An Evaluation of Color Selections to Accommodate Map Users with Color-Vision Impairments

Judy M. Olson* and Cynthia A. Brewer**

*Department of Geography, Michigan State University,
**Department of Geography, The Pennsylvania State University

An experiment shows that maps can be designed to accommodate the approximately 4 percent of the population with red-green color-vision impairments. The experiment used seven pairs of maps with seven different color schemes to determine the effects of color selection on the map-reading ability of people with impaired or normal color vision. One rendition in each pair had colors that were potentially confusing to people with red-green impairments; the other had colors selected specifically to accommodate this group. On the set with potentially confusing colors, people with red-green impairments were less accurate and took longer to respond than those with normal color vision. They were just as accurate as those with normal color vision on the set with accommodating colors but continued to have longer reaction times. Logit analysis of accuracy confirmed the interaction between vision group (normal, impaired) and rendition of the map (confusing, accommodating) and indicated that performance differed from one color scheme to another. An analysis of variance of reaction times on legend-matching questions yielded similar results for the same variables. A second part of the test asked participants to choose from each pair (confusing, accommodating) of the renditions the map that was easier or better to use. Those with red-green impairments overwhelmingly chose the accommodating renditions; those with normal color vision did not show a clear preference for one rendition over the other. Key Words: Cartography, color schemes, color selection, map design, red-green color-vision impairments.

Maps are a distinctive medium of communication among geographers and between geographers and the outside world. Because maps are such an integral part of geographic thinking, geographers and cartographers have devoted considerable effort to their study.

Mapmakers often use colors as symbols representing various elements of data. Color permits clear distinction of many different categories of map features and facilitates visual grouping. By using characteristics of color analogous to those of data, color readily conveys information; for example, darker colors generally signify more of something, red and blue symbolize a difference in kind.

Unfortunately, most color maps are produced with little if any consideration of color-vision impairments on the part of map users. It is estimated that color-vision impairments of one form or another afflict approximately 4 percent of the population, and, because the trait is carried by a sex-linked gene, almost all (95 percent) of the impaired are male (Wyszecki and Stiles 1982:464; Kalmus 1965). Although the percentage of affected people is small, it translates into more than 9,000,000 people in the U.S. who may be adversely affected by current map design.

Traditionally, it has been difficult to design maps for people with impaired color vision. One reason is that maps have almost always been duplicated by printing, a medium that is not very flexible. To be sure, colors suitable for people with impaired color vision can be chosen for printing, but preparation for printing is expensive and time consuming, and color control in the printing process is far from perfect. These problems have discouraged the experimentation necessary to develop sets of suitable colors.

Annals of the Association of American Geographers, 87(1), 1997, pp. 103–134

Published by Blackwell Publishers, 350 Main Street, Malden, MA 02148, and 108 Cowley Road, Oxford, OX4 1JF, UK.

Without the experimentation, it has not been clear whether colors chosen for people with impaired color vision will also serve those with normal color vision, and high cost discourages any thought of printing more than one rendition. Understandably mapmakers have been preoccupied with designing maps for those with normal color vision.

Technology has permitted inroads to the problem, however. Computerized mapping and high-resolution color CRTs have brought considerable flexibility into color mapping. Not only can we produce maps quickly, but we can choose from literally millions of colors in producing them. Furthermore, it is conceivable that a given map can be displayed in different colors for different users if it is displayed on a color CRT. So long as we know which colors should be displayed, actually substituting these colors is a minor matter. Technology has also had a positive influence on the printing of maps. Production procedures for printed products have changed so drastically that it is feasible to consider refinements of design, including color adjustments and even production of multiple maps of the same distributions.

Simultaneously, our attitudes toward physical defects and the rights of those who possess them have changed. We are no longer prone to mistake physical handicaps for mental inability, and we are much more likely to recognize handicaps as contextual in nature. Handicaps such as blindness, paraplegia, and deafness are problems associated with specific obstacles. We provide braille or talking books and raised-image or talking maps for people without sight, ramps for those in wheelchairs, and closed-circuit captioning and visual phones for those with hearing loss. We assume that people who are capable of benefiting from facilities or understanding information should have access to the facilities or information, regardless of physical impairment.

Given the possibilities of accommodating variation in color vision by adjusting colors on CRTs or by modifying the design of maps to be printed, it seems reasonable to provide map information access to the 4 percent or so of the population that is color-vision impaired. Since choosing color for maps is a process that requires care under any circumstances, choosing with color-vision constraints in mind is but a small additional burden.

The study of map design for the color impaired is important not just for the immediate practical goal of providing a subpopulation with maps, however. In cartography, and in geography at large, it is also one piece in a larger research puzzle. We know that some people use maps effectively and easily, while others have difficulty, real or imagined, and others avoid using them altogether. How cartographic products are used and how effectively they are used are venerable research themes. One explanation for the differences in these abilities might be such physical problems as impaired color vision. Although the research reported here does not explicitly address the effects of impaired color vision on map avoidance, it does focus on map design that would eliminate the effects of this handicap.

The applicability of the research presented here is not limited to geography and cartography. Because computers have made color displays of information a common occurrence, and because computers enable ready creation and change of CRT displays, it is inexcusable to display information in ways that are unsuited to the needs of the person viewing the screen. It is important to display information effectively for whoever is using it, whether we are displaying maps, color-coded separation of various windows on the screen, multicolor graphs, or color satellite imagery. The "average user" is of little relevance for a technology so easily adapted to the specific user. Technology provides the opportunity of joining the principles of design with user needs; exploiting the flexibility of that technology makes good sense.

Research Objectives

The research carried out in this project concerned the feasibility of choosing map colors to accommodate red-green color-vision problems, which account for almost all of the 4 percent of the population with color-vision impairments (Wyszecki and Stiles 1982). More specifically, the research compared map-reading performance for two groups of people, those with normal color vision and those with red-green impairments, on two renditions of several maps. One rendition had potentially confusing colors, and the other was designed to avoid such colors, that is, to accommodate the impaired user. The long-range goal of the research is to provide readable maps for those with red-green impairments. The immediate goal is to determine whether color-vision problems affect map-reading performance and whether proper selection of colors changes that performance.

The use of two renditions of each map, one with a deliberate attempt to use potentially confusing colors and one with accommodating colors, may be seen as "stacking the deck" in favor of finding a difference. The design is crucial to the goals of the research, however, for two major reasons. First, the alternative would be to use maps "the way we normally do them." Color selection for maps, however, is not standardized. Though there is considerable consistency in the nature of the schemes we use, that is, in the general characteristics of colors that represent particular characteristics of data, there are virtually infinite numbers of specific color sets that fit each scheme. Second, and perhaps even more important, the published materials indicating which colors should be confusing to those with color-vision impairments are based on averages and on stimuli that are far different from maps. The seminal work on color impairments by Pitt, for example, employed an instrument with two precisely calibrated fields of light, one of which was adjusted until a match was perceived by the viewer (Pitt 1935). This was an experimental context, and the size of the field was carefully controlled. We cannot assume that general findings about color vision will transfer to maps, even though the obvious working hypothesis is that they do. With many sets of colors possible on maps, then, and with maps differing from the stimuli used to define the general characteristics of color impairments, we need to test the extremes of potential difficulty that maps might pose for those with impaired color vision.

Literature and Background

The literature relevant to this study falls into several categories: color research in cartography, color on CRTs, psychophysical/physiological research into impaired color vision, and the Commission on Illumination, or Commission Internationale de l'Eclairage (CIE), color system.

Color Research in Cartography

Despite considerable past interest on the part of cartographers in the map user and in the psychology of map reading, research involving color has been relatively limited. Roadblocks have included the technology of map production (traditionally manual) and reproduction (traditionally printing) and the lack of availability (to cartographic researchers) of color measurement devices. Manual production and printing of maps were expensive, time consuming, and relatively inflexible compared to CRT display of maps. The lack of color measurement devices impeded our ability to relate colors on maps to color systems used in color science and in research on impaired color vision.

As technological conditions have changed, so has the treatment of color. Increased technical competence on the part of academic cartographers is another reason for the changing interest in color. Cartography students in the 1970s and 1980s were exposed routinely and increasingly to methods of multicolor production for printing, and academic cartographers were involved in the production of high-quality maps and atlases in full color. Having to face problems of color usage on a regular basis has given considerable impetus to academic research on the subject. Add to that the appearance of inexpensive but flexible and highly accessible computers with color CRTs, and there is ample evidence that the opportunities for color research have increased dramatically.

Most cartographic literature on color presents general guidelines and rules for color usage. Textbooks such as later editions of Robinson et al. (1978, 1984, 1995), Dent (1993), and Campbell (1984) have sections on the use of color on maps and include color illustrations. As expected, recent texts devote considerably more space to color than earlier ones such as Raisz (1962) and the early editions of the Robinson text (1953, 1960, 1969), in which color was given at most a few pages of discussion and there were no pages with multicolor illustrations.

Research on color and maps has risen in parallel with its coverage in texts. Early studies compared the visual size of colored graduated symbols (Williams 1956) and provided experimental evidence that traditional rules of color progression were effective (Cuff 1973). In the 1970s through early 1990s, research by cartographers and statisticians addressed the effectiveness of color choices on bivariate maps (Olson 1981; Mersey 1980; Wainer and Francolini 1980; Trumbo 1981), prediction of scientific designations of colors produced by standard printed methods used for maps (Kimerling 1980, 1981), the classification and effects of color schemes (Eastman 1986; Heyn 1984;

Mersey 1984, 1990; Olson 1987; Brewer 1994), development of a printed approximation of perceptually ordered charts (Brewer 1989), and the effects of color surround on the perception of color on maps (Brewer 1991).

Color on CRTs

Numerous works in the literature concern color usage on CRTs, and a few are related to the work reported here. Robertson (1988) developed models of color space specifically for map-color selection, although impaired color vision was not a central concern. Meyer and Greenberg (1988) used the same reasoning as used in our research to produce continuous-tone CRT images (pictures instead of maps) in colors that should be easily differentiated by persons with impaired red-green color vision. Ware's (1988) experiments with CRT continuous-tone images (including maps) sought to identify effective colors for those with normal color vision.

Psychophysical/Physiological Research on Color Vision

The research by psychologists and physiologists on problems of color vision is central to research on color vision in map reading. Many color science and psychology texts cover the subjects of color vision and color-vision impairments at least in a very general way. Agoston (1987) offers a sound and readable introduction to color with an applied emphasis, the work of Wyszecki and Stiles (1982) serves as a valuable reference tool, and Pokorny et al. (1979) and Travis (1991) present general treatments of color vision. Travis is also among the readable coverages that integrate the three-cone theory with the opponent process model of color. The three-cone theory explains perception of color as the result of varying levels of excitation of three types of cones (receptors) in the retina of the eye that are differentially sensitive to short, medium, and long wavelengths of light. The opponent process model recognizes that blue, yellow, red, and green are "unique hues," that is, they do not look like combinations of any other colors, and it postulates that since there are no apparent mixtures of blue and yellow or of red and green, color perception results from differential excitation of blue-yellow and red-green mechanisms. The physiological existence of the three types of cones is recognized (Rushton 1975), as is their effect on mediating cells that respond as predicted by the opponent process model (Travis 1991).

A summary of color-impairment knowledge based on these sources helps set the stage for discussion of the current experiment. It also reveals an important part of the logic that we use in selecting colors for the map stimuli.

Briefly, we know that color vision can be divided into two categories (Table 1). One is called "normal trichromatic," in which case three primaries (light sources) are needed in varying proportions to match any color, and people whose vision falls in this category make virtually identical matches. Most people (more than 95 percent) have normal color vision. Among color impairments, some require three primaries to match colors (anomalous trichromatic vision), others two primaries (dichromatic vision), and others only one primary (monochromatic vision). Regardless of how many primary colors are required, many of the colors seen as matching by people with impaired color

Table 1. Categories of Color Vision

Normal trichromatic (requires three primaries to match colors)
Impaired (matches are different than for the normal trichromat)
Anomalous trichromatic (requires three primaries but matches are different from those of the normal trichromat)
protanomalous (long-wavelength [red] weakness)
deuteranomalous (medium-wavelength [green] weakness)
tritanomalous (short-wavelength [blue] weakness)
Dichromatic (requires two primaries to match colors)
protanopic (lacking long-wavelength cones)
deuteranopic (lacking medium-wavelength cones)
tritanopic (lacking short-wavelength cones)
Monochromatic or achromatic (requires only one primary to match colors)

Source: Fletcher and Voke (1985:137–43).

vision will be seen as different by people with normal color vision.

Anomalous trichromatic vision is further classified as protanomalous, deuteranomalous, or tritanomalous depending on which set of cones (short-, medium-, or long-wavelength sensitive) differs from normal in sensitivity. Dichromatic vision is subdivided into protanopic, deuteranopic, and tritanopic vision according to which set of cones is missing or nonfunctional. Each subgroup is characterized by reasonably distinct matching functions, that is, characteristic proportions of primaries that are seen as matching other colors.

People with impaired color vision are often called "colorblind," but this term has fallen from favor because it erroneously suggests that these individuals cannot see colors. The color impaired can see colors so long as the background is some distinguishable color; they merely confuse certain colors that the majority of people are able to distinguish. The terms used in color science are "defective," or "deficient," color vision, yet these terms are also unfortunate choices when one considers that some people with "defective" color vision can detect camouflage, for example, more readily than people with normal color vision (Pickford 1965). The term "color impaired" (suggested by one of the reviewers of this paper) is preferable, and we have chosen to use it here.

The current experiment concerned the collective group of people with red-green impairments. That group is an amalgamation of people with protanopic, deuteranopic, protanomalous, and deuteranomalous color vision. The reason for combining these subcategories is that confusion patterns (i.e., the colors likely to be confused) are similar for these groups, and one set of colors on a map should accommodate all groups.

