Patterns

CellPress
OPEN ACCESS

Perspective

Principles of Effective Data Visualization

Stephen R. Midway[1,*]
[1]Department of Oceanography and Coastal Sciences, Louisiana State University, Baton Rouge, LA 70803, USA
*Correspondence: smidway@lsu.edu
https://doi.org/10.1016/j.patter.2020.100141

THE BIGGER PICTURE Visuals are an increasingly important form of science communication, yet many scientists are not well trained in design principles for effective messaging. Despite challenges, many visuals can be improved by taking some simple steps before, during, and after their creation. This article presents some sequential principles that are designed to improve visual messages created by scientists.

1 2 3 4 5 Mainstream: Data science output is well understood and (nearly) universally adopted

SUMMARY

We live in a contemporary society surrounded by visuals, which, along with software options and electronic distribution, has created an increased importance on effective scientific visuals. Unfortunately, across scientific disciplines, many figures incorrectly present information or, when not incorrect, still use suboptimal data visualization practices. Presented here are ten principles that serve as guidance for authors who seek to improve their visual message. Some principles are less technical, such as determining the message before starting the visual, while other principles are more technical, such as how different color combinations imply different information. Because figure making is often not formally taught and figure standards are not readily enforced in science, it is incumbent upon scientists to be aware of best practices in order to most effectively tell the story of their data.

INTRODUCTION

Visual learning is one of the primary forms of interpreting information, which has historically combined images such as charts and graphs (see Box 1) with reading text.[1] However, developments on learning styles have suggested splitting up the visual learning modality in order to recognize the distinction between text and images.[2] Technology has also enhanced visual presentation, in terms of the ability to quickly create complex visual information while also cheaply distributing it via digital means (compared with paper, ink, and physical distribution). Visual information has also increased in scientific literature. In addition to the fact that figures are commonplace in scientific publications, many journals now require graphical abstracts[3] or might tweet figures to advertise an article. Dating back to the 1970s when computer-generated graphics began,[4] papers represented by an image on the journal cover have been cited more frequently than papers without a cover image.[5]

There are numerous advantages to quickly and effectively conveying scientific information; however, scientists often lack the design principles or technical skills to generate effective visuals. Going back several decades, Cleveland[6] found that 30% of graphs in the journal *Science* had at least one type of error. Several other studies have documented widespread errors or inefficiencies in scientific figures.[7–9] In fact, the increasing menu of visualization options can sometimes lead to poor fits between information and its presentation. These poor fits can even have the unintended consequence of confusing the readers and setting them back in their understanding of the material. While objective errors in graphs are hopefully in the minority of scientific works, what might be more common is suboptimal figure design, which takes place when a design element may not be objectively wrong but is ineffective to the point of limiting information transfer.

Effective figures suggest an understanding and interpretation of data; ineffective figures suggest the opposite. Although the field of data visualization has grown in recent years, the process of displaying information cannot—and perhaps should not—be fully mechanized. Much like statistical analyses often require expert opinions on top of best practices, figures also require choice despite well-documented recommendations. In other words, there may not be a singular best version of a given figure. Rather, there may be multiple effective versions of displaying a single piece of information, and it is the figure maker's job to weigh the advantages and disadvantages of each. Fortunately, there are numerous principles from which decisions can be made, and ultimately design is choice.[7]

The data visualization literature includes many great resources. While several resources are targeted at developing design proficiency, such as the series of columns run by *Nature*

Box 1.

Regarding terminology, the terms *graph*, *plot*, *chart*, *image*, *figure*, and *data visual(ization)* are often used interchangeably, although they may have different meanings in different instances. *Graph*, *plot*, and *chart* often refer to the display of data, data summaries, and models, while *image* suggests a picture. *Figure* is a general term but is commonly used to refer to visual elements, such as plots, in a scientific work. A *visual*, or *data visualization*, is a newer and ostensibly more inclusive term to describe everything from figures to infographics. Here, I adopt common terminology, such as bar plot, while also attempting to use the terms figure and data visualization for general reference.

Communications,[10] Wilkinson's *The Grammar of Graphics*[11] presents a unique technical interpretation of the structure of graphics. Wilkinson breaks down the notion of a graphic into its constituent parts—e.g., the data, scales, coordinates, geometries, aesthetics—much like conventional grammar breaks down a sentence into nouns, verbs, punctuation, and other elements of writing. The popularity and utility of this approach has been implemented in a number of software packages, including the popular `ggplot2` package[12] currently available in R.[13] (Although the grammar of graphics approach is not explicitly adopted here, the term *geometry* is used consistently with Wilkinson to refer to different geometrical representations, whereas the term *aesthetics* is not used consistently with the grammar of graphics and is used simply to describe something that is visually appealing and effective.) By understanding basic visual design principles and their implementation, many figure authors may find new ways to emphasize and convey their information.

