EOS

EOS, TRANSACTIONS, AMERICAN GEOPHYSICAL UNION

REPRINTS

The End of the Rainbow? Color Schemes for Improved Data Graphics. A. Light and P.J. Bartlein, *EOS*, Vol. 85, No. 40, 5 October 2004, pp. 385 & 391.

End of the Rainbow Indeed. J.I. Sammons, *EOS*, Vol. 85, No. 47, 23 November 2004.

Comment on "Color Schemes for Improved Data Graphics," by A. Light and P.J. Bartlein. D.B. Stephenson, *EOS*, Vol. 86, No. 20, 17 May 2005, p. 196.

Reply. A. Light and P.J. Bartlein, *EOS*, Vol. 86, No. 20, 17 May 2005, p. 196.

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

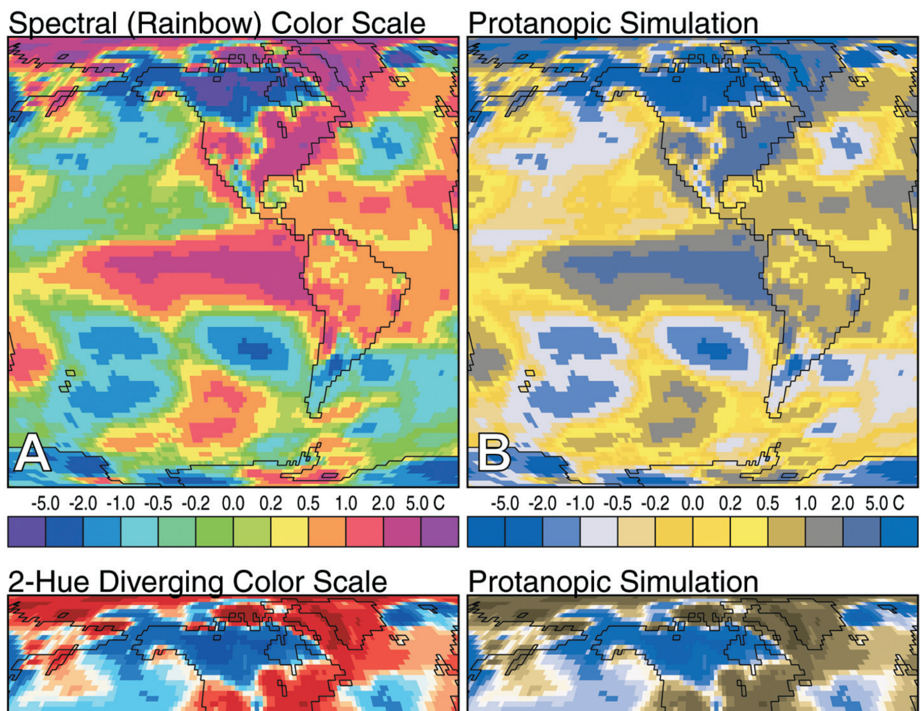


Fig. 1. Two-meter air temperature anomalies (i.e., differences from the 1971–2000 mean) for January 1998 (during a recent El Niño) using two different color schemes. (A) Data using a saturated spectral scheme similar to those used by many geoscience authors; (B) A simulation of the spectral image as 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.)

By A. Light and P.J. Bartlein

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 Accomodating 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 presentation and helps make even gray scale photocopies somewhat legible.