The CIE Color System

The CIE has defined several color spaces for different uses. Work with impaired color vision uses the CIE 1931 *xyY* color system. Extensive descriptions are found in Agoston (1987), Wyszecki and Stiles (1982), and Judd and Wyszecki (1975). The system is best understood by starting with the notion of primary colors. If we take three colors (let us assume three sources of light), no two of which combine to produce the third, we can mix the colors in varying amounts to produce a range of other colors. These colors can be arranged logically in a two-dimensional triangular space in which they gradually blend from one to another. If chosen carefully, which generally means choosing a red, a green, and a blue because they are widely separated in color space, the three primaries can produce a remarkably wide array of colors, but there is no set of three lights that can produce all other colors. To specify all colors from three primaries (without using negative numbers), we must use mathematical ones such as those in the CIE system represented by the diagram in Figure 1. The roughly horseshoe-shaped closed form in the figure shows the relative amounts of the three imaginary light sources—*X*, *Y*, and *Z*—that are combined to specify any visible color. The axes show only *x* (relative amount of *X*) and *y* (relative amount of *Y*), because *z* (relative amount of *Z*) is simply the complement, or one minus the sum of *x* and *y*. The smooth curved part around the edge of the figure shows the *x*,*y* positions of colors in the visible spectrum. The straight line connecting the end points of the curve is "the line of purples," that is, the colors that are not in the spectrum but result from

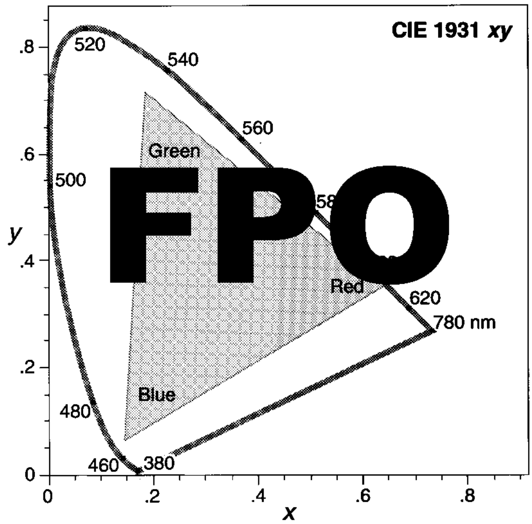


Figure 1. The CIE 1931 chromaticity diagram. The *x* and *y* labeled on the axes are mathematically constructed "primaries" (corresponding roughly to red and green, respectively). Dots represent positions of pure spectral colors at wavelengths designated in nanometers. The shaded triangle is the gamut of colors produced on our CRT. Source: Judd and Wyszecki (1975:130) (triangle based on our measurements).

mixtures of colors at the far ends of it (blue and red). Any other nonspectral color will plot somewhere within the figure, and its "chromaticity" (its look without considering how much energy is involved) can be specified by the x and y (and, implicitly, the z) values. The range of colors that could be produced on our CRT is represented by the shaded triangle with apexes at the red, green, and blue primaries of the CRT. An additional number (Y) allows the luminance (amount of energy) to be defined as well, in which case the color is fully specified. The system represented by the diagram in Figure 1 is known as CIE 1931 xyY even though only x and y are normally included in the visual representation.

The CIE 1931 xyY system is not a uniform color system, that is, equal distances on the diagram do not represent equal visual differences. Transformations were later developed that more closely approximate uniform systems (Agoston 1987; Judd and Wyszecki 1975), but because the 1931 system is the most common one encountered in the literature on color-vision impairments, and because it is an international standard, it was used in this project as well.

Color science literature indicates that people with various forms of color-vision impairments are confused by colors denoted by straight lines radiating across the CIE diagram from a given point (Figure 2). The location of the point depends on the category of impairment. If two colors of equal luminance fall along a single line in the diagram labeled R in Figure 2, protanopes ("red impaired") will not be able to distinguish them. Likewise, deuteranopes ("green impaired") will not be able to distinguish colors that fall along a single line in the diagram labeled G. Tritanopes ("blue impaired") will not be able to distinguish colors that fall along a single line in the diagram labeled B.

Pitt (1935) carried out the most crucial work in sorting out the capabilities and problems of people with red and green impairments. Pitt's experiments led to construction of the confusion-line diagrams for protanopes and deuteranopes. Wright (1952) plotted the corresponding confusion lines for tritanopes. Later research by Nimeroff (1970), Vos and Walraven (1971), Birch (1973), Estevez (1979), and Vos, Estevez, and Walraven (1990) led to modified convergence points (points from which the confusion lines radiate) for these groups. The subtle differences are less critical in the case of maps where colors differ sharply from one another, hence any of the convergence points would have served well to facilitate our choice of colors likely to be difficult or easy for the red-green impaired. The lines plotted in Figure 2 are those cited by Travis (1990): $x_p = .75$, $y_p = .25$; $x_d = 1.44$, $y_d = -.40$; $x_t = .18$, $y_t = .00$ (where x and y are the chromaticity axes of the CIE 1931 system; p, d, and t refer to protanopes, deuteranopes, and tritanopes). The diagram for tritanopes is included in Figure 2 to show that blue impairments result in confusions that are very different from those associated with red and green impairments.

The confusion lines for the various forms of impairment are generalizations derived from multiple subjects and would not necessarily provide a precise description of a single subject.

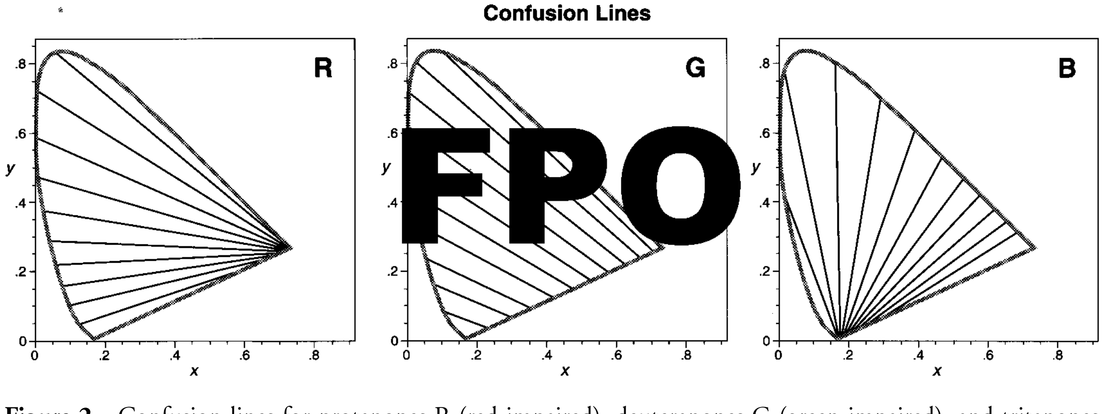


Figure 2. Confusion lines for protanopes R (red-impaired), deuteranopes G (green-impaired), and tritanopes B (blue-impaired). Source: Judd and Wyszecki (1975:133).

More important, many persons with color impairments will not confuse all colors along the line, but they will confuse those that are not sufficiently far apart, and that distance is considerably greater than required for those with normal color vision (Regan et al. 1994).

In selecting colors for maps, it follows that color pairs that fall along a confusion line on the diagram (and are not distinguishable on the basis of differing luminance) will be confusing, while colors that lie across the confusion lines instead of along them will be distinguishable. Note the similarity of pattern in the confusion lines for those with red impairments and those with green impairments; this similarity explains why the two categories are generally referred to as one ("red-green") and why our map experiment required only two sets of maps, one with confusing colors and one with accommodating colors, for both the red and green impairments. Given that any one map has only a few colors, confusing colors were chosen to align with an average confusion line for the two groups. Accommodating colors were chosen such that they were far from alignment with either set of confusion lines.

Application of the CIE System

Our designations of colors for the maps were in "RGB," that is, the red, green, and blue values that indicate the degree to which each color gun in the CRT is turned on. RGB is a standard way of defining color in computer contexts. It is CIE coordinates, however, rather than RGB, in which color impairment characteristics are defined. Applying the CIE system to maps created on a color CRT requires the use of suitable equipment to measure the displayed colors in CIE coordinates and to find RGB values that result in suitable color locations on the CIE diagram.

Our color selection procedure was as follows: We began with pencil sketches of desirable locations of colors on the CIE diagram. The sketch had the gamut of the CRT superimposed so that only displayable colors were selected (see Figure 1). We also plotted on the diagrams a selected sample of colors, such as red (R gun at maximum, G and B guns off), yellow (R and G guns maximum, B off), and green (G at maximum, R and B off). We then estimated, based on our samples, the RGB values needed to produce the chosen colors and displayed a map with those RGB colors on the CRT that we were to use in the testing (AED 1024×780 CRT with 256 displayable colors from a palette of 16.7 million). Settings of color adjustment dials were kept constant, and readings from one time to another of the same displays varied little relative to the large differences in the selected colors. Measurements were made from the CRT display using a spectroradiometer (Spectron 500) pointed at the relevant area and attached to a separate computer that calculated and displayed the CIE values for the RGB designations we had used. We then typed in modified RGB values, the displayed map colors changed instantly, and a new measurement was made. The adjustment procedure continued until all colors on a map were judged to be in satisfactory locations relative to the confusion lines on the CIE diagram. The final colors chosen were then recorded in the program that was to display the maps to subjects participating in the experiment.

The spectroradiometer available for the research was not normally used for calculating CIE values, and we performed some rough (and small) calibrations using an expected spectrum for a white calibration surface to adjust relative levels of energy across the spectrum. We used light sources with known "spikes" (highly localized areas of intense energy) to shift readings along the spectrum. Without going through an expensive check of the calibration, we could not express the degree of error in measurement, but after consultation with experts at the manufacturing company and with a lighting expert, and after observing that our measurements consistently "made sense" (our color gamut coincided with published CRT gamuts, and our readings were in the expected regions of the CIE diagram), we concluded that the error was acceptable relative to the large differences between colors on maps. The accuracy of instruments designed for matching colors (of paints, for example) would be well beyond what is needed to place well-separated colors into their appropriate locations on a CIE diagram to see if they fall "close to" or "far from" confusion lines. Precalibrated colorimeters with known levels of error have since become readily available at affordable prices and are used in color map research, but colorimeters have filters that yield approximations of primaries rather than precise readings along the spectrum in the manner of a spectroradiometer. We are confident that the accuracy and precision of our instrument compare well with those of a colorimeter.

Color Schemes and Test Stimuli

Knowing how to confuse or not confuse people with impaired color vision is not sufficient for selecting colors for a map experiment. One must also consider the cartographic, or representational, rationale. A given set of colors, no matter how clearly distinguishable, cannot be used for all maps we might want to construct. For example, if we map median income (a quantitative variable) by county, we are very likely to use a set of colors that vary from light (for low income) to dark (for high income). If we make a map to show dominant employment sector (a qualitative variable), we need a set of colors that suggests differences in kind and will probably use colors of different hue (green, red, yellow, blue). A map to show change in median income (a quantitative variable but with a natural dividing line at zero) requires yet a different set of colors, this time clearly distinguishing between increases and decreases as well as showing quantity of the increase or decrease.

Color choices on maps do not depend on data characteristics alone, however. In the first example, we wanted the set of colors to vary from light to dark, but we may want either a single-hue set or a multihue set. The characteristics of color being employed (hue, lightness, saturation) can also be involved in the classification of color schemes.

Note that in all of the examples, whether based on data characteristics or data plus color characteristics, only a general description of the colors has so far appeared; we have not indicated specific colors. We refer to a general class of colors as a color "scheme" and to a specific manifestation of the scheme as a "rendition" of it. Alternative renditions of the color scheme for the first data example, using monochrome colors, might be: (1) light green, medium green, dark green; or (2) light purple, medium purple, dark purple. For the same map using multiple hues, alternatives might be: (1) light yellow, medium orange, dark brown; or (2) light yellow, medium yellow-green, dark green.

The problem of selecting different color schemes to suit different maps was addressed by incorporating as many schemes into the project as possible. We then used two renditions of each, one with colors potentially confusing for the red-green impaired and referred to as "confusing" maps, the other with colors chosen carefully to avoid confusion and referred to as "accommodating" maps.

In identifying the color schemes, we tried to be reasonably inclusive of all that are used on maps, and our list was a refinement of a previous classification (Olson 1987). Our terminology is adapted from more recent work by Brewer (1994). Seven map color schemes were identified for use in the testing. The maps constructed to represent each color scheme (both the confusing and accommodating renditions) are described verbally here and are shown in Figure 3. The CIE locations of colors are depicted graphically and numerically in Appendix A. For completeness, both Figure 3a and the Appendix (Figure A-1) include the practice map used in the testing procedure ("Population Density, China, 1982," adapted from Pannell and Ma 1983). The maps were designed for viewing on a CRT. With one exception, all maps were adapted from actual printed maps, most of which appeared in published sources. No attempt was made to reproduce the original colors or even use the same color schemes as on the originals; existing maps were used only to provide a realistic setting within which to evaluate the effects of color choices on the red-green impaired. In describing our renditions of the various schemes, we use the perceptual vocabulary of hue, lightness, and saturation. Hue refers to the character of color represented by such names as yellow, orange, red, and green. Lightness refers to the relative amount of light, with white at the high end of lightness, black at the low end. Saturation refers to the colorfulness versus grayness of a color. The abbreviations for the various schemes are included in parentheses here and are used in later tables and diagrams.

Qualitative/area (Qa). A qualitative/area scheme shows differences in kind but not differences in quantity. The symbols occupy area on the map in proportion to the size of the original feature or perhaps some measure of its importance. The general cartographic convention is that variation in hue but not in lightness should be used to represent qualitative differences on maps. This rule is seldom followed precisely, however, because lightness variations help to provide differentiation and visual definition in maps. Nevertheless, hue differences are usually the most noticeable variation in the colors used to represent qualitative differences on maps. To represent the qualitative/area scheme, we chose the map shown in Figure 3b and c, which is a large-scale map from the *Atlas of California* depicting five categories of building use in a section of San Francisco (Donley et al. 1979:22). Three of the map colors on the confusing rendition (green, red-orange, and yellow) fall on a single average confusion line. Another confusing pair (blue and magenta) fall along another such line. Colors in each of the confusing sets were nearly equal in lightness. The accommodating rendition used both a series of hues located perpendicular to the confusion lines and lightness contrast (see Figure A-2.)

Qualitative/dot (Qd). A qualitative/dot scheme uses dots of different hue to represent quantities of more than one category of item. The scheme differs from the qualitative/area one because the symbols are equal in size. The quantity of each item is represented by the relative number of dots, and quantity variation is normally an important part of the map message; thus we might have given the scheme a label that reflects the qualitative/quantitative nature of the graphic information. The color choices, however, are solely qualitative. For the qualitative/dot scheme, we constructed the map in Figure 3d and e, "Livestock Inventory, South Dakota, 1982," which shows densities of cattle, hogs, and sheep and is based on census data (U.S. Bureau of the Census 1982). The confusing rendition used dots whose colors were a very dark red, which is hard for protanopes to see, and dots of purple and cyan, which fall on a confusion line. The accommodating rendition used blue, red, and white dots that varied in lightness as well as hue (see Figure A-3).

