, 1987, May.
78. Hartman, M., Knopman, D. S., and Nissen, M. J., Implicit learning of new verbal associations, *J. Exp. Psychol. Learning, Mem. Cognit.,* 15, 1070–1082, 1989.

79. Hayes, N. A. and Broadbent, D. E., Two modes of learning for interactive tasks, *Cognition*, 28, 249–276, 1988.
80. Hess, S. M., Detweiler, M. C., and Ellis, R. D., The effects of display layout on monitoring and updating system states, in *Proc. 38th Annu. Meet. Hum. Factors and Ergonomics Soc.*, Human Factors and Ergonomics Society, Santa Monica, CA, 1994, 1336–1340.
81. Hintzman, D. L., O'Dell, C. S., and Arndt, D. R., Orientation in cognitive maps, *Cognit. Psychol.* 11, 375–394, 1981.
82. Hodgson, M. E., What size window for image classification? A cognitive perspective, *Photogrammetric Engin. Remote Sensing*, 64, 797–807, 1998.
83. Hodgson, M. E. and Lloyd, R. E., Cognitive and statistical approaches to texture, *ASPRS Annu. Convention Tech. Papers*, 4, 407–416, 1986.
84. Hoffman, R. R., The problem of extracting the knowledge of experts from the perspective of experimental psychology, *AI Mag.* (1987, Summer) 53–67.
85. Hoffman, R. R., Remote perceiving: a step toward a unified science of remote sensing, *Geocarto Int. Multidisciplinary J. Remote Sensing*, 5, 3–13, 1990.
86. Hoffman, R. R., Human factors psychology in the support of forecasting: the design of advanced meteorological workstations, *Weather Forecasting*, 6, 98–110, 1991.
87. Hoffman, R. R., Doing psychology in an AI context, in *The Psychology of Expertise: Cognitive Research and Empirical AI,* Hoffman, R. R., Ed., Erlbaum, Mahwah, NJ, 1992, 3–11.
88. Hoffman, R. R. Ed., *The Psychology of Expertise: Cognitive Research and Empirical AI*, Erlbaum, Mahwah, NJ, 1992.
89. Hoffman, R. R. and Conway, J. A., Psychological factors in remote sensing: a review of some recent research, *GEOCARTO Int.*, 4, 3–21, 1989.
90. Hoffman, R. R. and Deffenbacher, K. A., An analysis of the relations of basic and applied science, *Ecological Psychol.*, 5, 315–352, 1993.
91. Hoffman, R. R. and Deffenbacher, K., A brief history of applied cognitive psychology, *Appl. Cognit. Psychol.*, 6, 1–48, 1992.
92. Hoffman, R. R., Detweiler, M., Conway, J. A., and Lipton, K., Some considerations in the use of color in meteorological displays, *Weather Forecasting*, 8, 505–518, 1993.
93. Hoffman, R. R. and Pike, R. J., On the specification of the information available for the perception and description of the natural terrain, in *Local Applications of the Ecological Approach to Human-Machine Systems,* Hancock, P., Flach, J., Caird, J., and Vicente, K., Eds., Erlbaum, Mahwah, NJ, 1995, 285–323.
94. Holz, R. K., Ed., *The Surveillant Science: Remote Sensing of the Environment,* 2nd. ed., Wiley, New York, 1985.
95. Homa, D. and Cultice, J., Role of feedback, category size, and stimulus distortion on the acquisition and utilization of ill-defined categories, *J. Exp. Psychol. Learning, Mem., Cognit.*, 10, 83–94, 1984.
96. Houghton, H. A. and Willows, D. M., Eds., *The Psychology of Illustration, Vol. 2: Instructional Issues,* Springer Verlag, New York, 1987.
97. Jacoby, L. L. and Dallas, M., On the relation between autobiographical memory and perceptual learning, *J. Exp. Psychol. Gen.*, 110, 306–346, 1981.
98. Jenkins, J. J., Acoustic information for objects, places and events, in *Persistence and Change: Proceedings of the First International Conference on Event Perception,* Warren, W. H. and Shaw, R. E., Eds., Erlbaum, Hillsdale, NJ, 1985, 115–138.