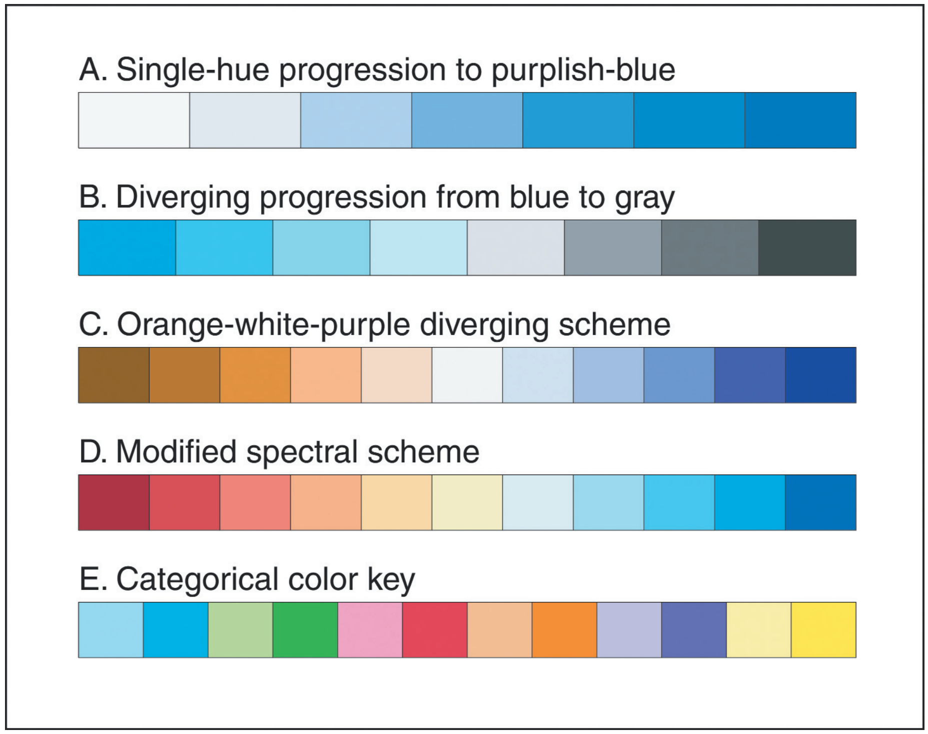


Fig. 2. Color schemes of one or two hues progressing from light to dark convey sequential data effectively. (A) A single hue progression suitable for sequential data. Diverging color schemes may be constructed by combining pairs of sequential schemes at the midpoint; (B) How color can extend a simple 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.

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.

Diverging color scheme

Display diverging data by using two complementary color schemes that diverge from a common hue.

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.

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.

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.

Sequential color scheme

Display data that do not contain midpoint critical values by using sequential color schemes

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.

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).

Context: Spectral color schemes work poorly for sequential data because spectral order produces no natural magnitude message in the viewer's mind.

Detailed information, including additional patterns and examples, is available at http://geography.uoregon.edu/datagraphics/.

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.

LETTERS

End of the Rainbow Indeed

PAGE 498

Many thanks to Adam Light and Patrick J. Bartlein for their helpful and succinct article, The End of the Rainbow? Color Schemes for Improved Data Graphics, regarding color in data illustration (*Eos*, *85*(40), 5 October 2004). I especially appreciate their point about the use of spectral schemes to illustrate sequential data.

As a science writer and reader, I have often struggled with the Crayola-inspired illustrations that are prevalent today. Overuse of color forces the reader into a myopic "paint by numbers" exercise to extract meaning. The situation reminds me of notes that I received just after Microsoft Windows developed the capability to print bit-mapped fonts. For a while, some authors seemed carried away by this new capability and marked each line, each change of voice, with a new font.

Here is one example of the need to use color correctly. In the early 1960s, I was a navigator aboard a U.S. Navy destroyer. During the time of my service, the Naval Hydrographic Office discontinued the use of red color on all U.S. charts and adopted a magenta and grey replacement. That made my life a whole lot easier, because then I could read the formerly red, now magenta, "lighted mark" symbols under military red battle lighting. I hope that this article is widely considered and that its clear message is taken to heart.

—JAMES I. SAMMONS, Sammons INK, Ltd., North Kingstown, R.I.; E-mail: jimsa@cox.net

FORUM

Comment on “Color Schemes for Improved Data Graphics,” by A. Light and P. J. Bartlein

PAGE 196

As a color-blind climatologist, I very much appreciated the important color-awareness issues raised by A. Light and P. J. Bartlein in their recent *Eos* article titled “The End of the Rainbow? Color Schemes for Improved Data Graphics,” (*85*(40), 5 October 2004, p. 385).