THE TEN PRINCIPLES

Principle #1 Diagram First

The first principle is perhaps the least technical but very important: before you make a visual, prioritize the information you want to share, envision it, and design it. Although this seems obvious, the larger point here is to focus on the information and message first, before you engage with software that in some way starts to limit or bias your visual tools. In other words, don't necessarily think of the geometries (dots, lines) you will eventually use, but think about the core information that needs to be conveyed and what about that information is going to make your point(s). Is your visual objective to show a comparison? A ranking? A composition? This step can be done mentally, or with a pen and paper for maximum freedom of thought. In parallel to this approach, it can be a good idea to save figures you come across in scientific literature that you identify as particularly effective. These are not just inspiration and evidence of what is possible, but will help you develop an eye for detail and technical skills that can be applied to your own figures.

Principle #2 Use the Right Software

Effective visuals typically require good command of one or more software. In other words, it might be unrealistic to expect complex, technical, and effective figures if you are using a simple spreadsheet program or some other software that is not designed to make complex, technical, and effective figures. Recognize that you might need to learn a new software—or expand your knowledge of a software you already know. While highly effective and aesthetically pleasing figures can be made quickly and simply, this may still represent a challenge to some. However, figure making is a method like anything else, and in order to do it, new methodologies may need to be learned. You would not expect to improve a field or lab method without changing something or learning something new. Data visualization is the same, with the added benefit that most software is readily available, inexpensive, or free, and many come with large online help resources. This article does not promote any specific software, and readers are encouraged to reference other work[14] for an overview of software resources.

Principle #3 Use an Effective Geometry and Show Data

Geometries are the shapes and features that are often synonymous with a type of figure; for example, the bar geometry creates a bar plot. While geometries might be the defining visual element of a figure, it can be tempting to jump directly from a dataset to pairing it with one of a small number of well-known geometries. Some of this thinking is likely to naturally happen. However, geometries are representations of the data in different forms, and often there may be more than one geometry to consider. Underlying all your decisions about geometries should be the data-ink ratio,[7] which is the ratio of ink used on data compared with overall ink used in a figure. High data-ink ratios are the best, and you might be surprised to find how much non-data-ink you use and how much of that can be removed.

Most geometries fall into categories: *amounts* (or comparisons), *compositions* (or proportions), *distributions*, or *relationships*. Although seemingly straightforward, one geometry may work in more than one category, in addition to the fact that one dataset may be visualized with more than one geometry (sometimes even in the same figure). Excellent resources exist on detailed approaches to selecting your geometry,[15] and this article only highlights some of the more common geometries and their applications.

Amounts or comparisons are often displayed with a bar plot (Figure 1A), although numerous other options exist, including Cleveland dot plots and even heatmaps (Figure 1F). Bar plots are among the most common geometry, along with lines,[9] although bar plots are noted for their very low data density[16] (i.e., low data-ink ratio). Geometries for amounts should only be used when the data do not have distributional information or uncertainty associated with them. A good use of a bar plot might be to show counts of something, while poor use of a bar plot might be to show group means. Numerous studies have discussed inappropriate uses of bar plots,[9,17] noting that "because the bars always start at zero, they can be misleading: for example, part of the range covered by the bar might have never been observed in the sample."[17] Despite the numerous reports on incorrect usage, bar plots remain one of the most common problems in data visualization.

Compositions or proportions may take a wide range of geometries. Although the traditional pie chart is one option, the pie geometry has fallen out of favor among some[18] due to the inherent difficulties in making visual comparisons. Although there may be some applications for a pie chart, stacked or clustered bar plots