Sequential/monochrome (Sm). The sequential/monochrome scheme represents a sequence of data values using differing lightnesses of a single hue. Saturation often changes as well, and there may be minor shifts in hue, but the most noticeable visual variation is in lightness. Our example, shown in Figure 3f and g, is "Timber Harvest, Pacific Northwest 1981" and was adapted from the *Atlas of the Pacific Northwest* (Kimerling and Jackson 1985:105). The "confusing rendition" had lightness steps of magenta (a reddish color). This sequential/monochrome scheme was not expected to cause great difficulty for those with impaired color vision, but because protanopes have difficulty seeing red light, steps of reddish colors would be the most likely to cause problems in this scheme. Also, our chosen colors followed a confusion line for deuteranopes, which means that saturation differences would not be accessible to this subgroup of red-green impaired. We used steps of gray, with which greater lightness differences are possible, on the accommodating rendition. Both the red-green impaired and normal groups should easily distinguish gray steps (see Figure A-4).

Sequential/polychrome (Sp). The sequential/polychrome scheme differs from the previous one in that it incorporates hue change as well as orderly variation in lightness. Hue variations are used to allow more contrast between categories, but lightness should still be the primary differentiating characteristic. Our example is shown in Figure 3h and i and is an adaptation of the map "Sales from Improved Farmland, Southern Ontario Townships" in the *Southern Ontario Atlas of Agriculture, Contemporary Patterns and Recent Changes* (Keddie and Mage 1985:24). The confusing rendition had two pairs of colors (purple and blue, green and yellow-orange) that fell on common confusion lines and that differed only slightly in lightness. The sequence progressed from dark to light colors starting with purple and passing through blue, green, yellow-orange, and red. The accommodating rendition used dark blue, less-dark blue, green-blue, light green, and very light yellow (see Figure A-5).

Diverging (D). The diverging scheme emphasizes extremes of data away from a critical mid-value or zero point. If one category represents data values that straddle the critical point, that category is represented either by a neutral or by a color that blends in both directions with the adjacent colors. A light gray might be used with increasingly darker reds on one side and increasingly darker blues on the other. Alternatively, yellow might give way to increasingly darker and more saturated greens on one side and increasingly darker and more saturated reds on the other. Diverging schemes in which the critical point is a category break have no neutral or third color that represents "near middle values." Rather, there is a distinct change in hue on either side of the critical point. Our example of a map with this scheme is shown in Figure 3j and k, "Change in Manufacturing Jobs, 1970–1982, Georgia" which is adapted from *The Atlas of Georgia* (Hodler and Schretter 1986:109).

Figure 3. The maps used in the testing (printed colors are only approximations of CRT colors): (a) the practice map; (b) qualitative/area, confusing; (c) qualitative/area, accommodating; (d) qualitative/dot, confusing; (e) qualitative/dot, accommodating; (f) sequential/monochrome, confusing; and (g) sequential/monochrome, accommodating. *(Continued on next page)*

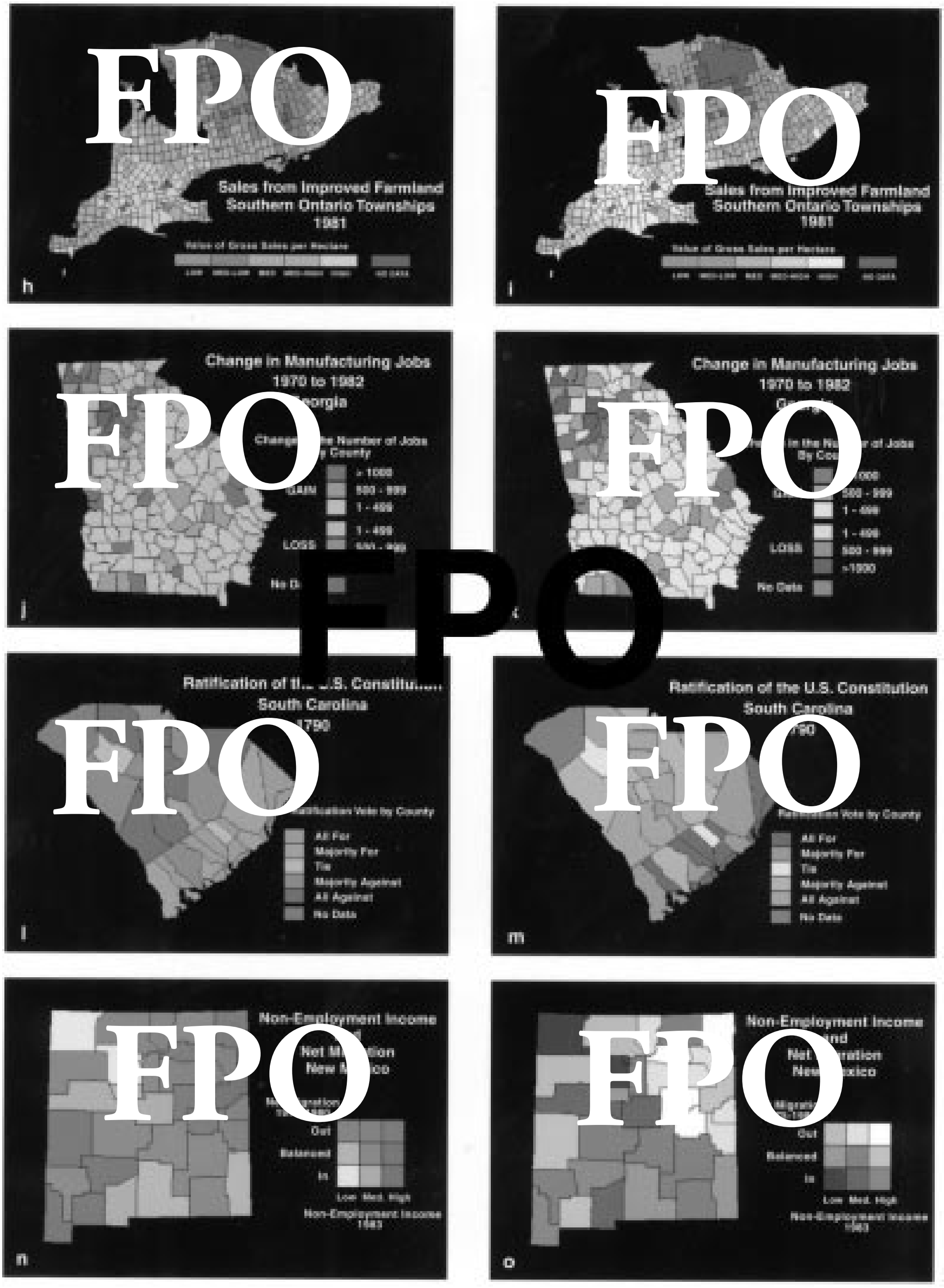


Figure 3, *continued.* (h) sequential/polychrome, confusing; (i) sequential/polychrome, accommodating; (j) diverging, confusing; (k) diverging, accommodating; (l) balance, confusing; (m) balance, accommodating; (n) two-variable, confusing; and (o) two-variable, accommodating.

With three categories on either side of the zero value, the confusing rendition used yellow-green, green, and dark green to represent increasing gains in jobs and yellow-orange, orange, and dark red to represent increasing losses. Since three color pairs with equal lightness (yellow-green and yellow-orange; orange and green; and dark red and dark green) fell on a single confusion line, these were predicted to be confusing. The colors in the accommodating set involved lightness steps of blue to represent increasing gains and steps of red to represent increasing losses (Figure A-6).

Balance (B). The balance scheme represents two sets of data that are complementary, that is, values must add up to a constant (usually 100 percent), and more of one means less of the other. Two colors at the extreme ends of a balance scheme legend are blended to represent the intermediate categories, and the midpoint should look like an even mixture of the two end colors. Our example of a map with this scheme is shown in Figure 3l and m and is an adaptation of "Ratification of the U.S. Constitution, South Carolina, 1790" from the *Atlas of Early American History* (Cappon et al. 1976:63). The proportions of "yes" and "no" votes are complementary and total 100 percent in every county. The confusing rendition had colors from magenta-red at one end through magenta, gray-purple, and gray-blue to blue at the other end. All three of the middle colors fell on the same confusion line. The accommodating rendition used blue, blue-gray, gray, gray-orange, and orange.

The balance scheme is probably the most problematic for retaining the logic of the scheme while using colors that fall at right angles to the confusion lines. Although we could have stayed closer to the line of purples (lower edge of the CRT gamut triangle; see Figure 1), it is difficult to get a light color between red and blue without moving toward gray and thus rendering a diverging rather than balance scheme (see Figure A-7).

Two-variable (2V). The two-variable map shows two variables at a time by superimposing a set of colors representing one variable onto the set of colors representing the other variable or otherwise coloring a matrix of value combinations. Our example, shown in Figure 3n and o, is entitled "Non-Employment Income and Net Migration, New Mexico" and was adapted from a map of the entire U.S. by James B. Moore ("non-employment income" refers to money received from such sources as retirement, investments, and welfare). The confusing rendition was a spectral arrangement with neutral (medium gray) in the middle of the legend (representing medium values on both variables), light yellow representing low values on both variables, and dark purple-blue representing high values on both. Remaining colors were other points in the hue circle with greens for high values on the first variable and low on the second, and reds for low values on the first variable and high on the second. There were three pairs of potentially confusing colors: green-blue and magenta, green and orange-red, and yellow-green and orange. The accommodating rendition used neutrals along the diagonal with medium gray for medium values on both variables, dark gray for low values, and white for high values. Blues were on one side of the diagonal and oranges on the other side. The two renditions of this scheme were quite different in the range of hues used, and their legends arranged lightness and darkness in opposite directions. Diverging logic was emphasized in the confusing rendition and sequential logic in the accommodating rendition (see Figure A-8).

Test Questions

Subjects were asked both content and matching questions. Content questions required the subject to consider spatial relationships and the relationship between colors and what they represented. With many questions, the display program would draw a line on the map, blink an arrow, or highlight the boundary of a subarea as needed to facilitate the question. Due to the vertical positioning of the CRT, we referred to top, bottom, right, and left on the map as needed. Examples of content questions include:

Describe the trend in sales from TOP to BOTTOM along the line displayed.

a) LOW to MEDIUM
b) LOW to MEDIUM-LOW
c) all LOW
d) all MEDIUM-LOW
e) MEDIUM to MEDIUM-HIGH
x) not answerable; relevant colors too similar

How many ENTERTAINMENT establishments are located on Grant Avenue?
a) none or one
b) two
c) three
d) four
e) five
f) six or more
x) not answerable; relevant colors too similar

Two content questions were asked for each rendition of each map. Two versions of the pair of questions were constructed with an effort to make them equal in difficulty and in general nature. The two versions were necessary because each subject would see each map twice, once with the confusing colors and once with the accommodating ones, and memory of previous answers might have rendered the test useless if identical questions had been asked. In the examples above, the alternate version of the first question simply displayed the line in a different position, and the answer choices were adjusted accordingly. The alternate version of the second question referred to a different street on the map. A copy of all content questions used in the experiment is available from the authors.

The matching questions asked the subject to identify the legend color that matched a selected areal unit on the map. The wording of all matching questions was of the form: "Find the blinking arrow on the map. Which legend color does this unit match?" The number of matches varied from map to map according to the number of colors it contained.

The reasons for including both content and matching questions are twofold. Asking only content questions would not have revealed whether differences in performance were associated with difficulty in distinguishing colors, or, if no differences were found, would not have revealed whether colors were equally distinguishable for the two groups, or if other factors equalized the performance. On the other hand, asking only matching questions would not have revealed whether subjects could read the maps. We wanted both color-discrimination and map-reading information.

All questions were in multiple-choice format. Subjects were to read the question, look at the map to determine the answer, press the down-arrow key until their choice of answer was highlighted, and then press the return key. Answer wording followed map wording as closely as possible (capital letters on the map were capitalized in answer choices, for example) and choices for matching questions were always in the same order as in the map legend. The last choice for every question was "not answerable; relevant colors too similar." This choice was included in an effort to equalize the difficulty level for the two groups of subjects based on earlier findings that forced-choice questions are extremely fatiguing for those with red-green impairments (Moore 1988). The "not answerable" choice allowed immediate expression of a subject's inability to distinguish colors and avoided prolonged and frustrating indecision. To be sure, inclusion of the choice meant running the risk of subjects using that choice to please us, since they knew the experiment had to do with maps and the color impaired. Had they been merely trying to please us, however, they would have used the answer just as often with the accommodating maps, and that proved not to be the case.

Test Order

Test order involved two components: (1) the sequence in which subjects encountered the fourteen maps (two renditions of each of seven schemes) and (2) the way in which alternate versions of questions were attached to the renditions of the maps. Two sequences of the maps were distributed equally within each group of thirty-two subjects: Sixteen subjects with normal color vision worked with sequence A, and sixteen others worked with sequence B; likewise, sixteen red-green impaired worked with sequence A, and sixteen others with sequence B. Sequence A of the maps was determined using a random number table with the constraint that alternate renditions (confusing, accommodating) of the same scheme were to be separated by at least three other maps. Sequence B simply placed the maps in reverse succession. The pairs of question versions were alternated for the two renditions of each color scheme such that a particular question was asked of one rendition half the time and of the other rendition half the time. Thus, four subsets of each group of subjects (normal color vision, red-green impaired), each with eight subjects, had identical map sequence and question versions. This balancing of the test design assured that the groups of subjects and renditions of maps could be compared without map sequence or question version biasing the comparison.

Subjects

Sixty-four subjects were included in the final tabulations. Thirty-two had red-green impairments and thirty-two had normal color vision. Subjects were obtained by advertising in the student newspaper, by posting notices at various locations on the Michigan State University campus, and by sending notices to places such as the local Army Reserve Office, men's social clubs, and eye-care centers. With off-campus groups, a phone call preceded our notices. From each group, we asked for pairs of volunteers; for example, if an Army Reservist with red-green impaired vision volunteered, we wanted an Army Reservist with normal color vision as well. The student newspaper ad asked only for "colorblind" volunteers, whereas the notices posted on campus asked for both colorblind and normal volunteers. The latter were expected to draw a sufficient number of those with normal color vision from the same general population. We used the term "colorblind" in our solicitation because the term is widely recognized by the general public.

Our efforts to solicit people with color impairments and with normal vision from the same groups (students, men's groups) helped to attract similar subjects. We kept track of their occupations, ages, and genders, and we used telephone screening by the office staff to assure reasonable similarity between the subjects with red-green impairments and those with normal color vision. Table 2 summarizes occupation, age, and gender of our subjects.