99. Kalyuga, S., Chandler, P., and Sweller, J., Levels of expertise and instructional design, *Hum. Factors,* 40, 1–17, 1998.
100. Kibbe, M. P. and Stiff, J., Operator performance in pattern matching as a function of reference material structure, in *Proc. 38th Annu. Meet. Hum. Factors and Ergonomics Soc.,* The Human Factors and Ergonomics Society, Santa Monica, CA, 1994, 1280–1284.
101. Klein, G. A. and Hoffman R. R., Perceptual-cognitive aspects of expertise, in *Cognitive Science Foundations of Instruction,* Rabinowitz, M., Ed., Erlbaum, Hillsdale, NJ, 1993, 203–226.
102. Klein, G. and MacGregor, D., Knowledge elicitation of recognition-primed decision making, Report MAD903-86-C-0170, U.S. Army Research Institute, Alexandria, VA, 1987.
103. Kolers, P. A., Reading a year later, *J. Exp. Psychol. Hum. Learning Mem.,* 2, 554–565, 1976.
104. Koubek, R. J., Clarkston, T. P., and Calvez, V., The training of knowledge structures for manufacturing tasks: an empirical study, *Hum. Factors,* 37, 765–780, 1994.
105. Lahav, O., Naim, A., Buta, R. J., Corwin, H. G., de Vaucouleurs, G., Dressler, A., Huchra, J. P., van den Bergh, S., Raychaudhury, S., Sodré, L., and Storrie-Lompardi, M. C., Galaxies, human eyes, and artificial neural networks, *Science,* 267, 959–962, 1995.
106. Larkin, J. H., The role of problem representation in physics, in *Mental Models,* Gentner, D. and Stevens, A., Eds., Erlbaum, Hillsdale, NJ, 1983, 75–98.
107. Larkin, J. H., Display-based problem solving, in *Complex Information Processing: The Impact of Herbert A. Simon,* Klahr, D. and Kotovsky, K., Eds., Erlbaum, Hillsdale, NJ, 1993, 319–341.
108. Larkin, J. H. and Simon, H. A., Why a diagram is (sometimes) worth ten thousand words, in *Models of Thought, Vol. 2,* Simon, H. A., Ed., Yale University Press, New Haven, CT, 1989, 413–437.
109. LeFevre, J. A. and Dixon, P., Do written instructions need examples? *Cognit. Instruction,* 3, 1–30, 1986.
110. Lesgold, A. M., Acquiring expertise, in *Tutorials in Learning and Memory: Essays in Honor of Gordon Bower,* Anderson, J. R. and Kosslyn, S. M., Eds., W. H. Freeman, San Francisco, 1984, 31–60.
111. Lesgold, A. M., Rubinson, H., Feltovich, P., Glaser, R., Klopfer, D., and Wang, Y., Expertise in a complex skill: diagnosing X-ray pictures, in *The Nature of Expertise*, Chi, M. T. H., Glaser, R., and Farr, M. J., Eds., Erlbaum, Mahwah, NJ, 1988, 311–342.
112. Levin, J. R. and Mayer, R. E., Understanding illustrations in text, in *Learning from Textbooks: Theory and Practice,* Britton, B. K., Woodward, A., and Binkley, M., Eds., Erlbaum, Hillsdale, NJ, 1993, 95–113.
113. Lewicki, P., Processing information about covariance that cannot be articulated, *J. Exp. Psychol. Learning, Mem. Cognit.*, 12, 135–146, 1986.
114. Lewicki, P., Czyzewska, M., and Hoffman, H., Unconscious acquisition of complex procedural knowledge, *J. Exp. Psychol. Learning, Mem. Cognit.*, 13, 523–530, 1987.
115. Lewicki, P., Hill, T., and Bizot, E., Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated, *Cognit. Psychol.,* 20, 24–37, 1988.

116. Lewis, M. W. and Anderson, J. R., Discovery of operator schemata in problem solving: learning from examples, *Cognit. Psychol.*, 17, 26–65, 1985.