CellPress
OPEN ACCESS

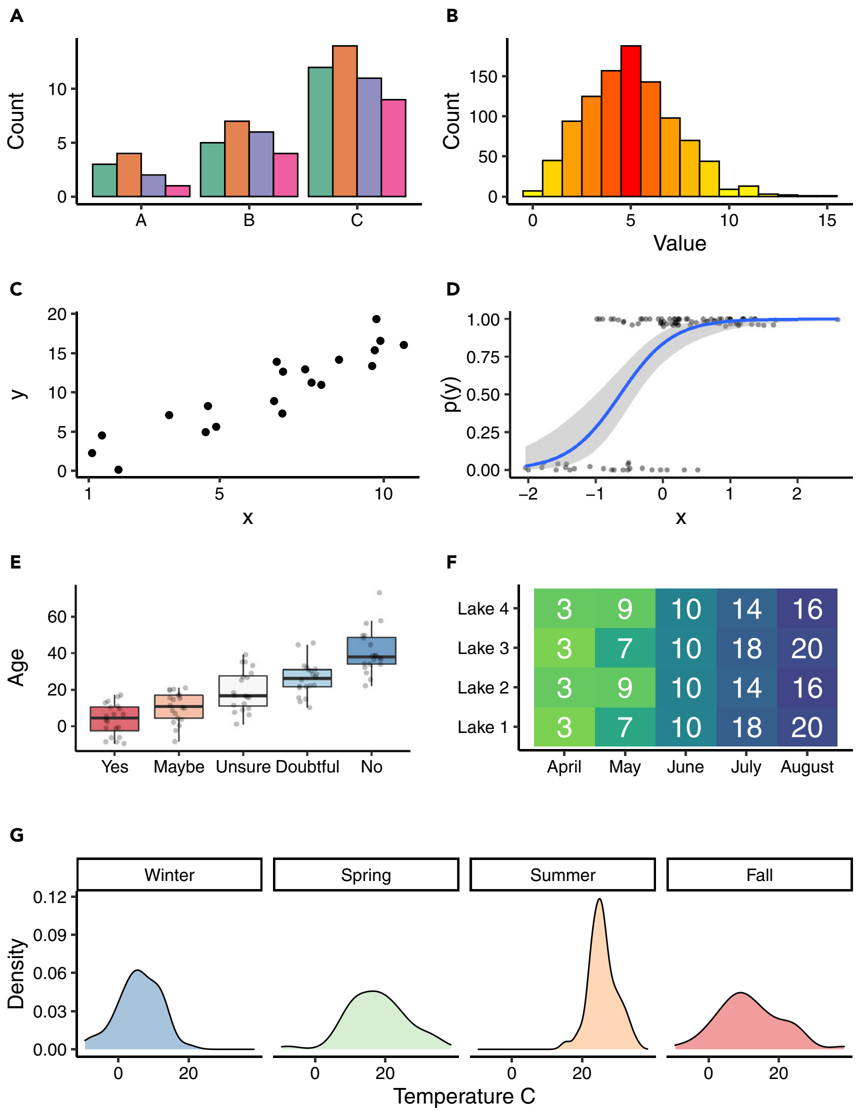

Figure 1. Examples of Visual Designs

(A) Clustered bar plots are effective at showing units within a group (A–C) when the data are amounts.

(B) Histograms are effective at showing the distribution of data, which in this case is a random draw of values from a Poisson distribution and which use a sequential color scheme that emphasizes the mean as red and values farther from the mean as yellow.

(C) Scatterplot where the black circles represent the data.

(D) Logistic regression where the blue line represents the fitted model, the gray shaded region represents the confidence interval for the fitted model, and the dark-gray dots represent the jittered data.

(E) Box plot showing (simulated) ages of respondents grouped by their answer to a question, with gray dots representing the raw data used in the box plot. The divergent colors emphasize the differences in values. For each box plot, the box represents the interquartile range (IQR), the thick black line represents the median value, and the whiskers extend to 1.5 times the IQR. Outliers are represented by the data.

(F) Heatmap of simulated visibility readings in four lakes over 5 months. The green colors represent lower visibility and the blue colors represent greater visibility. The white numbers in the cells are the average visibility measures (in meters).

(G) Density plot of simulated temperatures by season, where each season is presented as a small multiple within the larger figure.

For all figures the data were simulated, and any examples are fictitious.

(Figure 1A), stacked density plots, mosaic plots, and treemaps offer alternatives.

Geometries for distributions are an often underused class of visuals that demonstrate high data density. The most common geometry for distributional information is the box plot[19] (Figure 1E), which shows five types of information in one object. Although more common in exploratory analyses than in final reports, the histogram (Figure 1B) is another robust geometry that can reveal information about data. Violin plots and density plots (Figure 1G) are other common distributional geometries, although many less-common options exist.

