EDITORIAL

Three Simple Ways to Identify Data Sets in a Figure

ACCESS | Metrics & More | Article Recommendations

Scientists use graphics, rather than text, to visually and more succinctly present the data analysis of their experimental or computational studies. With the widespread availability of computer software and design tools, we see many colorful and visually appealing graphs presented as *x-y*, *x-y-y*, 2D, 3D plots, etc. in scientific articles. When graphs contain multiple data sets, a common practice is to employ different colored plots in order to distinguish the different sets. One common difficulty a reader may experience while going through these graphs occurs when the individual traces are identified solely on the basis of colors (e.g., blue trace, red trace) (Figure 1). Whereas these colors

Figure 1. Distinguishing individual data sets solely based on color identification can be challenging. (Source: Strummer/Shutterstock.com)

were obvious for the person who drew the figure, it may not be so obvious for readers with a visual imparity or those who print the articles on a black-and-white printer. In addition, some screen settings can result in online colors displaying differently.

The obvious question is, *How can you present multiple data sets that readers can easily identify?* In the days before personal computers, all figures were drawn by hand and were in black and white. With color being unavailable, the scientific community developed a way to distinguish different traces through the use of symbols, solid/dashed/dotted lines, and legends and labels. Interestingly, these well-established

protocols are still useful in the modern age and can greatly enhance the quality of graphical presentation in scientific articles.

Previous Editorials[1−5] and *The ACS Guide to Scholarly Communication*[6] address some of the key points in creating figures, figure captions, axis scales and units, and other related tips (see Related Readings). In this Editorial, we provide three effective ways to define individual data sets in a figure.

Best Practices. Identify the data sets by including either a legend, a key, or labels, whichever is most appropriate. Here are three simple ways to create effective identifiers.

Legends. A legend is a caption or brief description. Instead of relying on just different colors, identify each trace with either letters (a, b, c, etc.) or numbers (i, ii, iii, etc.). A description of the traces is included either as an inset within the figure (provided the text is large enough to be legible, Figure 2) or in the figure caption. The inclusion of a legend is just one additional step the researcher has to introduce after drawing the graph, and it helps to explain individual traces in the caption with no ambiguity.
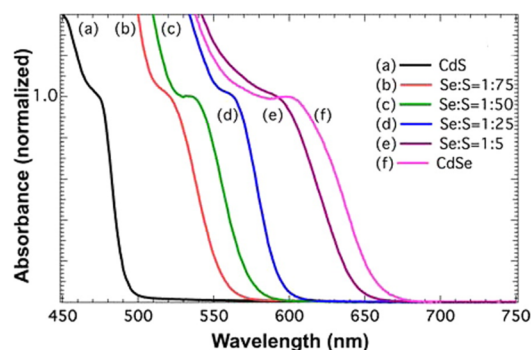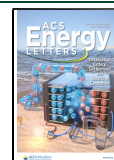

Figure 2. Use of legends to identify multiple traces and distinguish them with clarity. (Reprinted from Santra, P. K.; Kamat, P. V. *J. Am. Chem. Soc.* 2013, *135*, 877−885. DOI: 10.1021/ja310737m. Copyright 2012 American Chemical Society.)

Keys. When symbols are used, a key, which identifies data point symbols, is essential. Multiple sets of data points can be distinguished through the use of differently shaped data point symbols (e.g., ○, ●, □, ■, △, ▲). The key may be presented in either a figure inset or the figure caption (Figure 3). The use of different colors further enhances the presentation. Use of colored lines with symbols (-△-) in lieu of just a line (—) to represent keys is strongly encouraged.
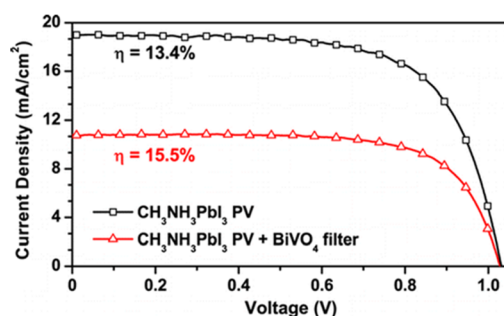

Figure 3. Use of a key to identify the data point symbols and distinguish the traces with clarity. (Reprinted from Chen, Y.-S.; Manser, J. S.; Kamat, P. V. *J. Am. Chem. Soc.* **2015**, *137*, 974−981. DOI: 10.1021/ja511739y. Copyright 2014 American Chemical Society.)

Labels. Attaching a short label next to a trace or data set (see, for example, Figure 4) is convenient way to identify or highlight the differences. Labels are useful when there are only a few data sets (two or three) in the figure. They are also useful to define individual columns in a bar diagram. It is important to avoid too much text in the figure as it leads to clutter. Note: Since colored text (lines or symbols) cannot be used in figure captions or in the main text, it is necessary to embed the color definition of the keys within the figure itself or describe them separately in the caption. For example, "(red triangles)" or "(red -△-)" can be used to define the keys.

Keep in mind different readership while composing a figure. It is not the quantity of data sets that you want to cram in a figure but the clarity of scientific presentation that matters while composin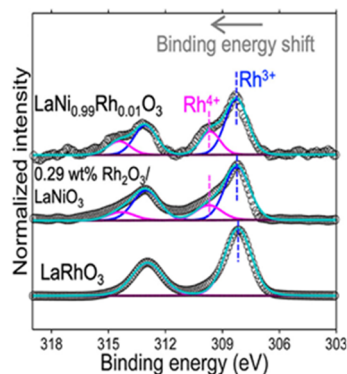g a figure. In addition, remember to explain all the data sets clearly in the caption along with experimental/analytical protocols employed.

Constance M. Biegel, Coordinating Editor, *ACS Energy Letters*

Prashant V. Kamat, Editor-in-Chief, *ACS Energy Letters* ⦿ orcid.org/0000-0002-2465-6819

■ AUTHOR INFORMATION

Notes

■ RELATED READINGS

(1) Kamat, P.; Hartland, G. V.; Schatz, G. C. Graphical Excellence. *J. Phys. Chem. Lett.* **2014**, *5*, 2118−2120.

(2) Buriak, J. M. Which Font Looks Best in a Figure? *Chem. Mater.* **2016**, *28*, 689−690.

(3) Buriak, J. M. Figure Size: Please Be Kind to Your Reader. *Chem. Mater.* **2017**, *29*, 8021−8022.

(4) Biegel, C. M.; Kamat, P. V. Ten Tips for Capturing Figures with Captions. *ACS Energy Lett.* **2019**, *4*, 637−638.

(5) Kamat, P. V. Absolute, Arbitrary, Relative, or Normalized Scale? How to Get the Scale Right. *ACS Energy Lett.* **2019**, *4*, 2005−2006.

(6) *ACS Guide to Scholarly Communication*; American Chemical Society, 2020; Part 4.1, Graphics & Multimedia. DOI: DOI: 10.1021/acsguide.40101.

Figure 4. Use of short labels to identify data sets. (Reprinted from Samira, S.; Camayang, J. C. A.; Patel, K.; Gu, X.-K.; Nikolla, E. *ACS Energy Lett.* **2021**, *6*, 1065−1072. DOI: 10.1021/acsenergylett.1c00102. Copyright 2021 American Chemical Society).