117. Lillesand, T. M. and Kiefer, R. W., *Remote Sensing and Image Interpretation,* Wiley, New York, 1979.
118. Lovett, M. C., Learning by problem solving versus by examples: the benefits of generating and receiving information, in *Proc. 14th Annu. Meet. Cognit. Sci. Soc.,* Erlbaum, Mahwah, NJ, 1992, 956–961.
119. Lowe, R., Constructing a mental representation from an abstract technical diagram, *Learning Instruction,* 3, 157–179, 1993.
120. Luney, P. L. and Dill, H. W., *Remote Sensing with Special Reference to Agriculture and Forestry,* National Academy of Sciences, Washington, D.C., 1970.
121. Lynch, M., Laboratory space and the technological complex: an investigation of topical contextures, *Sci. Context,* 4, 51–78, 1991.
122. Lynch, M. and Edgerton, S. Y., Aesthetics and digital image processing: representational craft in contemporary astronomy, in *Picturing Power: Visual Depictions and Social Relations,* Fyfe, G. and Law, J., Eds., Routledge, London, 1988, 184–220.
123. MacEachren, A. M., Buttenfield, B. P., Campbell, J. B., DiBiase, D. W., and Monmonier, M., Visualization, in *Geography's Inner Worlds: Pervasive Themes in Contemporary American Geography,* Abler, R. F., Marcus, M. G., and Olson, J. M., Eds., Rutgers University Press, New Brunswick, NJ, 1992, 99–137.
124. MacEachren, A. M. and Ganter, J. H., A pattern identification approach to cartographic visualization, *Cartographica,* 27, 64–81, 1990.
125. MacEachren, A. M. and Taylor, D. R. F. Eds., *Visualization in Modern Cartography,* Elsevier, Tarrytown, NY, 1994.
126. Mandl, H. and Levin, J. R., Eds., *Knowledge Acquisition from Text and Pictures,* North-Holland, Amsterdam, 1989.
127. Markman, A. B. and Makin, V. S., Referential communication and category acquisition, *J. Exp. Psychol. Gen.,* 127(4), 331–354, 1998.
128. Markman, A. B., Yamauchi, T., and Makin, V. S., The creation of new concepts: a multifaceted approach to category learning, in *Creative Thought: An Investigation of Conceptual Structures and Processes,* Ward, T. B., Smith, S. M., and Vaid, J., Eds., American Psychological Association, Washington, D.C., 1997, 179–208.
129. Mathews, R. C., Buss, R. R., Chinn, R., and Stanley, W. B., The role of explicit and implicit learning processes in concept discovery, *Q. J. Exp. Psychol.,* 40A, 135–165, 1988.
130. Mayer, R. E. and Anderson, R. B., Animations need narrations: an experimental test of the dual-coding hypothesis, *J. Educ. Psychol.,* 83, 484–490, 1991.
131. Mayer, R. E. and Gallini, J. K., When is an illustration worth ten thousand words? *J. Educ. Psychol.,* 82, 715–726, 1990.
132. Mayer, R. E. and Greeno, J. G., Structural differences between learning outcomes produced by different instructional methods, *J. Edu. Psychol.,* 63, 165–173, 1972.
133. McGuinness, C., Expert/novice use of visualization tools, in *Visualization in Modern Cartography,* MacEachren, A. M. and Taylor, D. R. F., Eds., Elsevier, Tarrytown, NY, 1994, 185–199.
134. Medin, D. L. and Smith, E. E., Strategies and classification learning, *J. Exp. Psychol. Hum. Learning Mem.,* 7, 29–37, 1981.

135. Moray, N. and Richards, M., Memory for radar-like patterns, Report EPRL-90-03, Engineering Psychology Research Laboratory, Department of Mechanical and Industrial Engineering, University of Illinois, Urbana-Champaign, 1990.
136. Murphy, G. L. and Medin, D. L., The role of theories in conceptual coherence, *Psychol. Rev.*, 92, 289–316, 1985.
137. Myles-Worsley, M., Johnston, W. A., and Simons, M. A., The influence of expertise on X-ray image processing, *J. Exp. Psychol. Learning, Mem. Cognit.*, 14, 553–557, 1988.