Relationships are the final category of visuals covered here, and they are often the workhorse of geometries because they include the popular scatterplot (Figures 1C and 1D) and other presentations of *x*- and *y*-coordinate data. The basic scatterplot remains very effective, and layering information by modifying point symbols, size, and color are good ways to highlight additional messages without taking away from the scatterplot. It is worth mentioning here that scatterplots often develop into line geometries (Figure 1D), and while this can be a good thing, presenting raw data and inferential statistical models are two different messages that need to be distinguished (see Data and Models Are Different Things).

Finally, it is almost always recommended to show the data.[7] Even if a geometry might be the focus of the figure, data can usually be added and displayed in a way that does not detract from the geometry but instead provides the context for the geometry (e.g., Figures 1D and 1E). The data are often at the core of the message, yet in figures the data are often ignored on account of their simplicity.

Principle #4 Colors *Always* Mean Something

The use of color in visualization can be incredibly powerful, and there is rarely a reason not to use color. Even if authors do not wish to pay for color figures in print, most journals still permit free color figures in digital formats. In a large study[20] of what makes visualizations memorable, colorful visualizations were reported as having a higher memorability score, and that seven or more colors are best. Although some of the visuals in this study were photographs, other studies[21] also document the effectiveness of colors.

In today's digital environment, color is cheap. This is overwhelmingly a good thing, but also comes with the risk of colors being applied without intention. Black-and-white visuals were more accepted decades ago when hard copies of papers were more common and color printing represented a large cost. Now, however, the vast majority of readers view scientific papers on an electronic screen where color is free. For those who still print documents, color printing can be done relatively cheaply in comparison with some years ago.

Color represents information, whether in a direct and obvious way, or in an indirect and subtle way. A direct example of using color may be in maps where water is blue and land is green or brown. However, the vast majority of (non-mapping) visualizations use color in one of three schemes: *sequential*, *diverging*, or *qualitative*. Sequential color schemes are those that range from light to dark typically in one or two (related) hues and are often applied to convey increasing values for increasing darkness (Figures 1B and 1F). Diverging color schemes are those that have two sequential schemes that represent two extremes, often with a white or neutral color in the middle (Figure 1E). A classic example of a diverging color scheme is the red to blue hues applied to jurisdictions in order to show voting preference in a two-party political system. Finally, qualitative color schemes are found when the intensity of the color is not of primary importance, but rather the objective is to use different and otherwise unrelated colors to convey qualitative group differences (Figures 1A and 1G).

While it is recommended to use color and capture the power that colors convey, there exist some technical recommendations. First, it is always recommended to design color figures that work effectively in both color and black-and-white formats (Figures 1B and 1F). In other words, whenever possible, use color that can be converted to an effective grayscale such that no information is lost in the conversion. Along with this approach, colors can be combined with symbols, line types, and other design elements to share the same information that the color was sharing. It is also good practice to use color schemes that are effective for colorblind readers (Figures 1A and 1E). Excellent resources, such as ColorBrewer,[22] exist to help in selecting color schemes based on colorblind criteria. Finally, color transparency is another powerful tool, much like a volume knob for color (Figures 1D and 1E). Not all colors have to be used at full value, and when not part of a sequential or diverging color scheme—and especially when a figure has more than one colored geometry—it can be very effective to increase the transparency such that the information of the color is retained but it is not visually overwhelming or outcompeting other design elements. Color will often be the first visual information a reader gets, and with this knowledge color should be strategically used to amplify your visual message.

Principle #5 Include Uncertainty

Not only is uncertainty an inherent part of understanding most systems, failure to include uncertainty in a visual can be misleading. There exist two primary challenges with including uncertainty in visuals: failure to include uncertainty and misrepresentation (or misinterpretation) of uncertainty.

Uncertainty is often not included in figures and, therefore, part of the statistical message is left out—possibly calling into question other parts of the statistical message, such as inference on the mean. Including uncertainty is typically easy in most software programs, and can take the form of common geometries such as error bars and shaded intervals (polygons), among other features.[15] Another way to approach visualizing uncertainty is whether it is included implicitly into the existing geometries, such as in a box plot (Figure 1E) or distribution (Figures 1B and 1G), or whether it is included explicitly as an additional geometry, such as an error bar or shaded region (Figure 1D).