Table 2. Summary of Occupation, Age, and Gender of Subjects in the Experiment

	Normal vision	R-G impaired
Occupation		
Undergraduate student	19	18
Graduate student	3	1
Professor	1	3
Military	4	3
Other	5	7
Age (in years)		
<20	1	3
20–29	20	16
30–39	6	8
>40	5	5
Gender		
Male	30	30
Female	2	2
n =	32	32

Testing Procedure

When volunteers appeared for testing, they first signed a consent form and filled out a background questionnaire. They were also given a brief verbal description of the steps in the procedure and were reminded that it would take approximately one hour. Subjects were then seated before a microcomputer in the testing room, which was illuminated with the same lamp for every testing session and with window blinds closed to keep lighting as consistent as possible. The microcomputer setup had both a monochrome and a color CRT, for question and map display, respectively, and specific instructions were given orally for the main test. Subjects answered a series of questions on the practice map and were encouraged to ask questions if anything was unclear. None had difficulty understanding the task.

Once a subject completed the practice map questions and confirmed readiness for the test, the main testing procedure began, and the examiner moved to the back of the small testing room, out of sight of the subject. The subject knew the examiner was present, but the examiner's position effectively discouraged any interaction. The examiner was to intervene only if the computer malfunctioned or if the subject had a problem. Computer malfunction fortunately did not occur. Other problems did occasionally arise. The most common one was confusion with the first matching question on the two-variable map; this was the subject's first encounter with the answer choices in multiple columns on the screen (to match the arrangement of the legend matrix on the map). The examiner made notes of such events and eleven reaction times (out of 6,144 for the sixty-four subjects) were eliminated from the final results.

The computer was programmed to display a new map for 20 seconds during which time the subject was to look at the map and become familiar with it. The map screen (color CRT) would then go blank until the subject pressed the space bar for a question. The map would redisplay on the color CRT while the question appeared on the monochrome CRT. As soon as the subject answered the question, both CRTs would go blank until the space bar was pressed for the next question.

Because it took a noticeable amount of time for the computer to produce each map for the first time, subjects were told about the pauses ahead of time and were asked to wait silently, using these times to rest. The monochrome screen informed the subject when a new map was being processed. No one indicated that the pauses were bothersome. A beep sounded when the new map appeared, assuring that the subject's attention was recaptured promptly. To the subject, the new map appeared instantaneously when the computer beeped; this was possible because the computer program had drawn out each map with all colors defined as black and then converted it to a full-color map by redefining the color table just as the beep sounded. When the map was turned off after each question, the reverse occurred, and the color table defined all colors as black once again.

The timer for the responses began the instant the map reappeared with a question. The computer recorded the elapsed time at the subject's first pressing of the down-arrow key and again at the pressing of the "enter" key. We later applied a correction that allows for the multiple-choice format. The correction was based on a regression of reaction time against answer position. The responses used were those near the median reaction times (within a time span equal to the median plus or minus the distance from the median to the fastest response for each position of answers) and was based on the assumption that longer answer times than those predicted by regression lines were due to additional decision time after first pressing the down-arrow key.

For each map, the two content questions were always presented first, followed by the matching questions, which varied in number depending upon the number of colors on a map. Matching questions included all potentially confusing colors plus all but one other color appearing on the map (except for the dot map, on which all three colors were matched). The omission of one color per map was an attempt to avoid getting answers to the last match by elimination. By the end of the test, each subject had seen both renditions of all seven schemes and had answered ninety-six questions (twenty-eight content, sixty-eight matching).

The appearance of the map on the screen for all questions meant that only the immediate readability of the map was being tested, not memory. Map memory would require a very different test design, and we felt the more immediate need was to understand how well people with red-green vision impairments could function while the map was in view.

Instructions for a second part of the test were given after completion of the main test. Subjects were asked which rendition from each pair was easier or better to use. Each of the seven maps was redisplayed and, for each one, subjects could press a key to toggle between the two renditions. They could toggle any number of times. They were asked to press the enter key when the CRT displayed the rendition that they found easier or better to use. This section of the test took only a few minutes.

The last item in the formal procedure was the Ishihara test (Ishihara 1984) to determine the status of color vision of the subject. It was administered from the published paper test booklet with natural or simulated daylight in a different room set up for this purpose. The Ishihara test is the familiar one with colored dots of various sizes with embedded numbers; on several of the plates, people with red-green impairments read different numbers than those with normal color vision. Only two persons were categorized differently than they had reported to office staff when signing up for the test: one who thought his color vision was normal was actually red-green impaired, and one who thought he was impaired turned out to have normal color vision.

Following the whole procedure, the subject was paid $10 for participating and was invited to comment on the test. Comments were generally positive about the goals of the project and about the clarity of the tasks and instructions. No one complained about the length of the test (roughly one hour), as the length had been made clear in the original advertising as well as in preliminary conversation when each subject arrived for testing.

Methods of Analysis and Expected Results

The testing generated a mass of data, and it was critical to decide on analyses that would focus on important relationships. There were eight potential variables in the main test: (1)

accuracy of response (right, wrong), (2) reaction time (seconds), (3) vision group (normal, red-green impaired), (4) rendition (confusing, accommodating), (5) map color scheme (qualitative/area, qualitative/dot, sequential/monochrome, sequential/polychrome, diverging, balance, two-variable), (6) question type (matching, content), (7) test map sequence (sequence 1, sequence 2), and (8) version of content questions (version A, version B). The central questions, however, concerned the comparison of accuracy and reaction times between the people in the two vision groups (normal, red-green impaired) and the comparison of their performance on the two renditions of the color schemes (confusing, accommodating). These two variables were expected to show the clearest results; people with normal color vision were expected to perform accurately and quickly on both confusing and accommodating renditions; people with red-green impairments were expected to perform with significantly lower accuracy and slower reaction times on the renditions with confusing colors but just as accurately and quickly on the accommodating renditions. Color scheme was considered an additional, but secondary, variable; at least one scheme (sequential/monochrome) was expected to show a weaker result (less difference in performance between the two subject groups on the rendition with potentially confusing colors). Both the sequence of the maps in the test and question version for content questions were carefully distributed among subjects so that an equal number of people in each vision category worked with each combination. Any influence of these variables, then, was expected to be evenly distributed over the two vision groups. The categories (matching, content) in the type-of-question variable are so drastically different that we chose to analyze results for the two question types separately.

Since there were two types of questions (content, matching) and two types of dependent variables (accuracy, reaction time), we chose to pursue four sets of analyses. For the accuracy results, we used logit analysis (Knoke and Bohrnstedt 1994). The logit technique is a class of log-linear analysis that allows a dichotomous variable (right and wrong responses in this case) to be distinguished as dependent. Log-linear analysis, in general, is a technique for analyzing relationships between categorical variables and involves logarithms of odds for combinations of conditions. In our case, for example, one set of conditions would be normal vision, confusing colors, and a diverging scheme; the question is whether the odds of a correct answer under these conditions is significantly greater than, say, for impaired vision, confusing colors, diverging scheme. In logit analysis, one fits models with varying combinations of variables included to find the model that gives the best fit relative to the degrees of freedom lost as variables are included. We used hierarchical models, in which all individual variables enter as well as any interactions specified. We compared models using the Bayesian information content (BIC) statistic, which is an index used to rank models and is a function of actual and expected frequencies with a correction for degrees of freedom and sample size. The ranking procedure is used by researchers with large sample sizes (ours were 1792 for content questions and 4350 for matching) because conventional tests almost always result in rejection when sample size is large, even when a model fits well (Raftery 1986; Cressie 1993:498; Knoke and Bohrnstedt 1994:381). All models in our case used accuracy of response as the dependent variable. We expected the best-fitting model to include: (1) the interaction between vision group (normal, impaired) and rendition of color scheme (confusing, accommodating), and (2) the interaction between those two independent variables and scheme (qualitative/area, qualitative/dot, etc.). That expectation applied to analysis of matching as well as content questions. We compared the fit of all combinations of the four variables (A, accuracy of response; V, vision group; R, rendition; and S, scheme), and the fully saturated model included the fifth variable, O, test order, which specified the particular configuration of the test for each subject.

Reaction time is a quantitative rather than categorical measure. It was included for observation in the experiment because accuracy is not always a good measure of difference in performance. People with red-green impairments may find it difficult to make distinctions between colors and take a longer time to react even if they are able to answer a question correctly. We analyzed only the reaction times for correct responses because those for incorrect responses are difficult at best to interpret. A subject could be answering quickly because it is not worth spending any more time on the question, not because the question is easy to answer.

Analysis of variance with a mixed design was used to compare mean reaction times. Repeated measures analysis (Snodgrass et al. 1985) was used for the scheme and rendition variables because the testing had been purposely structured so that all subjects answered questions about both a confusing and an accommodating rendition of every color scheme. Vision was a between-subjects variable because subjects had either normal or impaired color vision.

In the second part of the test, in which subjects chose which rendition was easier or better to use, a separate binomial test was employed with the results for each subject group to see if selections were significantly different from 50 percent (no bias in choice of renditions). Expected results were a significant difference (in favor of the accommodating maps) for the red-green impaired and a nonsignificant difference for the normal color-vision group.

Results and Analysis

The results of the testing are presented in Tables 3 and 4. They include the percentages of correct answers and reaction times for each type of question (content, matching) and indicate numbers of responses. The results are given by color scheme as well as for vision groups and renditions aggregated over schemes. These data summarize a large number of responses (6144) that are used in the analyses of accuracy and reaction times.

Accuracy

The results of the logit analyses are summarized in Table 5. The fully saturated model appears first, with increasingly simpler models following. The second model leaves out test order. The next model includes the interaction between accuracy of response, vision group, and rendition {AVR}; the interaction between accuracy, vision group, and scheme {AVS}; and the interaction between accuracy, rendition, and scheme {ARS}, plus all the interaction pairs and individual influences of the included variables, as all of these models are hierarchical. It leaves out not only test order but all of the high-order interactions between included variables. Like-

Table 3. Responses to Content Questions

		Percent Correct Responses (*n*)		Mean Reaction Time in Seconds (*n*)	
	Rendition:	Confusing	Accomm.[a]	Confusing	Accomm.[a]
	Vision				
All Schemes	Impaired	59 (448)	74 (448)	20.39 (262)	19.69 (332)
	Normal	79 (448)	74 (448)	16.50 (353)	16.88 (333)
Each Scheme					
Qualitative/area (Qa)	Impaired	42 (64)	72 (64)	19.76 (27)	18.80 (46)
	Normal	58 (64)	64 (64)	15.17 (37)	16.20 (41)
Qualitative/dot (Qd)	Impaired	36 (64)	67 (64)	15.23 (23)	13.03 (43)
	Normal	80 (64)	73 (64)	13.22 (51)	12.15 (47)
Sequential/monochrome (Sm)	Impaired	83 (64)	84 (64)	17.99 (53)	18.60 (54)
	Normal	81 (64)	84 (64)	15.66 (52)	15.79 (54)
Sequential/polychrome (Sp)	Impaired	78 (64)	67 (64)	18.85 (50)	15.58 (43)
	Normal	88 (64)	72 (64)	15.91 (56)	16.21 (46)
Diverging (D)	Impaired	45 (64)	69 (64)	21.04 (29)	21.56 (44)
	Normal	92 (64)	77 (64)	15.41 (59)	15.47 (49)
Balance (B)	Impaired	55 (64)	86 (64)	24.45 (35)	22.24 (55)
	Normal	75 (64)	80 (64)	18.12 (48)	18.59 (51)
Two-variable (2V)	Impaired	70 (64)	73 (64)	23.37 (45)	26.95 (47)
	Normal	78 (64)	70 (64)	22.07 (50)	24.03 (45)

[a]Accommodating rendition of color scheme.

Table 4. Responses to Matching Questions

	Rendition:	Percent Correct Responses (*n*) Confusing	Percent Correct Responses (*n*) Accomm.[a]	Mean Reaction Time in Seconds (*n*) Confusing	Mean Reaction Time in Seconds (*n*) Accomm.[a]
	Vision				
All Schemes	Impaired	79 (1,088)	96 (1,088)	6.39 (860)	5.48 (1,045)
	Normal	98 (1,088)	96 (1,088)	4.20 (1,063)	4.47 (1,045)
Each Scheme					
Qualitative/area (Qa)	Impaired	72 (128)	100 (128)	6.48 (92)	5.60 (128)
	Normal	100 (128)	99 (128)	3.67 (128)	4.68 (127)
Qualitative/dot (Qd)	Impaired	69 (64)	97 (64)	6.89 (44)	4.80 (62)
	Normal	100 (64)	98 (64)	4.22 (64)	4.06 (63)
Sequential/monochrome (Sm)	Impaired	96 (128)	100 (128)	6.06 (123)	5.18 (128)
	Normal	98 (128)	99 (128)	4.85 (125)	4.66 (127)
Sequential/polychrome (Sp)	Impaired	90 (160)	98 (160)	5.61 (144)	5.03 (157)
	Normal	99 (160)	93 (160)	4.06 (159)	4.20 (148)
Diverging (D)	Impaired	65 (192)	88 (192)	7.15 (124)	5.74 (169)
	Normal	96 (192)	98 (192)	4.55 (185)	3.82 (189)
Balance (B)	Impaired	74 (160)	99 (160)	6.73 (119)	5.25 (158)
	Normal	98 (160)	99 (160)	3.66 (156)	4.69 (158)
Two-variable (2V)	Impaired	84 (256)	95 (256)	6.34 (214)	5.99 (243)
	Normal	96 (256)	91 (256)	4.17 (246)	4.87 (234)

[a]Accommodating rendition of color scheme.

wise, each remaining model includes the interaction between variables included within each set of curly brackets plus all lower-order effects.

L^2 in Table 5 is the test statistic, *df* is the degrees of freedom, and BIC is the indication of goodness of fit. The lower the BIC value, the better the fit (Knoke and Bohrnstedt 1994:382). We considered the fit for both sets of data—content and matching questions—in selecting our "best model."

For content questions, the {AVR}{AS} ranks first and the even simpler {AVR} model is a close second. This means that the interaction between accuracy of response, vision group, and rendition were extremely important, and scheme was also important but only as an individual variable, not in its interaction with the other variables. In other words, there were notable differences in performance with schemes regardless of the category of color vision or rendition of the scheme.

