

The End of the Rainbow? Color Schemes for Improved Data Graphics

PAGES 385, 391

Modern computer displays and printers enable the widespread use of color in scientific communication, but the expertise for designing effective graphics has not kept pace with the technology for producing them. Historically, even the most prestigious publications have tolerated high defect rates in figures and illustrations [Cleveland, 1984], and technological advances that make creating and reproducing graphics easier do not appear to have decreased the frequency of errors. Flawed graphics consequently beget more flawed graphics as authors emulate published examples. Color has the potential to enhance communication, but design mistakes can result in color figures that are less effective than gray scale displays of the same data.

Empirical research on human subjects can build a fundamental understanding of visual perception [Ware, 2004] and scientific methods can be used to evaluate existing designs, but creating effective data graphics is a design task and not fundamentally a scientific pursuit. Like writing well, creating good data graphics requires a combination of formal knowledge and artistic sensibility tempered by experience: a combination of “substance, statistics, and design” [Tufte, 1983, p. 51].

Unlike writing, however, proficiency in creating data graphics is not a main component of secondary or postsecondary education. This article provides some concrete suggestions to help geoscientists use color more effectively when creating data graphics. The article explains factors to consider when designing for color-blind viewers, offers some example color schemes, and provides guidance for constructing and selecting color schemes in the form of design patterns. Because scientific authors frequently misuse spectral color, the particular focus here is on better alternatives to such schemes.

Designing for Color-blind Viewers

One of the commonly overlooked considerations in scientific data graphics is perception by individuals with color-deficient vision. The significance of “color-blindness” increases for geoscience publications whose readers are disproportionately male because, while 0.4% of women exhibit some form of color-vision deficiency, the figure is approximately 8% for Caucasian men. Among the predominantly male readership of *Eos*, as many as one in fifteen may have difficulty interpreting the rainbow hues frequently used in maps, charts, and graphs. Figure 1 simulates how

one spectral color scheme, and a better alternative, may appear to color-blind readers.

Color-blind individuals see some colored data graphics quite differently from the general population. The human visual system normally perceives color through photosensitive cones in the eye that are tuned to receive wavelengths in the red, green, and blue portions of the visible spectrum. People who lack cones sensitive to one of the three wavelengths are called dichromats [Fortner and Meyer, 1997]. Individuals whose

receptors are shifted toward one or the other end of the spectrum are called anomalous trichromats. The term “color deficient” encompasses dichromats and anomalous trichromats, as well as those who exhibit rarer forms of impaired color vision. Because a sex-linked recessive gene is implicated in the condition, color-vision deficiency is far more common in men than in women.

Algorithms based on psychophysical observations make it possible to simulate the appearance of colored images to color-deficient viewers [Brettel *et al.*, 1997]. Data maps of the type shown here serve two main purposes: detection of large-scale patterns and determination of specific grid-cell or point values. The saturated spectral scale (Figure 1a) creates a region of confusion centered on the North American continent where achieving either purpose becomes nearly impossible for dichromat