138. NASA, NASA color thermographs show how much heat homeowners are losing, Flyer produced by the National Aeronautics and Space Administration, Washington, D.C., 1977.
139. Nelson, I. and Lee, J., Conversations with graphics: implications for the design of natural language / graphics interfaces, *Int. J. Hum.-Comput. Stud.*, 40, 509–541, 1994.
140. Neisser, U. and Hirsch, N., Phantom flashbulbs: false recollections of hearing the news about Challenger, in *Affect and Accuracy in Recall: Studies of "Flashbulb Memories,"* Winograd, E. and Neisser, U., Eds., Cambridge University Press, New York, 1992, 9–31.
141. Niall, K. K., Reising, J. D., and Martin, E. L., Distance estimation with night vision goggles: a little feedback goes a long way, *Hum. Factors,* 41, 495–506, 1999.
142. Nissen, M. J. and Bullemer, P., Attentional requirements of learning: evidence from performance measures, *Cognit. Psychol.*, 19, 1–32, 1987.
143. Norman, G. R., Rosenthal, D., Brooks, L. R., Allen, S. W., and Muzzin, L. J., The development of expertise in dermatology, *Arch. Dermatol.*, 125, 1063–1068, 1989.
144. Nosofsky, R. M., Palmeri, T. J., and McKinley, S. C., Rule-plus-exception model of classification learning, *Psychol. Rev.*, 101, 53–97, 1994.
145. Olson, J. M., Color and the computer in cartography, in *Color and the Computer,* Durrett, J. H., Ed., Academic Press, New York, 1987, 205–219.
146. Oransky, N. A., Skedsvold, P. R., and Fisk, A. D., The interaction of bottom-up and top-down consistency in the development of skills, in *Proc. 31st Annu. Meet. Hum. Factors Soc.*, The Human Factors Society, Santa Monica, CA, 1987, 1044–1048.
147. Peeck, J., The role of illustrations in processing and remembering text, in *The Psychology of Illustration, Vol. 1: Basic Research,* Willows, D. M. and Houghton, H. A., Eds., Springer Verlag, New York, 1987, 115–151.
148. Perruchet, P., Galego, J., and Savy, L., A critical reappraisal of the evidence for unconscious abstraction of deterministic rules in complex experimental situations, *Cognit. Psychol.*, 22, 493–516, 1990.
149. Reber, A. S., Implicit learning of synthetic languages: the role of instructional set, *J. Exp. Psychol. Hum. Learning Mem.*, 2, 88–94, 1976.
150. Reber, A. S., Implicit learning and tacit knowledge, *J. Exp. Psychol. Gen.*, 118, 219–235, 1989.
151. Reber, A. S., Cassin, S. M., Lewis, S., and Cantor, G., On the relation between implicit and explicit modes of deriving a complex rule structure, *J. Exp. Psychol. Hum. Learning Mem.*, 6, 492–502, 1980.

152. Reber, A. S. and Milward, R. B., Probability learning and memory for event sequences, *Psychonomic Sci.*, 3, 431–432, 1968.
153. Rissland, E. L., Examples and learning systems, in *Adaptive Control of Ill-Defined Systems,* Selfridge, O. G., Rissland, E. L. and Arbib, M. A., Eds., Plenum, New York, 1981, 149–163.
154. Rogowitz, B. E. and Treinish, L. A., An architecture for perceptual rule-based visualization, *Proc. IEEE Comput. Soc. Visualization '93,* IEEE, New York, 1993, 236–243.
155. Rogowitz, B. E. and Treinish, L. A., How not to lie with visualization, *Comput. Phys.*, 10, 268–274, 1996.
156. Rosch, E. and Lloyd, B. B., Eds., *Cognition and Categorization,* Erlbaum, Hillsdale, NJ, 1978.
157. Rosch, E. and Mervis, C. B., Family resemblances: studies in the internal structure of categories, *Cognit. Psychol.*, 7, 573–605, 1975.