For matching questions the best fit is the {AVR}{AVS} model, and the {AVR}{AS} model is a close second. In this case the best model does include the interaction between accuracy of response, vision group, and scheme, although the one with scheme by itself, {AVR}{AS}, is almost as good.

Given the pattern of results, the simpler {AVR}{AS} seems to be the best overall model considering both types of questions. Generally, results were those expected except that the scheme variable does not interact with vision group and rendition as clearly as we anticipated; that interaction improves the model only for the matching questions.

A more detailed look at the logit results emerges from Table 6, where the coefficients, Z scores, and odds of a right versus wrong answer are shown by component term in the model. The coefficients column gives the effect parameters, which are positive for variables with a positive effect on the dependent variable (correct responses); the higher the value the stronger the effect. The Z score is the measure used to determine significance of the effect. The estimated odds give the expected ratio of correct to incorrect responses associated with each component term. The first line indicates that, overall, there were more correct than incorrect responses for both content and matching questions, the ratio was significant, and the odds of a correct versus incorrect answer was 2.65:1 for content and 21.6:1 for matching. The second line indicates that those with normal color vision had significantly more correct answers than those with

Table 5. Hierarchical Logit Models

Hierarchical Models	Content Questions n = 1,792			Matching Questions n = 4,352		
	L^2 [b]	*df* [c]	BIC [d]	L^2 [b]	*df* [c]	BIC [d]
{AVRSO}[a]	0.0	0	0.0	0.0	0	0.0
{AVRS}	180.6	84	–449.2	93.0	68	–476.8
{AVR}{AVS}{ARS}	189.6	90	–484.6	110.4	90	–643.7
{AVR}{AVS}	214.9	96	–504.3	138.3	96	–666.1[e]
{AVR}{ARS}	211.5	96	–507.7	146.4	96	–657.9
{AVS}{ARS}	211.9	91	–469.8	155.4	91	–607.0
{AVR}{AS}	**237.0**	**102**	–527.1[e]	**193.3**	**102**	–661.3[f]
{AVS}{AR}	236.2	97	–490.5	205.1	97	–607.6
{ARS}{AV}	233.5	97	–493.1	212.1	97	–600.6
{AVR}	286.5	108	–522.6[f]	271.3	108	–633.6
{AVS}	243.4	98	–490.7	305.9	98	–515.2
{ARS}	257.8	98	–476.3	361.1	98	–460.0
{AV}{AR}{AS}	257.8	103	–513.8	258.9	103	–604.1
{AV}{AS}	265.0	104	–514.1	358.2	104	–513.2
{AV}{AR}	306.7	109	–509.9	335.1	109	–578.1
{AR}{AS}	281.8	104	–497.3	406.8	104	–464.5
{AV}	313.6	110	–510.4	432.4	110	–489.2
{AR}	330.0	110	–494.0	480.4	110	–441.3
{AS}	288.8	105	–497.7	502.6	105	–377.1

[a]Definitions of variables included in the models are as follows:
Dependent dichotomous variable—**A,** accuracy (right, wrong).
Independent categorical variables—**V,** vision group (normal, impaired);
R, rendition (confusing, accommodating);
S, color scheme (Qa, Qd, Sm, Sp, D, B, 2V);
O, order (1, 2, 3, 4).
[b]L^2 is likelihood ratio.
[c]*df* is degrees of freedom.
[d]Bayesian information content (BIC) statistic = $L^2 - (df)(\ln n)$.
[e]Lowest BIC.
[f]Second lowest BIC.

impairments in both cases, and that the odds of a correct response on a content question for normals (controlling for the overall greater number of correct responses) were 1.29:1 for content questions and 1.87:1 for matching. The confusing renditions had a negative effect on accuracy of responses, while the odds of a correct answer on confusing renditions were .89:1 (content) and .71:1 (matching). The influence of individual color schemes varied, with the qualitative/area, qualitative/dot schemes having a significant negative effect on accuracy of responses to content questions and sequential/monochrome a significant positive effect. The sequential/monochrome has a similar effect in the matching data, and the diverging and two-variable schemes have a significant negative effect. The interaction term (correct answers, normal vision, confusing rendition) shows a significantly positive effect.

The estimated odds in Table 6 are supplemented in Table 7 with odds affected by interactions. These odds have been calculated using the procedure outlined in the SPSS manual (SPSS 1988:544). The estimated and observed odds match closely.

The relationships summarized by the logit model are visually illustrated in a series of interaction graphs (Figure 4). In Figure 4a, the relatively equal accuracy of the normal vision group on the two renditions is apparent, whereas the lines for the impaired group are steeply inclined with far better performance on the accommodating maps. Figure 4b shows the differential performance with different color schemes for content questions. The sequential/monochrome (Sm) line for the impaired group is relatively flat, as expected, and most other lines for that vision group are very steeply sloped with greater accuracy on the accommodating renditions.

Table 6. Logit Model {AVR}{AS}[a]

Model Terms	Content Questions Coeff.[b]	Z[c]	Odds[d]	Matching Questions Coeff.[b]	Z[c]	Odds[d]
{A}	.487	17.68*	2.65	1.536	33.34*	21.60
{AV}	.126	4.65*	1.29	.312	8.03*	1.87
{AR}	−.059	−2.19*	0.89	−.171	−4.40*	0.71
{AS}						
Qa	−.301	−4.92*	0.55	−.050	−0.59	0.91
Qd	−.190	−3.04*	0.69	−.180	−1.72	0.70
Sm	.337	4.40*	1.96	.720	4.84*	4.22
Sp	.112	1.63	1.25	.160	1.83	1.38
D	−.033	−0.51	0.94	−.421	−6.81*	0.43
B	.048	0.71	1.10	−.084	−1.11	0.85
2V	.027	0.41	1.06	−.146	−2.33*	0.75
{AVR}	.123	4.54*	1.28	.306	7.87*	1.84

[a]Model {AVR}{AS} provides the best balance between accuracy of response estimates, simplicity of model, and degrees of freedom for both content and matching questions. The model estimates response accuracy (**A**) using color vision group (**V**), rendition of color scheme (**R**), and interaction between vision and rendition. The model also includes general levels of accuracy for different schemes (**S**), but not interactions between the schemes and other variables. See Table 5 notes for further explanation of abbreviations of logit model variables.
[b]Coefficients for model terms.
[c]Z scores for model terms.
[d]Odds of a correct response for model terms.
*Significant term at $\alpha = .05$ ($|Z| > 1.96$).

One scheme, sequential/polychrome (Sp), for the impaired group shows a surprising decrease in performance on the accommodating rendition instead of the increase we expected. For the normal vision group, the lines are flatter, as expected, though it becomes apparent that the flat summary line for that group is actually the average of some far less flat lines by scheme and that it is the inconsistency of slope direction that averages out to the flat line.

The lines in Figure 4c are generally higher, indicating the generally higher accuracy of answers to matching questions than to content questions. Again, the slopes are steeper for the impaired group and even the sequential/polychrome (Sp) slopes in the expected direction. For those with normal color vision, the slopes are more consistent and generally opposite in direction to the lines for the impaired group. Given the more complex pattern for content questions (Figure 4b) than for matching questions (Figure 4c), it is not surprising that interaction between vision and scheme showed up more clearly for matching questions in the logit analysis.

Reaction times

Because the effects of the variables showed up clearly in the analysis of correct responses, reaction time was less important as a technique for

Table 7. Odds of Correct Responses: {AVR}

	Estimated	Observed
Content Questions (all schemes aggregated)		
normal vision with confusing rendition	3.89	3.72
impaired vision with confusing rendition	1.44	1.41
normal vision with accommodating rendition	3.01	2.90
impaired vision with accommodating rendition	2.99	2.86
Matching Questions (all schemes aggregated)		
normal vision with confusing rendition	52.77	42.52
impaired vision with confusing rendition	4.47	3.77
normal vision with accommodating rendition	30.75	24.91
impaired vision with accommodating rendition	30.26	24.90

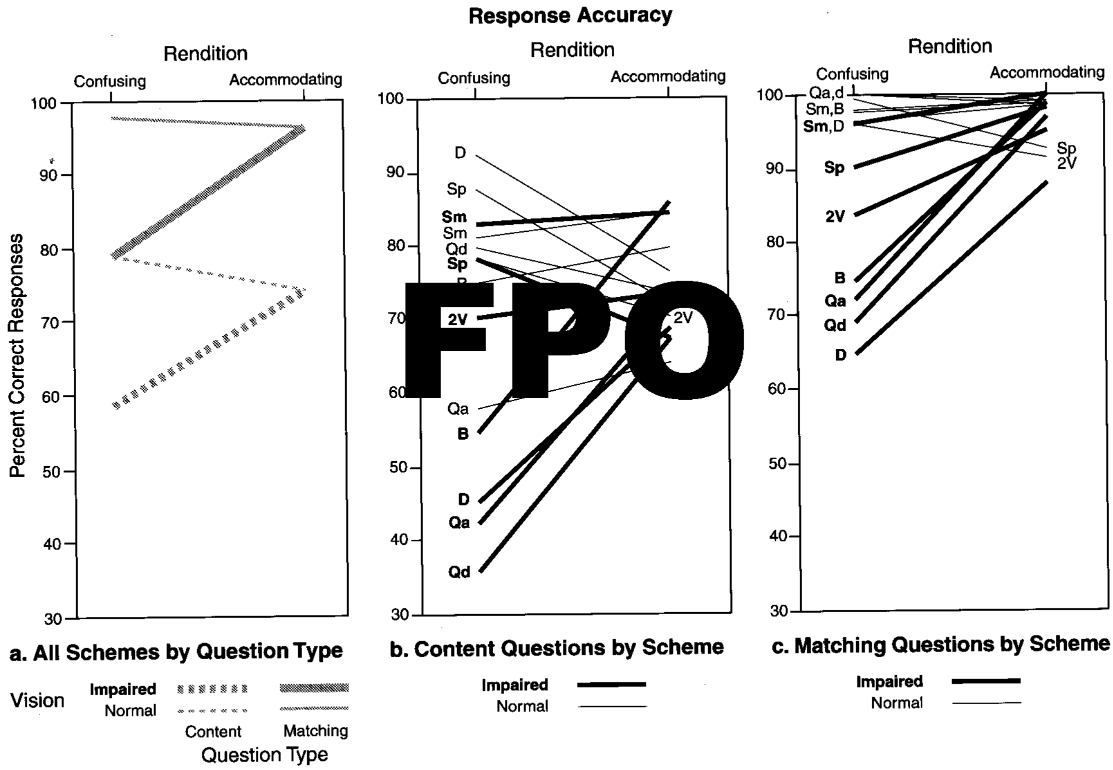


Figure 4. Graphs summarizing response accuracy: (a) all schemes by question type; (b) content questions by scheme; and (c) matching questions by scheme. Sources: See Tables 3 and 4.

determining differences in effects than we had anticipated. Also, the number of incorrect answers for content questions was high enough that it made analysis of the reaction times for correct responses impossible. The repeated measures in our analysis were for rendition and scheme, and to be included, a subject must have responses to all combinations of renditions and schemes. The small number of questions per map and the large number of incorrect responses eliminated many subjects. For the matching questions, the higher numbers of correct responses and higher numbers of questions per map provided sufficient data for analysis with only one normal subject and one impaired subject removed (for a total of sixty-two subjects).

We began by including all variables in the analysis, including test order. The overall effect of test order was not significant, however, and although some two- and three-way interaction terms involving test order were significant, they seemed to result from the small samples involved (limited numbers of questions and maps as well as subjects). We found no enlightening patterns involving test order in the graphs of these complicated interactions. We therefore reran the analysis, omitting the test order variable. The results of the analysis of variance for the matching questions using vision group, rendition, and scheme as the predictor variables are presented in Table 8.

At the .05 significance level, all terms in the analysis were significant except the three-way interaction, and even that term approached significance. In other words, all individual variables in the model were important in explaining differences in reaction times (those with impaired vision take longer than those with normal color vision, accommodating renditions result in shorter reaction times than confusing maps, reaction times differ for one or more schemes), all two-way interactions were important (by vision group, it matters whether a map is accommodating or not; by vision group the reaction times for schemes differ; by rendition, the reaction times for schemes differ), and the three-way interaction has potential (the combination of scheme and rendition may affect differential reaction times for vision groups, but cell frequen-

Table 8. Analysis of Variance: Reaction Times for Matching Questions[a]

	F[b]	df[c]	p[d]
Vision	15.019	1	0.000*
Rendition	16.966	1	0.000*
Scheme	2.748	6	0.013*
Vision × Rendition	41.431	1	0.000*
Vision × Scheme	2.752	6	0.013*
Rendition × Scheme	2.177	6	0.045*
Vision × Rendition × Scheme	1.887	6	0.082

[a]Results of the 2 × 2 × 7 mixed analysis of variance of reaction times for matching questions answered correctly. Vision is a between-subjects variable (subjects had either normal or impaired vision). Color scheme and rendition are within-subjects variables (reaction times of each of sixty-two subjects were measured at all levels of these variables). In Figure 5b, the Vision × Rendition interaction is represented by gray lines and the Vision × Rendition × Scheme interaction is represented by black lines.
[b]F is ratio of between- and within-group variances in reaction times.
[c]df is degrees of freedom.
[d]p is probability.
*Mean reaction times are significantly different at $p = .05$.

cies for this analysis are not sufficient to reach significance). It is interesting that the primary variables and their interactions give results exactly as expected, and that the vaguer expectation about the effects of schemes is paralleled by vague results.

Reaction time data are summarized for all schemes in Figure 5a. Reaction times for the accommodating rendition, compared to those for the confusing rendition, are lower for the impaired group and slightly higher for the normal group on both types of questions. The interactions are not nearly as striking, however, as in the accuracy graph in Figure 4a.

A graph illustrating the other relevant reaction-time interactions is shown in Figure 5b. Only matching questions are involved and the results over all schemes (gray lines) are shown on the same graph as the results for each individual scheme (black lines). The results for all schemes (gray lines) indicate that the normal-vision group has distinctly lower reaction times than the impaired group and that times for the normal group vary only slightly overall between the confusing and accommodating renditions. The impaired group, on the other hand, have longer reaction times and a steep decline between the confusing and accommodating renditions. For individual schemes (black lines), average reaction times for the normal-vision group are all below those for the impaired group. For the normal group, one scheme (diverging [D]), shows a relatively steep decline, and three (two-variable [2V], qualitative/area [Qa], and balance [B]) show relatively steep inclines. For the impaired group, however, all lines decline, most of them more steeply than any changes for the normal group. For the impaired group, even the sequential/monochrome (Sm) scheme results in lower reaction times when the colors are accommodating.