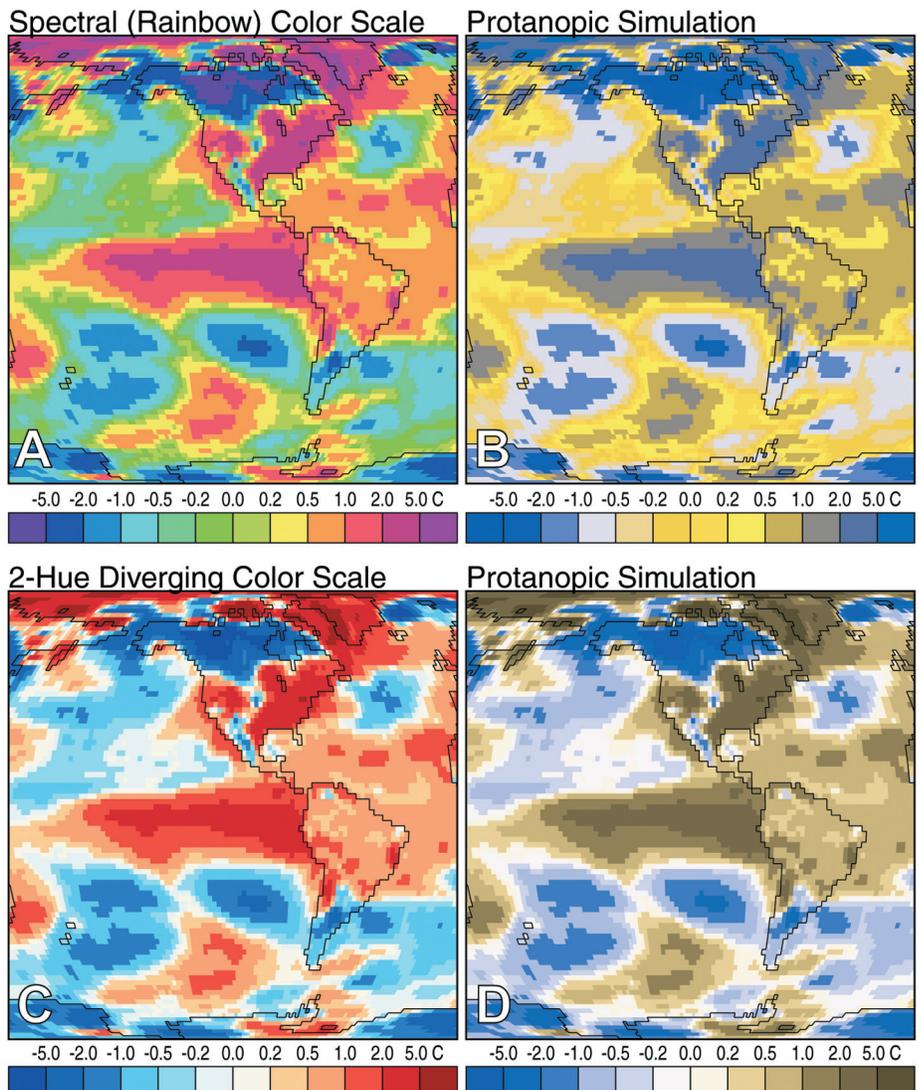


Fig. 1. Two-meter air temperature anomalies (*i.e.*, differences from the 1971–2000 mean) for January 1998 (during a recent El Niño) using two different color schemes. (A) Data using a saturated spectral scheme similar to those used by many geoscience authors; (B) A simulation of the spectral image as it might appear to individuals with protanopic vision, one of the most common types of color-vision deficiency in which the retina lacks red-sensitive cones; (C) The same data mapped using a red-white-blue diverging scale; and (D) The corresponding simulation for color-deficient readers. (NCEP/NCAR reanalysis data.)

readers. Large negative temperature anomalies keyed to violet and blue appear on the map adjacent to large positive temperature anomalies depicted in red and orange; but to the color-deficient viewer, the hues form a continuous progression of “blue” such that they can not distinguish large positive anomalies from large negative anomalies (Figure 1b). The two-hued red and blue image (Figure 1c) displays the same map region such that both color-deficient (Figure 1d) and normally sighted readers can detect patterns and look up values.

Improving Color Schemes While Accommodating Color Deficiency

Designing effective color schemes demands attention to the needs of readers who are unable to perceive certain colors. Color schemes that accommodate red or green-blind dichromats will accommodate most other forms of color deficiency [Rigden, 2002]. By designing with the severest forms of red and green color-blindness in mind, authors can create data graphics that work for all readers.

The following suggestions can help authors make rainbow-colored graphics accessible to more of their readers and can be used to improve both spectral and nonspectral color schemes.

1. Avoid the use of spectral schemes to represent sequential data because the spectral order of visible light carries no inherent magnitude message. Readers do not automatically perceive violet as greater than red even though the two colors occupy opposite ends of the color spectrum. Rainbow color schemes are therefore not appropriate if the data to be mapped or graphed represent a distribution of values ranging from low to high. With suitable modification, however [Brewer, 1997], spectral schemes can work for continuously distributed diverging data, such as anomalies and residuals (Figure 2c) and for the display of categorical data (Figure 2e).

2. Use yellow with care and avoid yellow-green colors altogether in spectral schemes. Readers with color-deficient vision often confuse yellow-green with orange colors. Yellow appears brightest among the primary colors and stands out visually for color-impaired and normally sighted readers alike. The yellow portion of the scheme should therefore be aligned with the midpoint of the data distribution if emphasizing the midpoint of a diverging distribution is a clear goal of the presentation. However, yellow may lead to misperceptions when the midpoint critical value of the data is not significant to the presentation or is not precisely known.