Meteorological and climate information is frequently communicated in the form of highly colored spatial maps (e.g., daily weather forecasts). The use of color is so endemic in climate research that it would be unthinkable to present scientific work at a climate conference without using color! The color schemes generally chosen to create such maps follow no universal convention and often consist of colors that cause confusion to both color-impaired and non-color-impaired people.

Light and Bartlein raised several pertinent points and made some good recommendations. However, a number of points require some more clarification.

First, Light and Bartlein's example in Figure 1 showed simulations of how images might appear to people with protanopic vision, which they claimed to be “one of the most common types of color-vision deficiency in which the retina lacks red-sensitive cones.” However, of the 8% of Caucasian males and 0.4% of Caucasian females with color-impaired vision, only 2% of all males and 0.01% of all females have this form of color-impairment.

By far the more common forms of color-impaired vision are deuteranopia/deuteranomaly (green-weakness), with 6% and 0.25% overall prevalence in Caucasian males and females, respectively. People with these conditions have cones that are less sensitive to medium wavelengths (greens) and so have reduced ability to discriminate the green (rather than the red) component in colors.

Deuteranopia and deuteranomaly are also known as Daltonism after the well-known scientist, John Dalton, in recognition of his groundbreaking scientific study on color vision [*Dalton,* 1798]. Dalton's diagnosis was confirmed as deuteranopia in 1995, some 150 years after his death, by DNA analysis of his preserved eyeball. Interestingly, Dalton was also a keen meteorological observer and so perhaps would be very sympathetic to the color issues now being raised here in *Eos*.

Since deuteranopia is much more prevalent than protanopia, it makes sense to focus attention on color schemes that work for deuteranopia, and it would have been more relevant if Light and Bartlein had presented a deuteranopic simulation (green-deficient) rather than a protanopic (red-deficient) simulation in their Figure 1.

This perhaps explains why as someone with deuteranomalous vision, I find the color schemes in Light and Bartlein's Figure 2 rather difficult to identify. For example, I would have referred to their purple colors as blue, and this can easily cause much embarrassment in, say, a scientific talk. Their categorical color key in Figure 2e is particularly difficult—a color-blind colleague and I could each only successfully name nine out of the 12 colors.

Light and Bartlein presented three suggestions to improve color schemes: Avoid spectral schemes, use yellow with care and avoid yellow-green colors, and use color intensity to reinforce hue. Suggestions 1 and 3 make good sense to me, but I disagree with suggestion 2—yellow is one of the colors I can distinguish the most easily, and so I favor this as a middle color in color schemes (and in my PowerPoint talks).

A few other recommendations to help improve color plots are:

1. Avoid using red and green together in a color scheme (remember the old adage: “Red and green should never be seen …”!). It is best to avoid using green completely if possible. Distinguishing between green and brown is especially difficult, and so these two colors should be avoided.
2. Use a limited number of high-contrast primary colors (e.g., white, black, blue, yellow, red) rather than a large number of pastel shades. Bright colors such as those commonly used for water sport equipment are good, whereas pastel shades such as Monet's water lily colors are a nightmare!
3. Check your figures (and Web pages) using the excellent and simple-to-use image checking tools that are now available on the Internet (e.g., http://www.vischeck.com).

More information on color vision and how to design for it is given by *Viénot et al.* [1995] and on Christine Rigden's informative Web pages (http://more.btexact.com/people/rigdence/colors).

References

Dalton, J. (1798), Extraordinary facts relating to the vision of colours: With observations," *Mem. Proc. Manchester Lit. Philos. Soc., 5,* 28–45.

Light, A., and P. J. Bartlein (2004), The end of the rainbow? Color schemes for improved data graphics, *Eos, 85*(40), 385, 391.

Viénot, F., H. Brettel, L. Ott, A. Ben M'Barek, and J. D. Mollon (1995), What do colour-blind people see?, *Nature, 376,* 127–128.