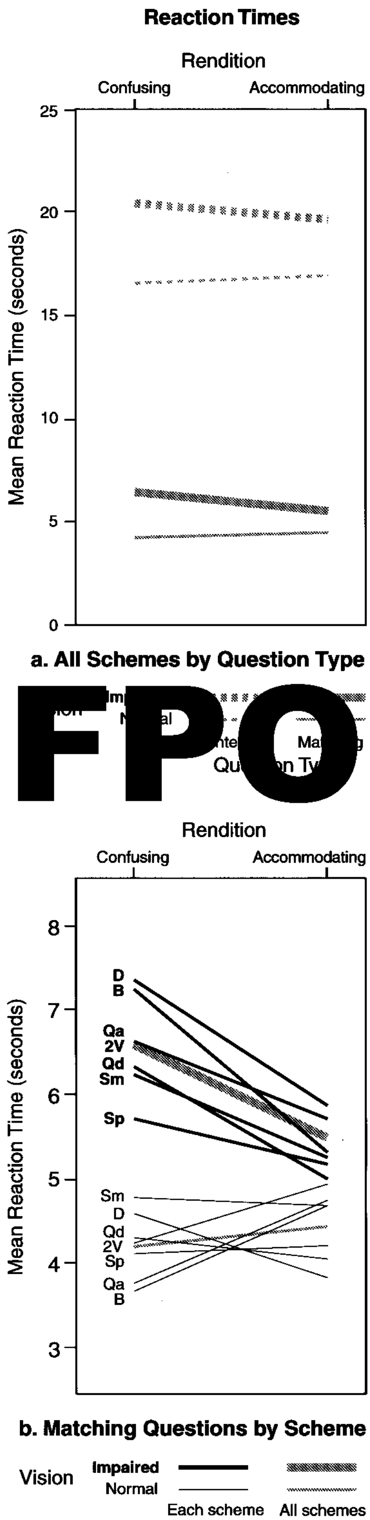
Selections between Renditions

The normal vision group chose the accommodating renditions of color schemes 56 percent of the time when asked which rendition within each pair was easier or better to use. The color-impaired group chose the accommodating renditions 87 percent of the time. As expected, the binomial test to determine if percentages differed from 50 percent (the figure one would expect if there is no bias in the selection) showed that those with normal vision did not choose one rendition over the other. Those with red-green impairments chose the accommodating rendition significantly more than 50 percent of the time.

Table 9 shows scheme-by-scheme results for this part of the testing. For those with color impairments, all seven of the choices favored the accommodating rendition and five were significant. For those with normal vision, four of seven choices favored the accommodating rendition and two of those were significant. Of the three choices favoring the confusing rendition, one was significant.

Discussion

Given the results of this experiment, there is no question that adjusting the colors on maps to accommodate the red-green impaired is important in allowing people in this group to use



maps effectively. The color confusions that are found under the much more controlled and artificial conditions of psychophysical testing of people with red-green impairments do carry over into conditions under which people look at maps (at least on CRTs). The confusions are especially obvious in the results for the matching task, but the content questions are also answered more accurately when colors are accommodating.

The reaction time results did not conform entirely to expectations. Even though people with red-green impairments have shorter reaction times when using the accommodating renditions of the maps (compared to confusing renditions), their reaction times are still longer than those for people with normal color vision. For people with normal color vision, the reaction time is slightly longer on average with accommodating renditions, with some schemes eliciting longer and others shorter times than the confusing renditions. Although we predicted no change in reaction time for the normal-vision group, it would have been reasonable to predict that their reaction times would be significantly worsened (perhaps "worsened to match those of the impaired group"), since adjustment of colors means choosing from a far more restricted color space, and the smaller the color space, the less opportunity for selecting highly contrasting colors. Fortunately, subjects with normal color vision did not take drastically longer with accommodating renditions compared to the confusing ones.

But why do those with red-green impairments continue to take more time than normals when accommodating colors are used? Why did they not perform even better than the normal group in that a restricted overall color space throughout life might have led to sharper abilities? An incidental conversation (not directly connected to the testing in the project) with a person with a red-green impairment revealed a very plausible explanation for the continued slowness of response. Shown a map on which several categories of values were represented by the very same color, this person was convinced

Figure 5. Graphs summarizing reaction times for correct answers: (a) all schemes by question type, and (b) matching questions by scheme. Sources: (a) Tables 3 and 4; (b) original data used in analysis of variance.

Table 9. Colors Chosen as Easier/Better

	Vision	Percent Choosing Rendition	
		Confusing	Accommodating
All Schemes[a]	Impaired	13	87*
	Normal	44	56
Each Scheme[b]			
Qualitative/area (Qa)	Impaired	3	97*
	Normal	66	34
Qualitative/dot (Qd)	Impaired	3	97*
	Normal	19	81*
Sequential/monochrome (Sm)	Impaired	6	94*
	Normal	53	47
Sequential/polychrome (Sp)	Impaired	38	62
	Normal	75*	25
Diverging (D)	Impaired	3	97*
	Normal	16	84*
Balance (B)	Impaired	3	97*
	Normal	38	62
Two-Variable (2V)	Impaired	38	62
	Normal	44	56

[a]Total number of choices by all subjects in each vision group was 224.
[b]Total number of choices by all subjects in each vision group for each scheme was 32.
*Percentage significantly greater than 50 percent at $\alpha = .05$.

that they were differentiable colors (to those with normal vision) that he could not distinguish. It was not until he was told that the colors really were identical and had been selected to illustrate what it is like to see a map on which categories are not discriminable that he discontinued his efforts to detect some difference between them. In other words, he had learned not to trust what he saw; there must be a difference, and if he looked long enough perhaps he could detect it. It is quite possible that, in general, those with color-vision impairments have become so accustomed to having difficulty with colors that they do not trust their first impressions, and they take longer in color discrimination tasks than do people with normal color vision. We must remember, too, that in past experience some colors that looked identical to those with normal vision have looked different to the impaired. Whether colors look alike or different, then, the color-vision impaired may have acquired distrust of their reactions.

A more likely clue, however, comes from color vision research that compares wavelength discrimination by people with normal and impaired color vision (Pokorny et al. 1979:198). Those with impairments are less sensitive across the entire spectrum, which is not surprising because all three cones are sensitive to some degree to virtually all visible wavelengths. We should have expected, then, that all color distinctions would be more difficult for subjects with impaired color vision even though the adjusted colors would be a great improvement over colors that fell close to confusion lines.

Another hint at the milieu in which these results were obtained comes from consideration of the results for questions on the background questionnaire, which included several questions about map usage and attitudes. The results were not considered major ones in this study, but we did run chi-square tests on frequency data and Mann-Whitney U tests on attitude ratings. Although none of the questions resulted in significant differences between those with normal color vision and those with red-green impairments (n was only thirty-two for each group), the direction in every case except one (other map use experience) showed that those with red-green impairments were less experienced with maps and held more negative attitudes toward them. The differences, if real, are small enough to require a larger sample size to confirm with a significance test, but one would expect a group with less experience and more negative attitudes to respond more slowly and with less accuracy when encountering new maps. It is plausible, then, that both nature and nurture play a role in the persistence of longer reaction times for subjects with impaired color vision.

Our results also show that accommodating renditions of color schemes lower slightly the accuracy of those with normal color vision. The difference is small, yet if we were to weight the matching accuracy figures according to the proportion of the population with normal color vision and with red-green impairments, the high gains with adjustment for the red-green group (4 percent of the population) are more than offset by the small deterioration that occurs for the much larger normal vision group (96 percent of the population). In other words, for the population as a whole, the renditions with potentially confusing colors used in this study result in slightly better performance than the accommodating renditions. The accommodating renditions, which are obviously worthwhile for the color impaired, need not be used for those with normal color vision, given the flexibility of modern technology. Alternatively, slight improvements over our color selections may well restore or even improve the performance of the normal color-vision group as well as accommodate the color impaired and thus improve overall population performance. Improved choices of map colors can be shared among mapmakers, saving considerable design time while serving users with both normal and red-green impaired vision.

The differences in accuracy and reaction time for different color schemes did not interact strongly with color-vision group and rendition. This result is not surprising when one considers that the quantitative/monochrome scheme is the only one on which the two vision groups would generally be expected to perform equally well on most any rendition of the scheme. We tried to use the one color area (reds) that potentially gives problems to protanopes and align the selections on a deuteranopic confusion line (eliminating saturation differences for them) to produce a difference between renditions even on this scheme. The reddish colors we used were probably not, however, in the most confusing part of color space (or possibly we had too few protanopes), and the lightness distinctions were probably distinguishable enough to deuteranopes without having saturation differences. In any event, the red-green group did well with both renditions. With all other schemes, we would expect that confusing and accommodating renditions would elicit differential performance on the part of the impaired. Even though results showed otherwise (the accommodating renditions of the sequential/polychrome and two-variable schemes did not result in improvements), the differences from scheme to scheme relative to group performances were not on the whole a strong effect.

The differences in accuracy between schemes may seem to suggest that some schemes are inherently better than others because they require no accommodation, are more conducive to accommodation, or elicit better performance on the part of both vision groups. Although there were some schemes that gave less difficulty to impaired users (e.g., sequential/polychrome), resulted in larger improvement (e.g., balance), or worked well for both vision groups (sequential/monochrome), we cannot necessarily label these as "better schemes." The scheme chosen to represent data depends primarily on the nature of the data represented, or the characteristics of it that we want to emphasize. For example, although we can use a sequential/monochrome scheme for any data to which we could apply a diverging scheme, the message of the map would be different. The question of whether the apparent differences that appear in our results are real or not is an interesting and worthwhile one and would require samples of maps representing each scheme (we used scheme as a way of developing a stratified sample of maps, but we had only one map in each stratum). It would also require a far larger sample of questions than we used in this experiment. A troublesome part of designing such a test would be developing appropriate questions; can we, for example, ask about areas above and below the mean on a map with a sequential/monochrome scheme, which normally has no clear indication of the mean? But even if the results of such an exercise would not tell us to use one scheme over another, perhaps it would alert us to ways in which we could modify map design and color selections to yield improvements. For example, perhaps stating the mean value on a map with a sequential/monochrome scheme would lessen the differences between it and a map with a diverging scheme. At any rate, we cannot draw conclusions about which scheme is best from the sample of one map and four content questions that we used for each scheme in this experiment.

Reflection on the design of a complex test inevitably suggests ways it might have been improved. In our case, the particular reddish hue used on the potentially confusing rendition of the sequential/monochrome scheme was probably not the best (i.e., most confusing) that could have been used. For the two-variable scheme, we used a light-to-dark arrangement on one rendition and a dark-to-light arrangement on the other, and careful consideration of which to use may help in developing an improved accommodating rendition for the CRT. The use of multiple-choice questions necessitated correction of reaction times for position of the answer given, and even if the correction had precisely eliminated the time to press the key down to the selected answer, it would not have corrected for the differential processing in eliminating, say, three wrong choices (answer *d* recognized as correct) versus no wrong choices (answer *a* recognized as correct). Fortunately, only one pair of correct answers for corresponding versions was that widely separated and most (nine of fourteen pairs) were in the same position or only one choice apart. Fortunately, too, the corrected data contained considerable information because a substantial portion of reaction time came before first pressing the arrow key, and that part of the time was unaffected by any correction. Using touch-screen technology and alternate question formats could eliminate the need for correction altogether, however.

Another issue is relevance of our results for the production of printed maps. Although the study involved CRT technology, the maps were similar to those we encounter on paper, with similar numbers of colors and real data. In fact, all but one of the maps were adapted from paper maps. The exact color selections for the CRT displays are not those that would be used on paper maps because of differences in the media; paper maps have colors produced by reflected light, CRT colors emanate from light sources. But the effects of color adjustment should be similar, and the test results suggest that it would be worthwhile to develop prototype color sets for paper maps as well. Considerable progress in that direction is represented by the work of Brewer, MacEachren, and Pickle (1995). With the desk-top color separation software commonly available today, we have unprecedented flexibility in color selection and separation construction for printed maps in small as well as large projects.

Finally, adding criteria to our already complex repertoire of cartographic rules for selecting colors for maps brings up the issue of aesthetics. The experiment reported here did not explicitly consider aesthetics (subjects were never asked which renditions "looked good"), yet the experience of attempting to select colors for seven different map pairs was instructive. In our judgment, the accommodating renditions of our maps were generally the more aesthetically pleasing of the pairs. That observation does not necessarily mean that accommodating maps are likely to be more aesthetically pleasing than the ones we would make without consideration of people with impaired color vision, however. The constraints to produce the potentially confusing renditions were just as stringent, and presumably color selections freed of these constraints would provide more leeway for deriving aesthetically pleasing sets. Indeed the severity of constraints imposed by the scheme itself (e.g., choosing colors that look appropriate for a balance scheme) may be greater than those imposed by consideration of color-vision impaired users. In short, we see no reason to expect noticeable aesthetic deterioration when choosing colors that accommodate color impairments.

Appendix A: Details on Colors Used in the Testing

The colors for each map used in the testing are shown in graphic form here (Figure A1-8). The legend arrangement appears in the upper left, with arrows connecting categories that were potentially troublesome (on the confusing rendition) to the red-green impaired. A verbal description of each color on the confusing map is followed by the *xyY* and RGB values. The *x* and *y* are the CIE chromaticity coordinates, *Y* is luminance factor (a measure of lightness), and R, G, and B are the red, green, and blue levels specified when producing the color on the CRT. The same information is then given for the accommodating rendition. On the CIE diagrams just below the numerical values, each color is plotted in its *xy* location, with circle size scaled according to the *Y* value. The confusion lines for protanopes and deuteranopes (Figure 2) are plotted on each CIE diagram for reference. Only one rendition of the practice map (with one pair of potentially confusing colors) was used; it is included first here because it was the first map encountered by all subjects. The re-

People per Square Mile by Province	Color name	Confusing Rendition x	y	Y	R	G	B
> 1000	light orange	.476	.417	54	255	170	130
750-1000	orange	.595	.396	24	220	100	70
500-749	dark orange	.668	.334	6	150	60	20
* 250-499	dark red	(not measurable)			80	30	0
* 0-249	dark red (repeated)	(not measurable)			80	30	0

*Potentially confusing color pairs

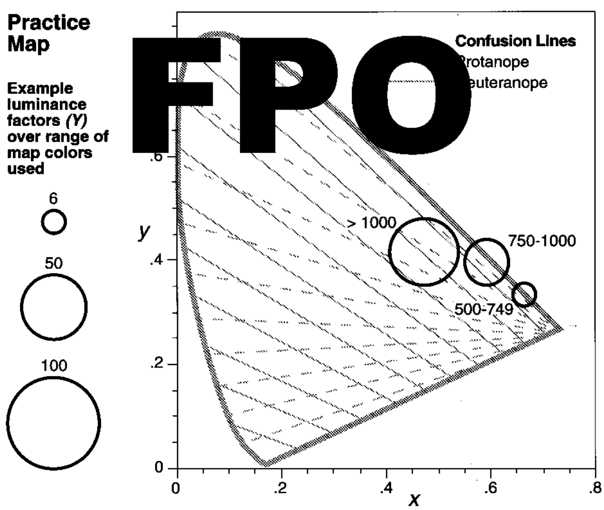


Figure A-1. Practice map: China, Population Density, 1982.