3. Use color intensity (or value) to reinforce hue as a visual indicator of magnitude. Hue is what we typically refer to as color; red, blue, green, and orange are all hues. Intensity or value may also be referred to as lightness, brightness, or luminosity. While it is possible to select hue sequences that are distinguishable by individuals with either color-deficient or normal vision, intensity readily provides perceptual ordering for all readers. Using intensity as well as hue also makes the quality of color reproduction less critical to the

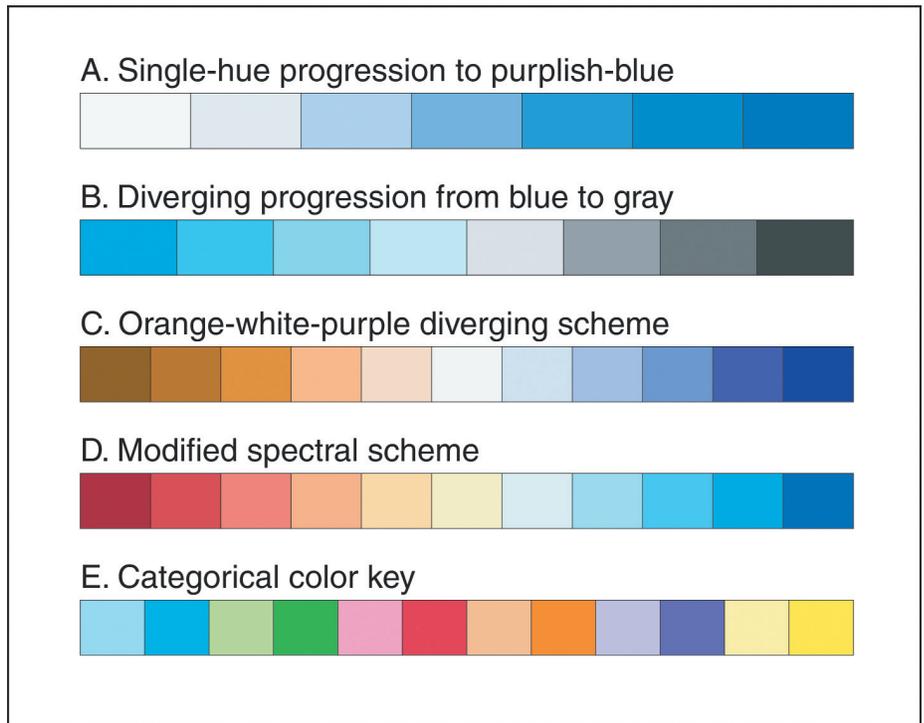


Fig. 2. Color schemes of one or two hues progressing from light to dark convey sequential data effectively. (A) A single hue progression suitable for sequential data. Diverging color schemes may be constructed by combining pairs of sequential schemes at the midpoint; (B) How color can extend a simple intensity scale, making it two-sided and therefore suitable for displaying either sequential or diverging data; (C) Combination of an orange-to-white sequence with a white-to-purple sequence, a scheme that appears very similar to both color-deficient and normally sighted readers. The rainbow spectrum appeals visually to many authors and readers, but an unmodified spectral color sequence proves ineffective for most purposes. The color scheme depicted in Figure 2D avoids yellow-green and varies intensity as well as hue, employing spectral color while avoiding the shortcomings of rainbow displays. For authors wishing to depict categorical rather than continuously distributed data, Figure 2E combines 12 bright colors that are mutually distinguishable from one another.

presentation and helps make even gray scale photocopies somewhat legible.

Design Patterns

The results of human-subject studies may inform designers, but design skill comes primarily from experience and example. Design patterns [Alexander, 1979; Gamma et al., 1995] for data graphics can communicate experience to nonexpert designers and may help integrate scientific knowledge into the design process. Design patterns make well-proven design knowledge explicit using a specific literal form: “Each pattern is a three-part rule, which expresses a relation between a certain context, a problem, and a solution” [Alexander, 1979, p. 247]. Table 1 documents two design patterns for the use of color in data graphics. Every design pattern carries a descriptive name (in bold italic in Table 1) and contains instructions to the designer for creating a specific instance of the pattern. Collectively, the names form a shared vocabulary for communicating about designs.

Example Color Schemes

Computer graphics tools make changing color palettes easy, but devising new color schemes for effective data graphics remains

a challenging design task. The color schemes in Figure 2 have been tried and tested for print and online display and evaluated for legibility by readers with color-deficient vision. They should work well with a range of Earth science data. Accurate reproduction of color for print and computer displays is a complex problem in its own right. Computer displays typically reproduce colors using an additive (RGB) color model, while print reproduction usually uses a subtractive (CMYK) color model [Fortner and Meyer, 1997]. Detailed specifications for reproducing these and other color schemes using both RGB and CMYK models are available on the Web site: <http://geography.uoregon.edu/datagraphics/>. Others are provided by Brewer et al. [2003].