158. Ross, B. H., Discrimination of superficial similarities: different effects on the access and use of earlier problems, *J. Exp. Psychol. Learning, Mem. Cognit.*, 15, 456–468, 1989.
159. Ross, B. H., The use of categories affects classification, *J. Mem. Language,* 37, 240–267, 1997.
160. Ross, B. H., Post-classification category use: the effects of learning to use categories after learning to classify, *J. Exp. Psychol. Learning, Mem. Cognit.*, 25(3), 743–757, 1999.
161. Ross, B. H., The effects of category use on learned categories, *Mem. Cognit.*, 28(1), 51–63, 2000.
162. Rothfusz, L. P., McLaughlin, M. R., and Rinard, S. K., An Overview of NWS Weather Support for the XXVI Olympiad, *Bull. Am. Meteorological Soc.*, 79, 845–860, 1998.
163. Rouse, W. B. and Morris, N. M., On looking into the black box: prospects and limits in the search for mental models, *Psychol. Bull.*, 100, 349–363, 1986.
164. Rudwick, M. J. S., The emergence of a visual language for geological science, 1760–1840, *Hist. Sci.*, 14, 149–195, 1976.
165. Sadalla, E. K., Burroughs, W. J., and Staplin, L. J., Reference points in spatial cognition, *J. Exp. Psychol. Hum. Learning Mem.*, 6, 516–528, 1980.
166. Sanderson, P. M., Verbalizable knowledge and skilled task performance: association, dissociation, and mental models, *J. Exp. Psychol. Learning, Mem. Cognit.*, 15, 729–747, 1989.
167. Schneider, W., Training high-performance skills: fallacies and guidelines, *Hum. Factors,* 30, 539–566, 1985.
168. Schyns, P. G., Goldstone, R. L., and Thibaut, J. P., The development of features in object concepts, *Behav. Brain Sci.*, 21(1), 1–54, 1998.
169. *Science News*, April 21, 1984; Vol. 125, No. 16, 244.
170. Scribner, S., Thinking in action: Some characteristics of practical thought, in *Practical Intelligence: Nature and Origins of Competence in the Everyday World*, Sternberg, R. J. and Wagner, R. K., Eds., Cambridge University Press, Cambridge, 1986, 14–30.
171. Shanteau, J., Some unasked questions about the psychology of expert decision makers, in *Proc. 1984 IEEE Conf. Syst., Man, Cybernetics,* El Hawary, M. E., Ed., IEEE, NY, 1984, 23–45.

172. Shiffrin, R. M., and Dumais, S. T., The development of automatization, in *Cognitive Skills and Their Acquisition,* Anderson, J. R., Ed., Erlbaum, Hillsdale, NJ, 1981, 111–140.
173. Short, N. M., *The LANDSAT Tutorial Workbook: Basics of Satellite Remote Sensing,* National Aeronautics and Space Administration, Washington, D.C., 1982.
174. Short, N. M., Lowman, P. D., Freden, S. C., and Finch, W. A., *Mission to Earth: LANDSAT Views the World,* National Aeronautics and Space Administration, Washington, D.C., 1976.
175. Siegal, B. S. and Gillespie, A. R., *Remote Sensing in Geology,* Wiley, New York, 1980.
176. Silverman, J., Mooney, J. M., and Shepherd, F. D., Infrared video cameras, *Sci. Am.* March, 1992, 78–83.
177. Slocum,T. A. and Egbert, S. L., Cartographic data display, in *Geographic Information Systems: The Microcomputer and Modern Cartography,* Taylor, D. R. F., Ed., Pergamon Press, Oxford, 1991, 167–199.
178. Tarmizi, R. and Sweller, J., Guidance during mathematical problem solving, *J. Edu. Psychol.,* 80, 424–436, 1988.
179. Tennson, R. D. and Park, O.-C., The teaching of concepts: a review of instructional design research literature, *Rev. Edu. Res.,* 50, 55–70, 1980.
180. Thorndyke, P. W. and Hayes-Roth, B., Differences in spatial knowledge acquired from maps and navigation, *Cognit. Psychol.,* 14, 560–589, 1982.
