POINTS OF VIEW

Design of data figures

Data figures or graphs are essential to life-science communication. Using these tools authors encode information that readers later decode. It is imperative that graphs are interpreted correctly. Despite the importance and widespread use of graphs, we primarily rely on our intuition, common sense and precedent in published material when creating them—a largely unscientific approach.

Because accurately interpreting visual variables is such a vital step in understanding graphs, a rational framework for creating effective graphs would accommodate the needs of the reader and focus on the strengths of human perception. Conversely, we want to avoid displays of data that are misleading or difficult to discern. For example, it can be tough to accurately judge the differences between two curves (**Fig. 1a**). The disparity is actually constant but our perceptual system is attuned to detecting minimal distances so the divergence appears to decrease. Another shortcoming limits our ability to accurately judge relative area. This diminishes the usefulness of bubble charts. For example, the larger circle in **Figure 1a** is 14 times larger than the smaller circle.

In 1967, the French cartographer Jacques Bertin provided a wide theoretical framework for information visualization[1]. His analysis focused on the visual properties of graphical elements such as shape, orientation, color, texture, volume and size for displaying quantitative variation. He defined several visual operations needed to extract information stored in graphs. Cleveland and McGill were one of the first to measure people's ability to efficiently and accurately carry out these elementary perceptual tasks[2] (**Table 1**).

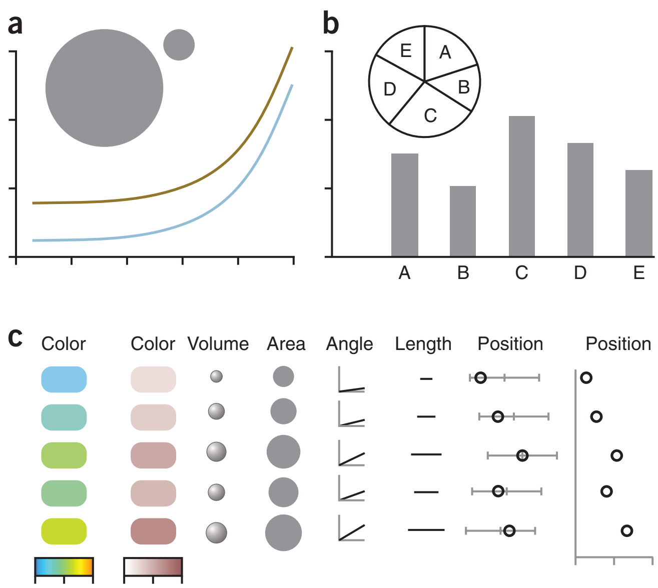

Figure 1 | Some visual estimations are more easily carried out than others. (**a**) Examples illustrating the difficulty in interpreting graphs and charts accurately. (**b**) Same data presented in a bar chart and in a pie chart. (**c**) Different visual variables encoding the same five values.

Table 1 | Elementary perceptual tasks

Rank	Aspect to compare
1	Positions on a common scale
2	Positions on the same but nonaligned scales
3	Lengths
4	Angles, slopes
5	Area
6	Volume, color saturation
7	Color hue

Tasks are ordered from most to least accurate. Information adapted from ref. 2.

When communicating with graphs, we want readers to perceive patterns and trends. This is distinct from conveying information through tables in which we report precise names and numbers. Cleveland and McGill's study assessed people's ability to judge the relative magnitude between two values encoded with a particular visual variable (for example, length, angle and others). In other words, they asked people to estimate how many times bigger *A* is when compared to *B*. Accuracy in their study does not imply reading out precise values from data points in graphs.

Different graph types depend on different visual assessments to uncover underlying trends. Pie charts are a common way to show parts of a whole. Most readers will likely judge angle when extracting information from pie charts, but they could also compare areas and arc length of the slices (**Fig. 1b**). Each of these perceptual tasks ranks low in efficiency and accuracy (**Table 1**). Plotting the same data as a bar chart effectively shows relative values (**Fig. 1b**).

When we occasionally need to invent new ways to graph data, we ideally want to use perceptual tasks that rank high in efficiency and accuracy (**Table 1**). In **Figure 1c**, I plotted the same five values using different encoding. In some cases, identifying magnitude and direction of change is laborious. In other cases, the trends are readily apparent. Encodings on the right more efficiently and accurately display the magnitude and direction of change. Though we can detect slight shifts in color hue, the relationship between hue and quantitative value is not obvious (see also ref. 3), making color hue one of the weaker methods to illustrate relative values.

Communicating with graphs depends on authors encoding information for readers to decode. Graphs' effectiveness can benefit from attention to their visual design. Composing figures with strong visual cues and relying on accurate perceptual tasks supports the visual assessment critical for interpreting information from graphs. Next month we will explore salience, the use of visual properties as differentiators.

Bang Wong

1. Bertin, J. *Semiology of Graphics*, English translation by W.J. Berg (University of Wisconsin Press, Madison, Wisconsin, USA, 1983).
2. Cleveland, W.S. & McGill, R. *Science* **229**, 828–833 (1985).
3. Wong, B. *Nat. Methods* **7**, 573 (2010).

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.