Acknowledgments

The authors gratefully acknowledge the contributions of Aileen Buckley, Peter Killoran, Sarah Shafer, and Jacqueline Shinker to this work, and the improvements to the manuscript suggested by Cindy Brewer and Henry Shaw. National Centers for Environment Prediction (NCEP) reanalysis data were provided by the National Oceanic and Atmospheric Administration-Cooperative Institute for Research and Environmental Sciences' Climate Diagnostics Center,

Table 1. Design Patterns for the Use of Color in Data Graphics.

| |
|---|
| <i>Diverging color scheme</i> |
| <i>Display diverging data by using two complementary color schemes that diverge from a common hue.</i> |
| <p>Problem: Many data distributions include a midpoint critical value where both ends of the data distribution are of interest. The midpoint critical value "...may be a mean, median, or zero value..." and the reader "...is often interested in patterns in the data that show clusters both above and below the critical value." [Brewer, 1996, p. 79] Examples of diverging data distributions include anomalies (see Figure 1) and residuals.</p> <p>Solution: Align perceptual orderings with logical orderings by displaying diverging data using a diverging (double-ended) color scheme. Craft color schemes using intensity to indicate magnitude and hue to indicate sign. In cases where the magnitude is relevant but the sign is not, consider using a more parsimonious gray scale intensity scheme. The combination of various pairs of single-hue sequential schemes with common end points produces many useful diverging schemes.</p> <p>Context: Choosing an appropriate color scheme for diverging data need not be a difficult or complicated process. For some variables, convention suggests the color scheme. Air temperature, for example, is often displayed using a diverging scale of red and blue with zero on the Celsius scale marking the transition. Cartographers have traditionally displayed surface elevation using a scheme of browns or greens and ocean depth using progressively darker blues. Brewer [1996] suggests the following hue combinations for diverging schemes: red/blue, orange/blue, orange/purple, yellow/purple, brown/blue, and yellow/blue.</p> |
| <i>Sequential color scheme</i> |
| <i>Display data that do not contain midpoint critical values by using sequential color schemes</i> |
| <p>Problem: Many data distributions include a range of values without a significant midpoint. Absolute critical values may bound such distributions, as in the case of percentages, or the range of sampled data may arbitrarily define the end points, but there is no significant value within the range of the data.</p> <p>Solution: Display sequential data using a sequence of lightness steps combined with a single hue (e.g., Figure 2a) or with a hue transition (e.g., Figure 2c).</p> <p>Context: Spectral color schemes work poorly for sequential data because spectral order produces no natural magnitude message in the viewer's mind.</p> |
| <i>Detailed information, including additional patterns and examples, is available at http://geography.uoregon.edu/datagraphics/.</i> |

Boulder, Colorado (<http://www.cdc.noaa.gov/>). Research was supported by U.S. National Science Foundation grant ATM-9910638.

References

- Alexander, C. (1979), *The Timeless Way of Building*, Oxford Univ. Press, New York.
- Brettel, H., F. Viénot, and J. D. Mollon (1997), Computerized simulation of color appearance for dichromats, *J. Opt. Soc. Am.*, 14(10), 2647–2655.
- Brewer, C. A. (1996), Guidelines for selecting colors for diverging schemes on maps, *Cartogr. J.*, 33(2), 79–86.
- Brewer, C. A. (1997), Spectral schemes: controversial color use on maps, *Cartogr. Geogr. Inf. Syst.*, 24(4), 203–220.
- Brewer, C. A., G. W. Hatchard, and M. A. Harrower (2003), Colorbrewer in print: A catalog of color schemes for maps, *Cartogr. Geogr. Inf. Sci.*, 30, 5–32. (<http://www.ColorBrewer.org>).
- Cleveland, W. S. (1984), Graphs in scientific publications, *Am. Stat.*, 38(4), 261–269.
- Fortner, B., and T. H. Meyer (1997), *Number by Colors: A Guide to Using Color to Understand Technical Data*, 349 pp., Springer-Verlag, New York.
- Gamma, E., R. Helm, R. Johnson, and J. Vlissides (1995), *Design Patterns: Elements of Reusable Object-Oriented Software*, 395 pp., Addison-Wesley-Longman, Reading, Mass.
- Rigden, C. (2002), Now you see it, now you don't, *IEEE Comput.*, 35(7), 104–105.
- Tufte, E. R. (1983), *The Visual Display of Quantitative Information*, 197 pp., Graphics Press, Cheshire, Conn.
- Ware, C. (2004), *Information Visualization: Perception for Design*, 486 pp., Morgan Kaufmann, Burlington, Mass.

Author Information

Adam Light and Patrick J. Bartlein, University of Oregon, Eugene
For additional information, contact P. J. Bartlein;
E-mail: bartlein@uoregon.edu.