—David B. Stephenson, University of Reading, U.K.

Reply

PAGE 196

By calling attention to the perception of data graphics among color-impaired readers, we hoped to raise awareness of an even more prevalent phenomenon: the misuse of spectral, or "rainbow," color schemes. David B. Stephenson is quite right regarding the incidence of different forms of color-vision impairment, which varies among racial and ethnic groups and is much lower for females [*Sharpe et al.*, 1999]. Our article highlighted simulations of the less common protanopia to dramatize the general peril of spectral schemes.

Figure 1 shows simulations of spectral, diverging, and progressive color schemes as they might appear to readers with either of the most common forms of color deficiency. Figures 1a and 1b illustrate why the adage "red and green should never be seen" is a good one. The simulations in the right-hand column show how deuteranopes may find it difficult or impossible to distinguish between red and green.

Spectral schemes guarantee zones of confusion for those with color-impaired vision, but also limit comprehension by nearly everybody else. The situation arises, as in our example, from the similarity of end-member colors (dark reddish purple versus dark purplish red), and the off-center location of the most vivid colors (yellow and yellow-green), which limit a reader's ability to comprehend gradients, overall patterns, or values for individual locations. In contrast, both color-deficient and normally sighted readers should be able to distinguish the color schemes in Figures 1c (blue-red) and 1d (green-magenta).

Creating effective data graphics for communication and analysis is a complex task; designing for color-deficient readers is only one of many challenges. Adages like "red and green should never be seen" and "the end of the rainbow" may help authors recognize obvious problems, but cannot provide sufficient guidance for them to reliably produce effective data graphics from first principles.

Fortunately, designing data graphics from first principles seldom proves necessary. Scientific authors should approach data graphics design from a position of strength—as masters of their own data—rather than from an uncertain mastery of color. Specialists have tested color schemes and identified those that work well; most authors will do well to select from the proven alternatives (e.g., http://www.ColorBrewer.org). Editors and publishers could help authors by recommending menus of color schemes for use with specific types of data.

Further information may be found at http://geography.uoregon.edu/datagraphics/.

Reference

Sharpe, L. T., et al. (1999), Opsin genes, cone photopigments, color vision, and color blindness, in *Color Vision: From Genes to Perception*, edited by K. R. Gegenfurtner and L. T. Sharpe, pp. 3–51, Cambridge Univ. Press, New York.