Representing uncertainty is often a challenge.[23] Standard deviation, standard error, confidence intervals, and credible intervals are all common metrics of uncertainty, but each represents a different measure. Expressing uncertainty requires that readers be familiar with metrics of uncertainty and their interpretation; however, it is also the responsibility of the figure author to adopt the most appropriate measure of uncertainty. For instance, standard deviation is based on the spread of the data and therefore shares information about the entire population, including the range in which we might expect new values. On the other hand, standard error is a measure of the uncertainty in the mean (or some other estimate) and is strongly influenced by sample size—namely, standard error decreases with increasing sample size. Confidence intervals are primarily for displaying the reliability of a measurement. Credible intervals, almost exclusively associated with Bayesian methods, are typically built off distributions and have probabilistic interpretations.

Expressing uncertainty is important, but it is also important to interpret the correct message. Krzywinski and Altman[23] directly address a common misconception: "a gap between (error) bars does not ensure significance, nor does overlap rule it out—it depends on the type of bar." This is a good reminder to be very clear not only in stating what type of uncertainty you are sharing, but what the interpretation is. Others[16] even go so far as to recommend that standard error not be used because it does not provide clear information about standard errors of differences among means. One recommendation to go along with expressing uncertainty is, if possible, to show the data (see Use an Effective Geometry and Show Data). Particularly when the sample size is low, showing a reader where the data occur can help avoid misinterpretations of uncertainty.

Principle #6 Panel, when Possible (Small Multiples)

A particularly effective visual approach is to repeat a figure to highlight differences. This approach is often called *small multiples*,[7] and the technique may be referred to as paneling or faceting (Figure 1G). The strategy behind small multiples is that because many of the design elements are the same—for example, the axes, axes scales, and geometry are often the same—the differences in the data are easier to show. In other words, each panel represents a change in one variable, which is commonly a time step, a group, or some other factor. The objective of small multiples is to make the data inevitably comparable,[7] and effective small multiples always accomplish these comparisons.

Principle #7 Data and Models Are Different Things

Plotted information typically takes the form of raw data (e.g., scatterplot), summarized data (e.g., box plot), or an inferential statistic (e.g., fitted regression line; Figure 1D). Raw data and summarized data are often relatively straightforward; however, a plotted model may require more explanation for a reader to be able to fully reproduce the work. Certainly any model in a study should be reported in a complete way that ensures

CellPress
OPEN ACCESS

reproducibility. However, any visual of a model should be explained in the figure caption or referenced elsewhere in the document so that a reader can find the complete details on what the model visual is representing. Although it happens, it is not acceptable practice to show a fitted model or other model results in a figure if the reader cannot backtrack the model details. Simply because a model geometry can be added to a figure does not mean that it should be.

Principle #8 Simple Visuals, Detailed Captions

As important as it is to use high data-ink ratios, it is equally important to have detailed captions that fully explain everything in the figure. A study of figures in the *Journal of American Medicine*[8] found that more than one-third of graphs were not self-explanatory. Captions should be standalone, which means that if the figure and caption were looked at independent from the rest of the study, the major point(s) could still be understood. Obviously not all figures can be completely standalone, as some statistical models and other procedures require more than a caption as explanation. However, the principle remains that captions should do all they can to explain the visualization and representations used. Captions should explain any geometries used; for instance, even in a simple scatterplot it should be stated that the black dots represent the data (Figures 1C–1E). Box plots also require descriptions of their geometry—it might be assumed what the features of a box plot are, yet not all box plot symbols are universal.

Principle #9 Consider an Infographic

It is unclear where a figure ends and an infographic begins; however, it is fair to say that figures tend to be focused on representing data and models, whereas infographics typically incorporate text, images, and other diagrammatic elements. Although it is not recommended to convert all figures to infographics, infographics were found[20] to have the highest memorability score and that diagrams outperformed points, bars, lines, and tables in terms of memorability. Scientists might improve their overall information transfer if they consider an infographic where blending different pieces of information could be effective. Also, an infographic of a study might be more effective outside of a peer-reviewed publication and in an oral or poster presentation where a visual needs to include more elements of the study but with less technical information.

Even if infographics are not adopted in most cases, technical visuals often still benefit from some text or other annotations.[16] Tufte's works[7,24] provide great examples of bringing together textual, visual, and quantitative information into effective visualizations. However, as figures move in the direction of infographics, it remains important to keep chart junk and other non-essential visual elements out of the design.

Principle #10 Get an Opinion

Although there may be principles and theories about effective data visualization, the reality is that the most effective visuals are the ones with which readers connect. Therefore, figure authors are encouraged to seek external reviews of their figures. So often when writing a study, the figures are quickly made, and even if thoughtfully made they are not subject to objective, outside review. Having one or more colleagues or people external to the study review figures will often provide useful feedback on what readers perceive, and therefore what is effective or ineffective in a visual. It is also recommended to have outside colleagues review only the figures. Not only might this please your colleague reviewers (because figure reviews require substantially less time than full document reviews), but it also allows them to provide feedback purely on the figures as they will not have the document text to fill in any uncertainties left by the visuals.