Building Use	Color name	Confusing Rendition x	y	Y	R	G	B	Color name	Accommodating Rendition x	y	Y	R	G	B
Residential	blue	.160	.258	18	0	140	170	blue	.152	.062	12	0	0	255
Commercial	magenta	.403	.200	18	200	0	170	light green	.255	.589	59	120	220	120
Entertainment	green	.279	.654	37	115	185	0	greenish-white	.247	.362	100	190	255	220
Public	red-orange	.590	.398	36	255	110	0	gray-magenta	.415	.302	21	180	100	140
Other	yellow	.448	.514	36	188	158	0	dark green	(not measurable)			0	90	0

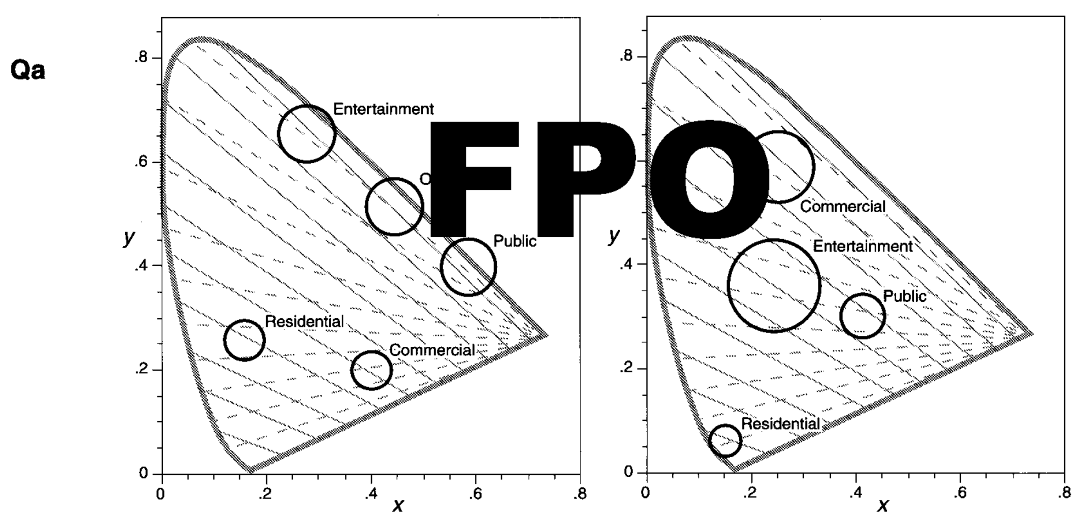


Figure A-2. Qualitative/area scheme (Qa): Building Use, San Francisco, California, 1913.

Each Dot Represents 5000 Animals	Color name	Confusing Rendition x	y	Y	R	G	B	Color name	Accommodating Rendition x	y	Y	R	G	B
Cattle	dark red	(not measurable)			100	0	0	blue	.155	.115	8	0	90	180
Hogs	purple	.299	.196	18	160	100	180	red	.624	.378	12	170	70	80
Sheep	blue	.161	.267	17	40	140	170	white	.287	.328	48	180	180	180

Qd

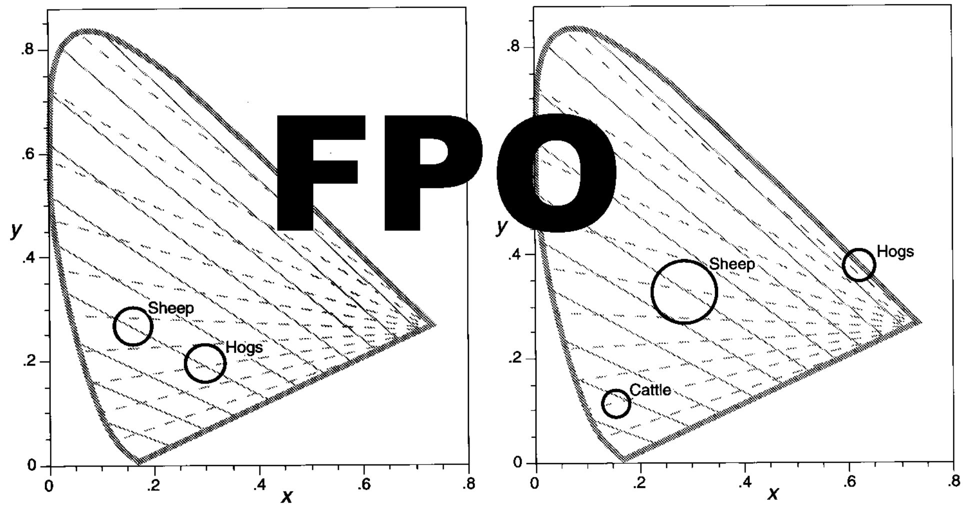


Figure A-3. Qualitative/dot scheme (Qd): Livestock Inventory, South Dakota, 1982.

Harvest by County (Million Board Feet)	Color name	Confusing Rendition x	y	Y	R	G	B	Color name	Accommodating Rendition x	y	Y	R	G	B
400-999	white	.292	.300	72	240	220	240	white	.285	.320	99	255	255	255
100-399	light gray-magenta	.311	.306	54	220	190	210	light gray	.294	.338	57	200	200	200
50-99	gray-magenta	.334	.299	28	190	150	180	gray	.337	.420	19	150	150	150
0.1-49	dark gray-magenta	.357	.281	13	160	115	155	dark gray	.394	.567	7	110	110	110
No Harvest	dark magenta	.395	.266	6	130	80	125	near black	(not measurable)			80	80	80

Sm

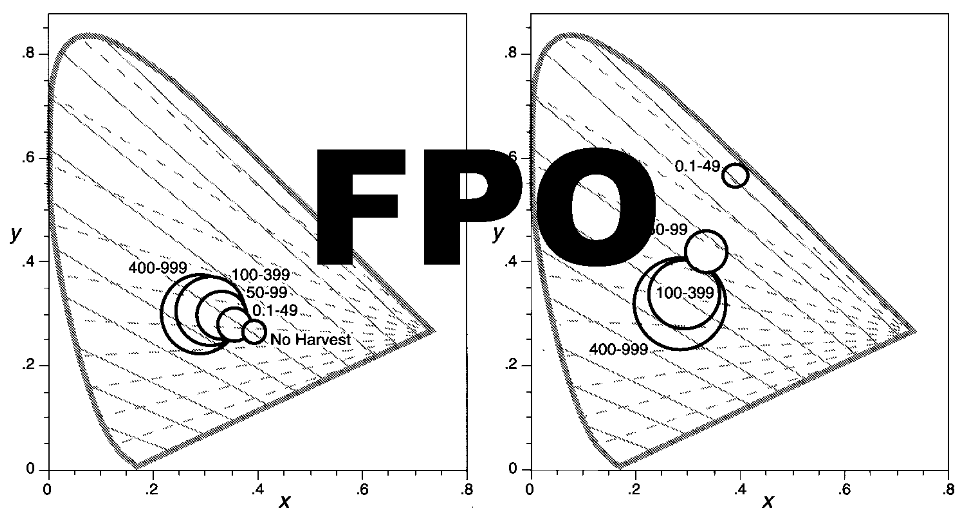


Figure A-4. Sequential/monochrome scheme (Sm): Timber Harvest, Pacific Northwest, 1981.

Value of Gross Sales Per Acre	Color name	Confusing Rendition x	y	Y	R	G	B	Color name	Accommodating Rendition x	y	Y	R	G	B
Low	purple	.299	.196	18	160	100	180	dark blue	.152	.098	7	50	90	190
Medium-Low	blue	.161	.267	17	40	140	170	blue	.182	.224	23	90	145	195
Medium	green	.318	.619	34	130	175	70	green-blue	.187	.349	40	100	215	210
Medium-High	yellow-orange	.519	.456	34	220	150	60	light green	.246	.448	66	170	255	200
High	pink	.420	.352	53	255	170	170	yellow	.388	.512	80	255	255	150
No Data	dark gray	(not measurable)			70	70	100	dark gray	(not measurable)			70	80	100

Sp

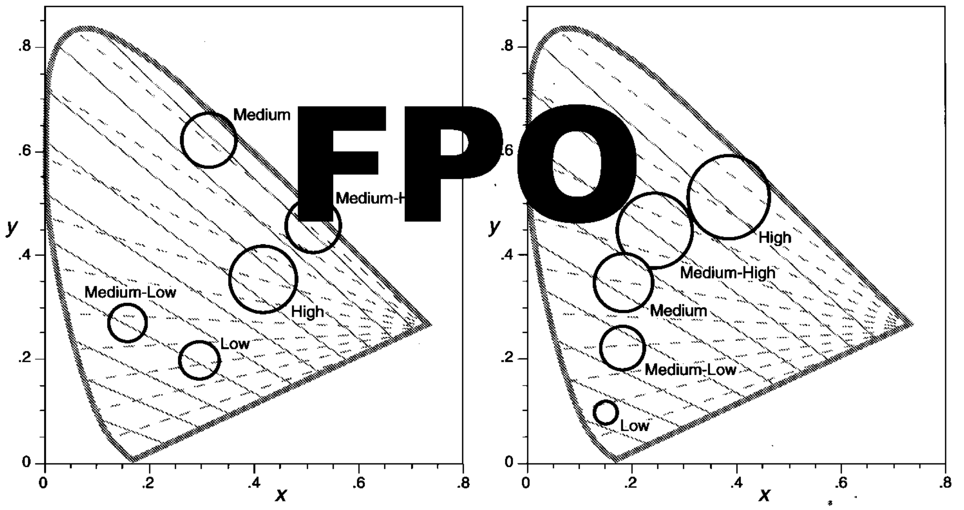


Figure A-5. Sequential/polychrome scheme (Sp): Sales from Improved Farmland, Southern Ontario Townships, 1981.

Change in the Number of Jobs by County	Color name	Confusing Rendition x	y	Y	R	G	B	Color name	Accommodating Rendition x	y	Y	R	G	B
Gain > 1000	dark green	.175	.756	13	0	145	0	dark blue	.153	.064	9	50	50	250
Gain 500-999	green	.268	.658	30	115	185	0	blue	.205	.164	29	150	150	250
Gain 1-499	yellow-green	.390	.554	61	210	215	0	light blue	.255	.261	65	210	210	250
Loss 1-499	yellow-orange	.480	.484	59	255	190	0	light orange-red	.413	.360	47	255	180	180
Loss 500-999	orange	.603	.385	31	255	110	0	orange-red	.519	.380	28	230	130	130
Loss > 1000	red	.646	.353	13	190	60	0	red	.658	.342	13	200	0	0
No Data	gray	.328	.411	19	150	150	150	dark gray	(not measurable)			90	90	90

D

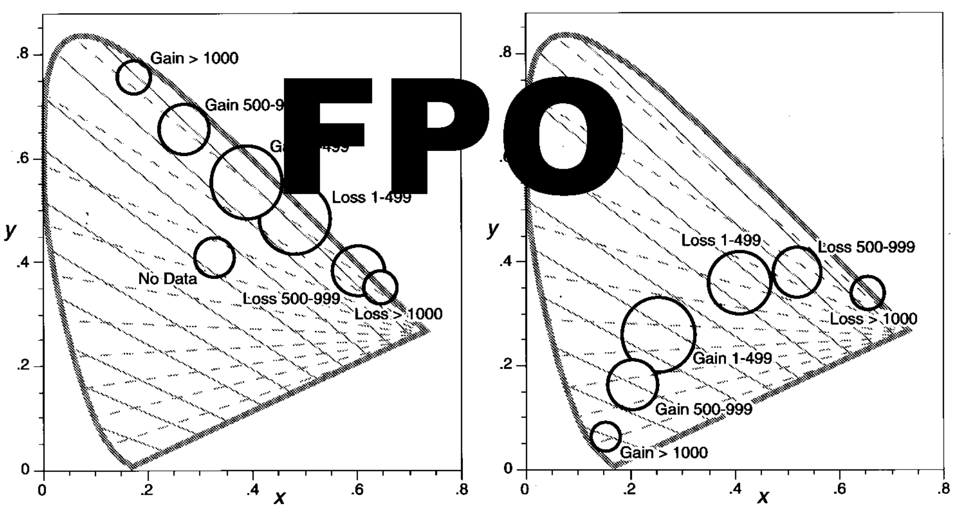


Figure A-6. Diverging scheme (D): Change in Manufacturing Jobs, 1970 to 1982, Georgia.