—ADAM LIGHT and PATRICK J. BARTLEIN, University of Oregon, Eugene

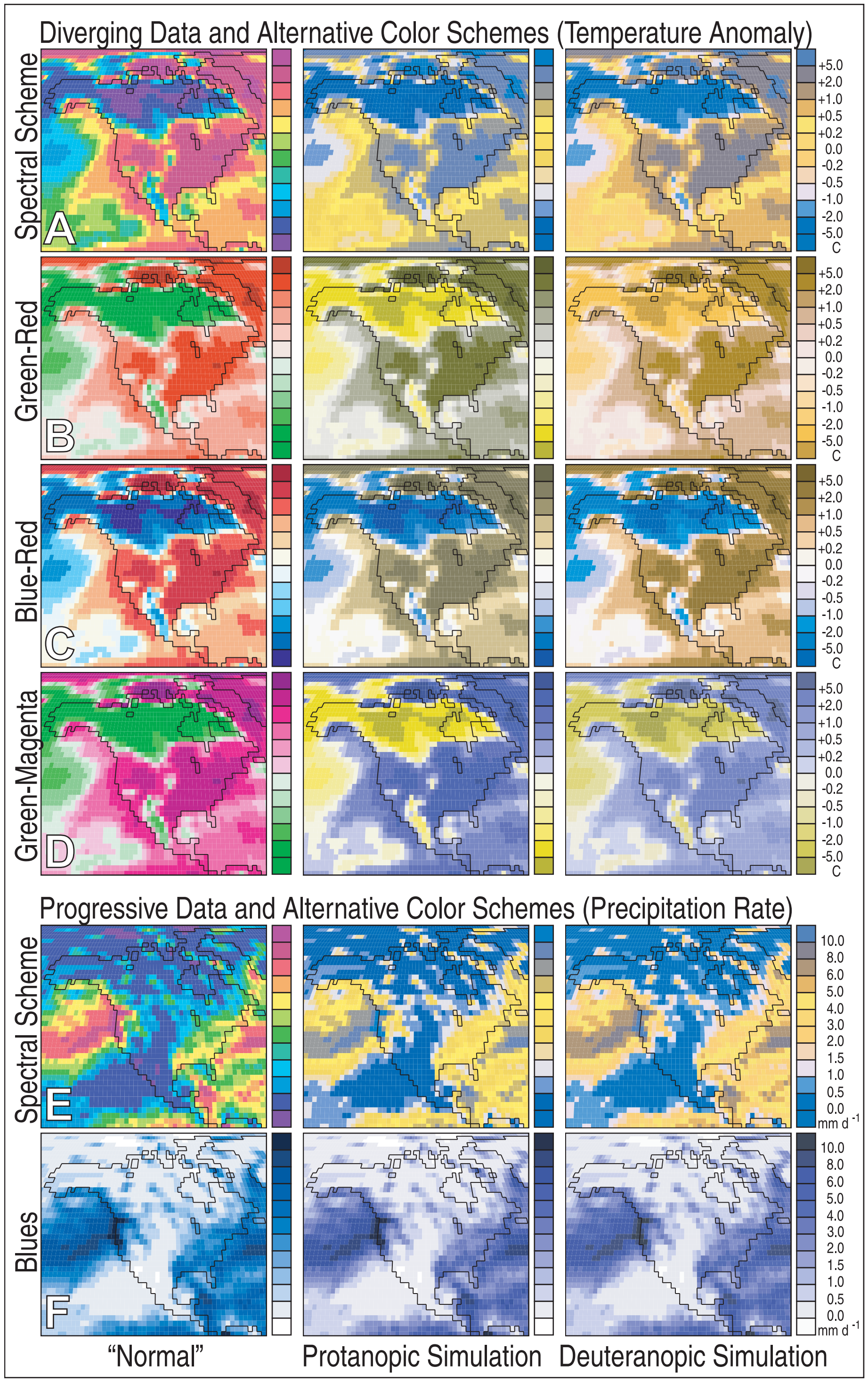
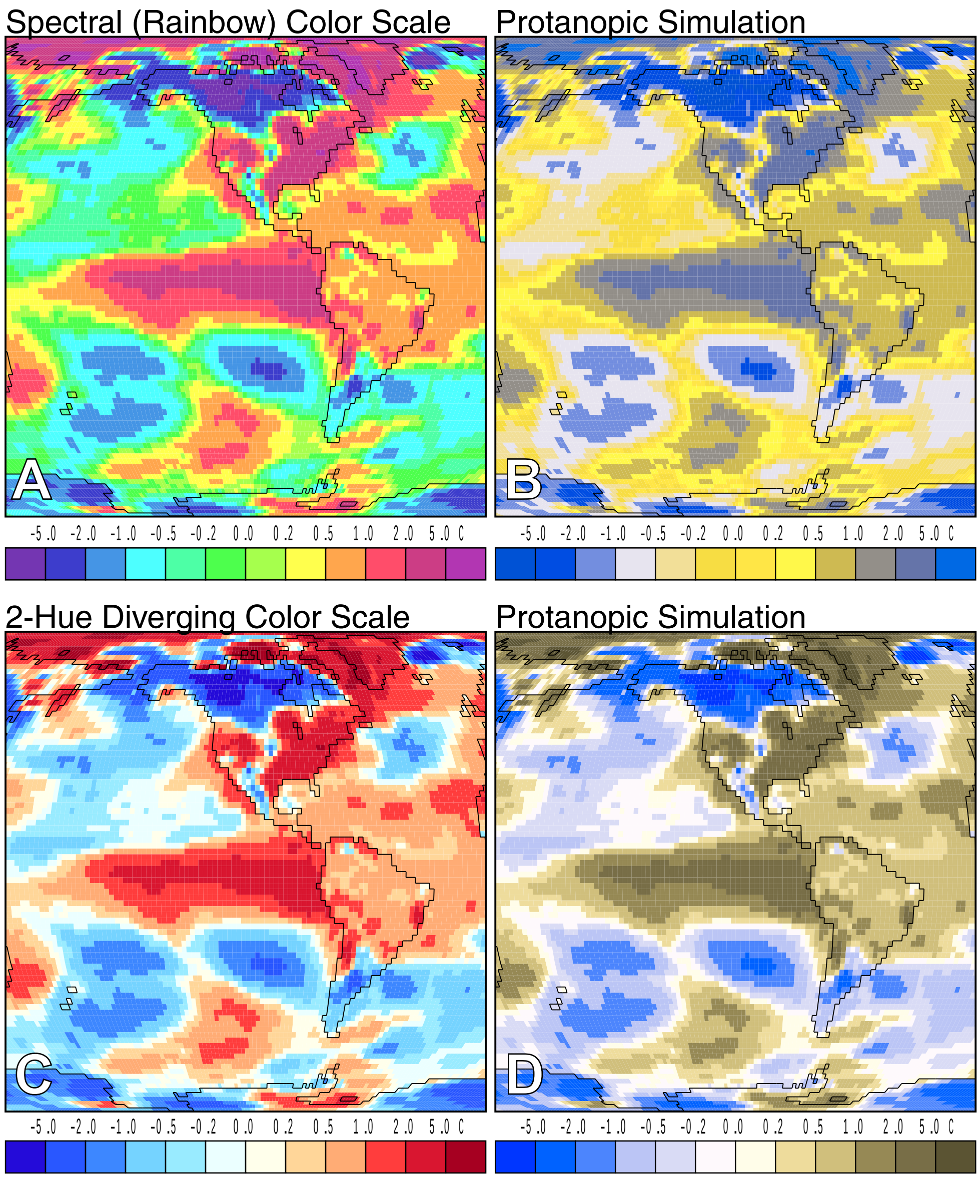


