POINTS OF VIEW

Color coding

Color can add dimensionality and richness to scientific communications. In figures, color is typically used to differentiate information into classes. The challenge is picking colors that are discriminable. A systematic approach to choosing colors can help us find a lineup effective for color coding.

Occasionally, authors use a sequence of colors, such as the 'rainbow' color scheme, to represent a range of values. Color, however, is not ideal for encoding quantitative data because of the inherent ambiguity in how the different colors should be ordered. For instance, does yellow represent a smaller value than blue? One could pattern the sequence after the ordering of visible light by wavelength (remembered by the mnemonic ROYGBIV), but use of this color spectrum is inherently problematic. The transitions from red to yellow to green and so on are uneven, breaking the correspondence between color and numerical value. Visually, certain colors in the rainbow spectrum seem to run on, whereas others are short lived. Even when we limit the spectrum to just a few colors, the incremental change in mapped value still might not translate to the magnitude of change we see.

In contrast, color is well suited to represent categorical data when it is used properly—for example, to distinguish between experimental conditions. If used improperly, such as by assigning intense or weak colors to specific categories, color can bias the reader. Because color is such a potent differentiator, the appropriate strategy is to choose colors that are discernible from one another but comparable in visibility.

Color is a relative medium, and neighboring colors can affect visual perception. For example, it is possible to make the same color look different or different colors appear the same (or nearly the same) by changing only the background color (**Fig. 1a,b**). The perception of color depends on context, and manipulating the attributes of neighboring colors affects how we see the original color[1]. A heat map requires us to judge the relative brightness of colors in a matrix. The interaction of color can cause a profound effect that makes this graphical representation suffer (**Fig. 1c**).

Every color is described by three properties: hue, saturation and lightness. Hue is the attribute we use to classify a color as red or yellow. Saturation describes the neutrality of a color; a red object with little or no white is said to be very saturated. The lightness of a color tells us about its relative ordering on the dark-to-light scale.
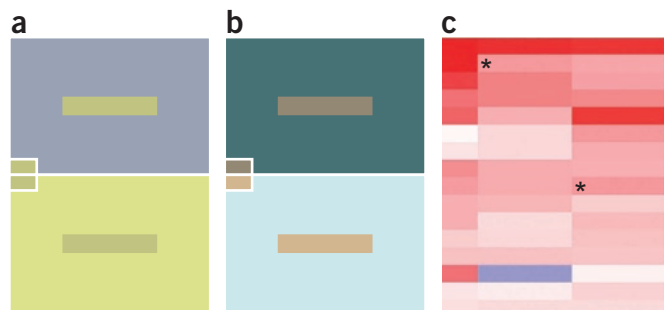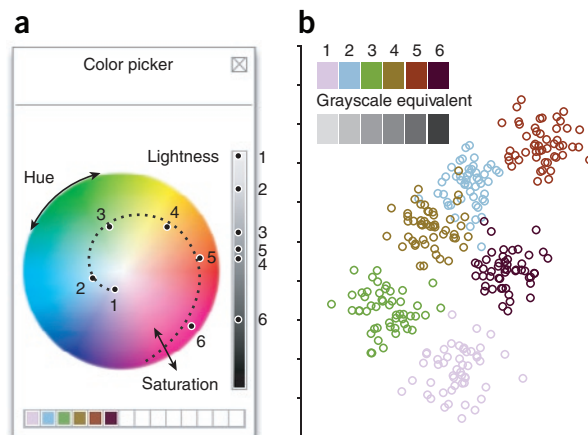
Figure 2 | Color has hue, saturation and brightness. (**a,b**) Colors can be tuned using a color picker (**a**). Spiraling through hue and saturation while varying lightness can generate a discernible color set distinguishable even in grayscale (points labeled 1–6).

On a computer, we can tune color attributes using the color picker (**Fig. 2a**). On a Mac or PC and in software such as Adobe Illustrator and Photoshop, the color picker is based on the traditional color wheel. In this system, hues are arranged around a circle with saturation increasing from the center outward. The 'true' color (hue) is near the ring midway from the center. On a PC and in Adobe products, the color wheel is transformed into a square with hue arrayed across the top and saturation decreasing from top to bottom. In all cases, lightness is controlled by a separate slider.

To pick colors easily discernible from each other, whether in color or converted to grayscale, spiral through the color wheel while varying the lightness (**Fig. 2**). We can achieve wide dynamic range by adjusting all three attributes of color. Our perceptual system is highly sensitive to grayscale, and the lightness property makes it possible to differentiate colors when photocopied to black and white. In this way, we can define a group of 6–8 colors. Beyond this number, the task of picking distinctive colors becomes difficult. To show more categories, we can rely on textural differences in addition to color. For example, we can encode data for two categories as red crosses and red circles.

Just picking suitable colors is not always sufficient, though. The size of the 'visual objects' in the figure also matters; the smaller the objects (or the thinner the lines) the greater the variations in hue, saturation and lightness that are needed. Finally, to test for comparable visibility of the selected colors, squint at the graphic and look for general evenness.

Color is a familiar and widely used design element. Poor color choices can introduce bias and unwanted artifacts into our presentations. Careful consideration when choosing colors will help us make the most of the communication and enable readers to discern the encoded information. Next month, we will focus on the design of data graphs.

Bang Wong

1. Albers, J. *Interaction of Color* (Yale University Press, New Haven, Connecticut, USA, 1975).

Bang Wong is the creative director of the Broad Institute of the Massachusetts Institute of Technology and Harvard and an adjunct assistant professor in the Department of Art as Applied to Medicine at The Johns Hopkins University School of Medicine.

Figure 1 | Perception of color can vary. (**a,b**) The same color can look different (**a**), and different colors can appear to be nearly the same by changing the background color (**b**)[1]. (**c**) The rectangles in the heat map indicated by the asterisks (*) are the same color but appear to be different.