Ratification Vote by County	Confusing Rendition: Color name	*x*	*y*	*Y*	R	G	B	Accommodating Rendition: Color name	*x*	*y*	*Y*	R	G	B
All For	blue	.165	.126	7	80	90	175	orange	.607	.389	10	165	85	85
Majority For	gray-blue	.206	.178	13	108	108	183	gray-orange	.467	.373	15	167	110	123
Tie	gray-purple	.278	.217	15	145	105	170	gray	.343	.321	26	175	140	165
Majority Against	magenta	.372	.210	13	170	70	160	gray-blue	.214	.180	20	135	130	210
All Against	magenta-red	.639	.321	8	160	50	110	blue	.158	.073	12	85	85	255
No Data	dark gray	.295	.310	8	100	90	120	dark gray	(not measurable)			90	80	110

B

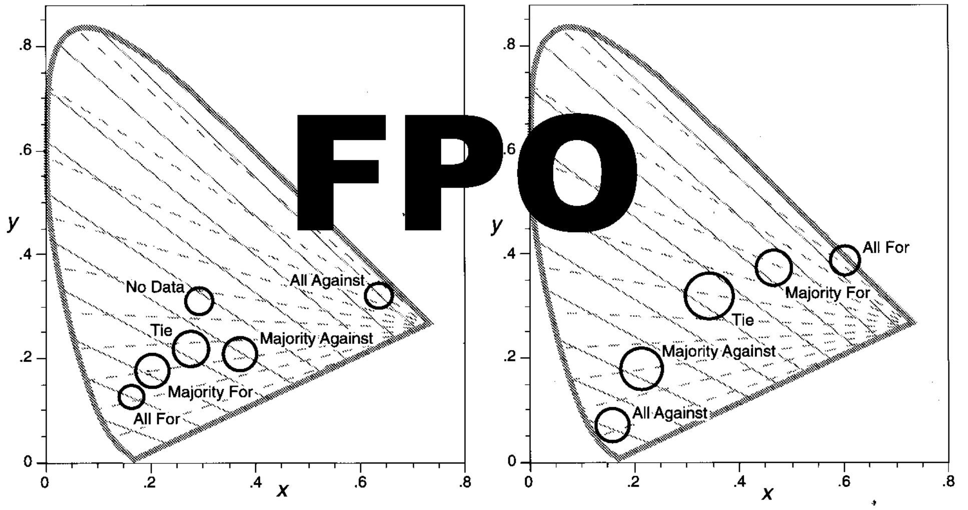


Figure A-7. Balance scheme (B): Ratification of the U.S. Constitution, South Carolina, 1790.

Non-Employment Income, 1983	Net Migration, 1975-1980	Confusing Rendition: Color name	*x*	*y*	*Y*	R	G	B	Accommodating Rendition: Color name	*x*	*y*	*Y*	R	G	B
Low	Out	green	.193	.726	29	0	177	0	blue	.166	.092	17	96	96	255
Med.	Out	green-blue	.161	.307	16	0	135	155	light blue	.210	.202	43	147	167	240
High	Out	purple-blue	.179	.078	5	90	60	180	white	.280	.319	74	210	210	210
Low	Balanced	yellow-green	.324	.614	47	148	198	0	dark blue	.184	.209	7	60	90	190
Med.	Balanced	gray	.303	.369	29	150	150	150	gray	.306	.353	17	128	128	138
High	Balanced	magenta	.430	.213	16	195	0	155	light orange	.409	.403	43	210	160	135
Low	In	yellow	.397	.547	79	230	230	0	dark gray	(not measurable)			80	80	100
Med.	In	orange	.509	.464	47	240	160	0	dark orange	.603	.389	7	135	75	20
High	In	orange-red	.617	.377	28	240	85	85	orange	.602	.391	17	185	85	75

2V

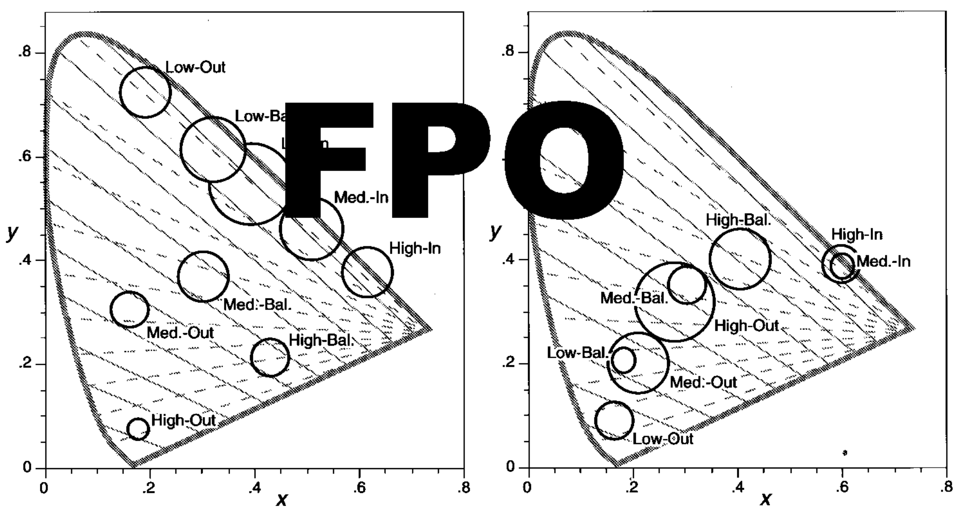


Figure A-8. Two-variable scheme (2V): Non-employment Income and Net Migration, New Mexico.

maining maps appear in the order in which they are described in the text; their sequencing in the testing is described in the test order section of the text.

Acknowledgments

This research was funded by NSF Grant No. SES-8712109. The assistance of the Ameritech Foundation is also gratefully acknowledged. Our thanks to the many students who influenced or worked on this project with us, especially Daniel Heist, James P. Moore, Michael H. Martel, and Charles Rader, and to the subjects who volunteered for the testing. We also thank the Geography Department office staff at Michigan State University for handling all scheduling of subjects.

References

Agoston, George A. 1987. *Color Theory and Its Application in Art and Design,* 2nd ed. New York: Springer-Verlag.

Birch, Jennifer. 1973. Dichromatic Convergence Points Obtained by Subtractive Colour Matching. *Vision Research* 13:1755–65.

Brewer, Cynthia A. 1989. The Development of Process-printed Munsell Charts for Selecting Map Colors. *The American Cartographer* 16:269–78.

———. 1991. The Prediction of Surround-Induced Changes in Map Color Appearance. Ph.D. dissertation, Department of Geography, Michigan State University.

———. 1994. Color Use Guidelines for Mapping and Visualization. In *Visualization in Modern Cartography,* ed. A. M. MacEachren and D. R. F. Taylor, pp. 123–47. New York: Pergamon Press.

Brewer, Cynthia A.; MacEachren, Alan M.; and Pickle, Linda W. 1995. Evaluation of Map Color Schemes for the NCHS Mortality Atlas. In *Cognitive Aspects of Statistical Mapping,* Cognitive Methods Staff, Working Paper Series No. 18, ed. Linda Williams Pickle and Douglas J. Herrmann, pp. 191–200. Hyattsville, MD: National Center for Health Statistics, Office of Research and Methodology. Referenced with authors' permission.

Campbell, John. 1984. *Introductory Cartography.* Englewood Cliffs, NJ: Prentice Hall.

Cappon, L. J.; Petchenik, B. B.; and Long, J. H., eds. 1976. *Atlas of Early American History, The Revolutionary Era, 1760–1790.* Princeton, NJ: Princeton University Press.

Cressie, Noel A. C. 1993. *Statistics for Spatial Data,* rev. ed. New York: John Wiley & Sons.

Cuff, David J. 1973. Shading on Choropleth Maps: Some Suspicions Confirmed. *Proceedings, Association of American Geographers* 5:50–54.

Dent, Borden D. 1993. *Principles of Thematic Map Design,* 3rd ed. Dubuque, IA: Wm. C. Brown Publishers.

Donley, M. W.; Allan, S.; Caro, P.; and Patton, C. P. 1979. *Atlas of California.* Culver City, CA: Pacific Book Center.

Eastman, J. Ronald. 1986. Opponent Process Theory and Syntax for Qualitative Relationships in Quantitative Series. *The American Cartographer* 13:324–34.

Estevez Uscanga, Oscar. 1979. On the Fundamental Data-base of Normal and Dichromatic Color Vision. Ph.D. dissertation, University of Amsterdam.

Fletcher, Robert, and Voke, Janet. 1985. Defective Colour Vision: Fundamentals, Diagnosis and Management. Bristol, U.K.: Adam Hilger Ltd.

Heyn, Barbara. 1984. An Evaluation of Map Color Schemes for Use on Cathode Ray Tubes. Master's thesis, Department of Geography, University of South Carolina.

Hodler, Thomas W., and Schretter, Howard A. 1986. *The Atlas of Georgia.* Athens, GA: Institute of Community and Area Development, University of Georgia.

Ishihara, Shinobu. 1984. *The Series of Plates Designed as a Test for Colour-blindness.* Tokyo: Kanehara & Co.

Judd, Deane B., and Wyszecki, Gunter. 1975. *Color in Business, Science, and Industry,* 3rd ed. New York: John Wiley & Sons.

Keddie, P. D., and Mage, J. A. 1985. *Southern Ontario Atlas of Agriculture, Contemporary Patterns and Recent Changes.* Occasional Papers in Geography, No. 7. Guelph, Ontario: Department of Geography, University of Guelph.

Kimerling, A. Jon. 1980. Color Specification in Cartography. *The American Cartographer* 7:139–54.

———. 1981. Process Color Diagrams. *The American Cartographer* 8:180–81.

———, and Jackson, P. L., eds. 1985. *Atlas of the Pacific Northwest,* 7th ed. Corvallis, OR: Oregon State University Press.

Knoke, David, and Bohrnstedt, George W. 1994. *Statistics for Social Data Analysis,* 3rd ed. Itasca, IL: F. E. Peacock Publishers.

Mersey, Janet E. 1980. An Analysis of Two-Variable Choropleth Maps. Master's thesis, Department of Geography, University of Wisconsin-Madison.

———. 1984. Color on Choropleth Maps: The Effects of Color Scheme and Number of Classes on Map Communication. Ph.D. dissertation, Department of Geography, University of Wisconsin-Madison.

———. 1990. *Colour and Thematic Map Design:*

The Role of Colour Scheme and Map Complexity in Choropleth Map Communication. Cartographica. Monograph 41.

Meyer, Gary W., and Greenberg, D. P. 1988. Color-Defective Vision and Computer Graphics Displays. *IEEE Computer Graphics and Applications* 8:28–40.

Moore, James B. 1988. Color Vision Deficiencies and Printed Color Maps: Is There a Problem? Seminar report, Department of Geography, Michigan State University. Referenced with author's permission.

Nimeroff, I. 1970. Deuteranopic Convergence Point. *Journal of the Optical Society of America* 60:966–69.

Olson, Judy M. 1981. An Evaluation of Spectrally Encoded Two-Variable Maps. *Annals of the Association of American Geographers* 71:259–76.

———. 1987. Color and the Computer in Cartography. In *Color and the Computer,* ed. H. John Durrett, pp. 205–20. Orlando, FL: Academic Press.

Pannell, Clifton W., and Ma, Lawrence J. C. 1983. *China: The Geography of Development and Modernization.* Washington: V. H. Winston and Sons.

Pickford, R. W. 1965. The Genetics of Colour Blindness. In *Colour Vision: Physiology and Experimental Psychology* (Ciba Foundation Symposium), ed. A. V. S. de Reuck and Julie Knight, pp. 228–44. Boston: Little, Brown and Co.

Pitt, F. H. G. 1935. *Characteristics of Dichromatic Vision.* Medical Research Council, Reports of the Committee upon the Physiology of Vision, 14. London: His Majesty's Stationery Office.

Pokorny, Joel; Smith, Vivianne C.; Verriest, Guy; and Pinckers, A. J. L. G. 1979. *Congenital and Acquired Color Vision Defects.* New York: Grune and Stratton.

Raftery, Adrian E. 1986. Choosing Models for Cross-Classifications. *American Sociological Review* 51:145–46.

Raisz, Erwin. 1962. *Principles of Cartography.* New York: McGraw-Hill Book Company.

Regan, B. C.; Reffin, J. P.; and Mollon, J. D. 1994. Luminance Noise and the Rapid Determination of Discrimination Ellipses in Colour Deficiency. *Vision Research* 34:1279–99.

Robertson, Philip K. 1988. Visualizing Color Gamuts: A User Interface for the Effective Use of Perceptual Color Spaces in Data Displays. *IEEE Computer Graphics & Applications* 8:50–64.

Robinson, Arthur H. 1953. *Elements of Cartography.* New York: John Wiley & Sons.

———. 1960. *Elements of Cartography,* 2nd ed. New York: John Wiley & Sons.

Robinson, Arthur H., and Sale, Randall D. 1969. *Elements of Cartography,* 3rd ed. New York: John Wiley & Sons.

Robinson, Arthur H.; Sale, Randall D.; and Morrison, Joel. 1978. *Elements of Cartography,* 4th ed. New York: John Wiley & Sons.

Robinson, Arthur H.; Sale, Randall D.; Morrison, Joel; and Muehrcke, Phillip C. 1984. *Elements of Cartography,* 5th ed. New York: John Wiley & Sons.

Robinson, Arthur H.; Morrison, Joel L.; Muehrcke, Phillip C.; Kimerling, A. Jon; and Guptill, Stephen C. 1995. *Elements of Cartography,* 6th ed. New York: John Wiley & Sons.

Rushton, W. A. H. 1975. Visual Pigments and Color Blindness. *Scientific American* 232:64–74.

Snodgrass, Joan Gay; Levy-Berger, Gail; and Haydon, Martin. 1985. *Human Experimental Psychology.* New York: Oxford University Press.

SPSS. 1988. *SPSS^X User's Guide,* 3rd ed. Chicago: SPSS.

Travis, David. 1991. *Effective Color Displays: Theory and Practice.* London: Academic Press (Harcourt Brace Jovanovich).

Trumbo, Bruce E. 1981. A Theory for Coloring Bivariate Statistical Maps. *The American Statistician* 35:220–26.

U.S. Bureau of the Census. 1982. *U.S. Agricultural Census, 1982.* Washington.

Vos, J. J., and Walraven, P. L. 1971. On the Derivation of the Foveal Receptor Primaries. *Vision Research* 11:799–818.

Vos, J. J.; Estevez, O.; and Walraven, P. L. 1990. Improved Color Fundamentals Offer a New View on Photometric Additivity. *Vision Research* 30:937–43.

Wainer, Howard, and Francolini, Carl M. 1980. An Empirical Inquiry Concerning Human Understanding of Two-Variable Color Maps. *The American Statistician* 34:81–93.

Ware, Colin. 1988. Color Sequences for Univariate Maps: Theory, Experiments, and Principles. *IEEE Computer Graphics and Applications* 8:41–49.

Williams, Robert L. 1956. *Statistical Symbols for Maps: Their Design and Relative Values.* New Haven, CT: Map Laboratory, Yale University.

Wright, W. D. 1952. The Characteristics of Tritanopia. *Journal of the Optical Society of America* 42:509–21.

Wyszecki, Gunter, and Stiles, W. S. 1982. *Color Science: Concepts and Methods, Quantitative Data and Formulae,* 2nd ed. New York: John Wiley & Sons.

Correspondence: Department of Geography, Michigan State University, East Lansing, MI 48824-1115 (Olson); Department of Geography, The Pennsylvania State University, University Park, PA 16802-5011 (Brewer).