181. Tindall-Ford, S., Chandler, P., and Sweller, J., Cognitive load theory and instructional design, *Cognit. Tech.,* 2, 49–59, 1997.
182. Treinish, L. A., Unifying principles of data management for scientific visualization, in *Animation and Scientific Visualization Tools and Applications,* Earnshaw, R. and Watson, D., Eds., Academic Press, New York, 1993, 141–169.
183. Treinish, L. A., Visualization of disparate data in the earth sciences, *Comput. Phys.,* 8, 664–671, 1994.
184. Treinish, L. A. and Rothfusz, L., Three-Dimensional visualization for support of operational forecasting at the 1996 Centennial Olympic Games, in *Proc. 13th Inter. Conf. Interactive Information Proc. Syst. Meteorology, Oceanography and Hydrology,* American Meteorological Society, Boston, 1997, 31–34.
185. Uttal, W. R., Baruch, T., and Allen, L., Psychological foundations of a model of amplified night vision in target detection tasks, *Hum. Factors,* 36, 488–502, 1994.
186. Ward, M. and Sweller, J., Structuring effective worked examples, *Cognit. Instruction,* 7, 1–39, 1990.
187. Wagner, R. K. and Sternberg, R. J., Tacit knowledge and intelligence in the everyday world, in *Practical Intelligence: Nature and Origins of Competence in the Everyday World,* Sternberg, R. J. and Wagner, R. K., Eds., Cambridge University Press, Cambridge, 1986, 51–83.
188. Ward, T. B. and Becker, A. H., Learning categories with and without trying: does it make a difference? in *Percepts, Concepts, and Categories,* Burns, B., Ed., Elsevier Science Publishers, Amsterdam, 1992, 451–493.
189. Ware, C. and Beatty, J. C., Using color dimensions to display data dimensions, *Hum. Factors*, 30, 127–142, 1988.
190. Wickens, C. D., Ed., Attention and Situation Awareness: a NATO AGARD Workshop, Aviation Research Laboratory, University of Illinois, Urbana-Champaign, 1996.

191. Williams, L. G., The effect of target specification on objects fixated during visual search, *Percept. Psychophysics*, 1, 315–318, 1966.
192. Williams, R. S. and Carter, E. D., Eds., *ERTS-1: A New Window on Our Planet,* U.S. Geological Survey, Department of the Interior, Washington, D.C., 1976.
193. Willows, D. M. and Houghton, H. A., Eds., *The Psychology of Illustration, Vol. 1: Basic Research,* Springer Verlag, New York, 1987.
194. Winn, W., Charts, graphs, and diagrams in educational materials, in *The Psychology of Illustration, Vol. 1: Basic Research,* Willows, D. M. and Houghton, H. A., Eds., Springer Verlag, New York, 1987, 152–198.
195. Winn, W., The design and use of instructional graphics, in *Knowledge Acquisition from Text and Pictures,* Mandl, H. and Levin, J. R., Eds., North-Holland, Amsterdam, 1989, 125–144.
196. Wisniewski, E. J. and Medin, D. L., On the interaction of theory and data in concept learning, *Cognit. Sci.,* 18, 221–282, 1994.
197. Woodward, A., Do illustrations serve an instructional purpose in U.S. text-books? in *Learning from Textbooks: Theory and Practice,* Britton, B. K., Woodward, A., and Binkley, M., Eds., Erlbaum, Hillsdale, NJ, 1993, 115–134.
198. Yamauchi, T. and Markman, A. B., Category learning by inference and classification, *J. Mem. Language,* 39, 124–148, 1998.
199. Yamauchi, T. and Markman, A. B., Learning categories composed of varying instances: the effect of classification, inference, and structural alignment, *Mem. Cognit.,* 28, 64–78, 2000.
200. Young, A. T., What color is the solar system? *Sky and Telescope,* May, 1985, 399.
201. Zhu, X. and Simon, H. A., Learning mathematics from example and by doing, *Cognit. Instruction,* 4, 137–166, 1987.