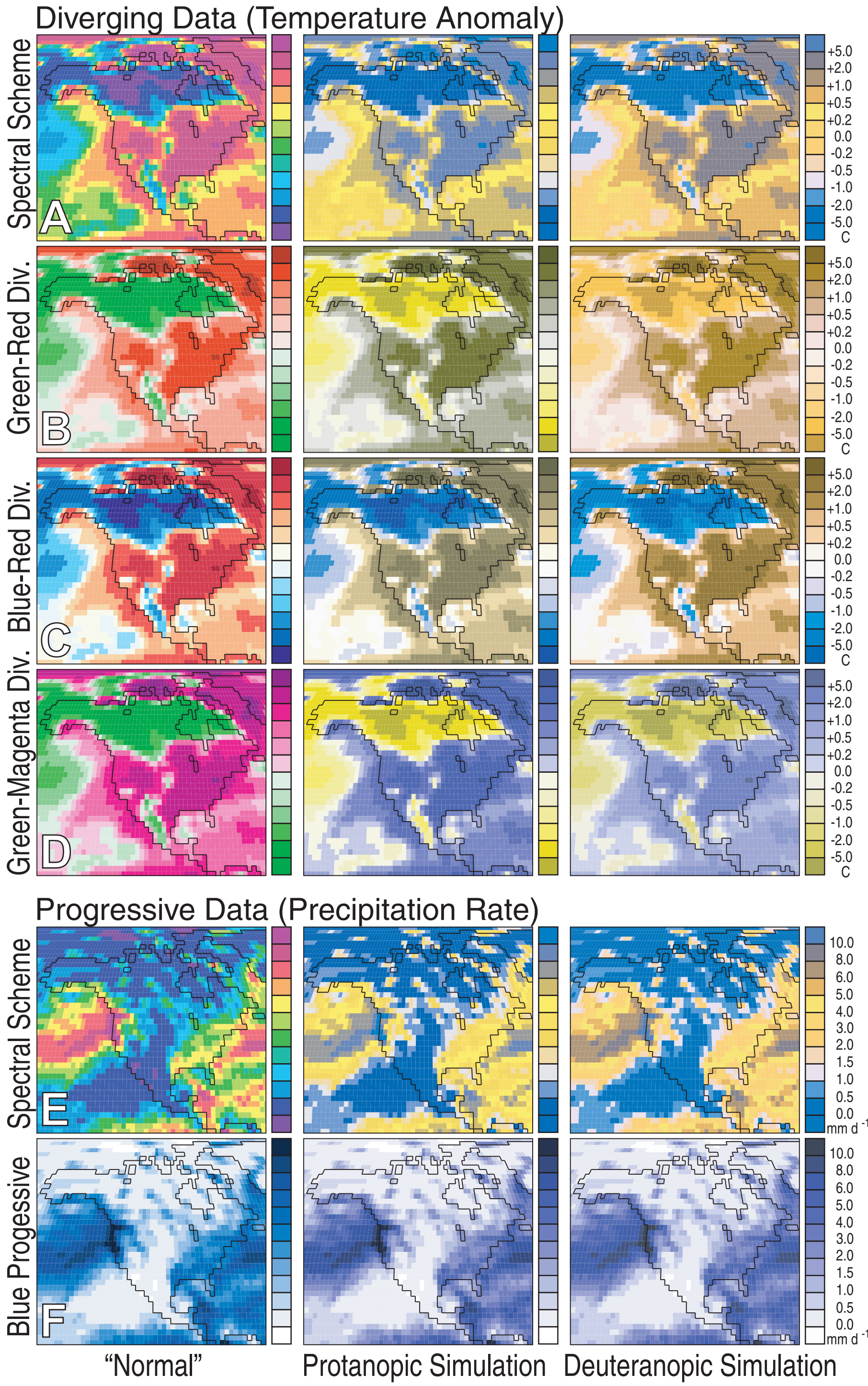
Fig. 1. The left column in each row displays a specific color scheme, while the middle and right columns display simulations of the scheme as it might appear to viewers who exhibit the two most common forms of color-vision deficiency. For "diverging" data, such as the temperature anomalies shown here, the locations of extrema and of the zero isopleth (the points where the temperature anomaly is zero) are significant. Figures 1a and 1b show how the spectral scheme may be slightly less confusing for deuteranopic viewers than protanopic ones, but also illustrate that the most prominent color (yellow) appears off-center in the range of scale values. Figures 1c and 1d show how properly chosen diverging color schemes can increase interpretability for all readers by making the location of extrema and zeros obvious; Figure 1d additionally demonstrates that should there be a conceptual reason for employing green, pairing it with magenta improves interpretability. For "progressive" data like precipitation rate, the spectral scheme (Figure 1e) arbitrarily emphasizes the middle range of values for all viewers, and makes it difficult to infer spatial gradients. Progressive data require only a single hue because intensity (or value) encodes the level of the data (Figure 1f). Blue connotes water in many cultures and is a more intuitive choice than, say, brown or red for representing precipitation, but a simple gray scale would also suffice. (Figures 1c, 1d, and 1f illustrate "good" color schemes.)

Color Scales and Color-Deficient-Viewer Simulations



Data: Jan 1998 2-m Air Temperature Anomalies (NCEP/NCAR Reanalysis Data)

Light, A and P.J. Bartlein (2004) The end of the rainbow? color schemes for improved data graphics. *EOS* Transactions of the American Geophysical Union 85(40):385,391.



Light, A and P.J. Bartlein (2005) Reply. *EOS* Transactions of the American Geophysical Union 86(20):196.

Light, A and P.J. Bartlein (2004) The end of the rainbow? color schemes for improved data graphics. *EOS* Transactions of the American Geophysical Union 85(40):385,391.