WHAT ABOUT TABLES?

Although often not included as data visualization, tables can be a powerful and effective way to show data. Like other visuals, tables are a type of hybrid visual—they typically only include alphanumeric information and no geometries (or other visual elements), so they are not classically a visual. However, tables are also not text in the same way a paragraph or description is text. Rather, tables are often summarized values or information, and are effective if the goal is to reference exact numbers. However, the interest in numerical results in the form of a study typically lies in comparisons and not absolute numbers. Gelman et al.[25] suggested that well-designed graphs were superior to tables. Similarly, Spence and Lewandowsky[26] compared pie charts, bar graphs, and tables and found a clear advantage for graphical displays over tabulations. Because tables are best suited for looking up specific information while graphs are better for perceiving trends and making comparisons and predictions, it is recommended that visuals are used before tables. Despite the reluctance to recommend tables, tables may benefit from digital formats. In other words, while tables may be less effective than figures in many cases, this does not mean tables are ineffective or do not share specific information that cannot always be displayed in a visual. Therefore, it is recommended to consider creating tables as supplementary or appendix information that does not go into the main document (alongside the figures), but which is still very easily accessed electronically for those interested in numerical specifics.

CONCLUSIONS

While many of the elements of peer-reviewed literature have remained constant over time, some elements are changing. For example, most articles now have more authors than in previous decades, and a much larger menu of journals creates a diversity of article lengths and other requirements. Despite these changes, the demand for visual representations of data and results remains high, as exemplified by graphical abstracts, overview figures, and infographics. Similarly, we now operate with more software than ever before, creating many choices and opportunities to customize scientific visualizations. However, as the demand for, and software to create, visualizations have both increased, there is not always adequate training among scientists and authors in terms of optimizing the visual for the message.

Figures are not just a scientific side dish but can be a critical point along the scientific process—a point at which the figure maker demonstrates their knowledge and communication of the data and results, and often one of the first stopping points

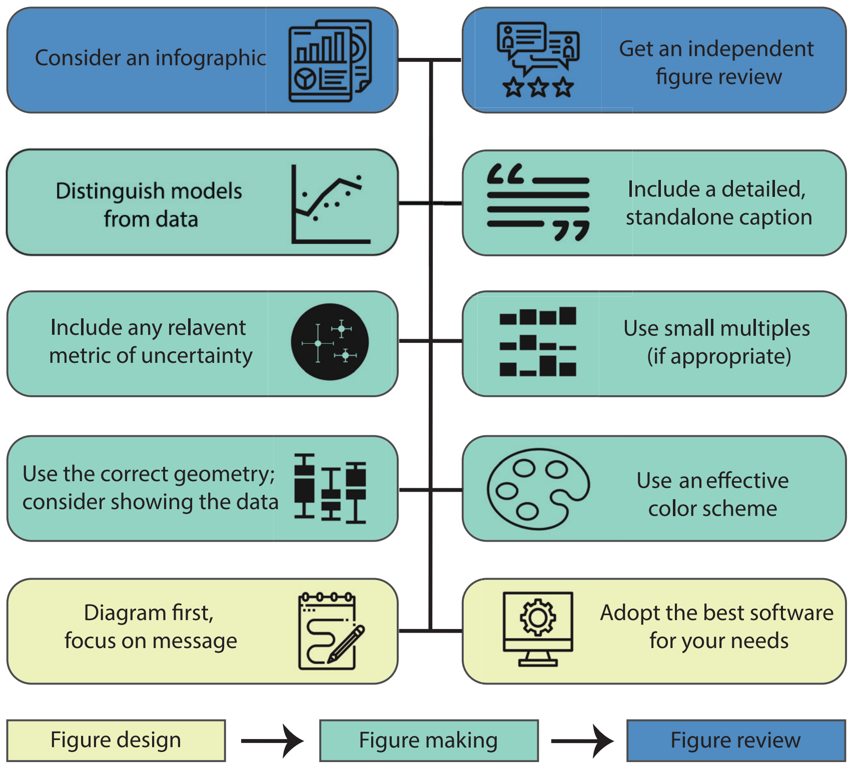

Figure 2. Overview of the Principles Presented in This Article
The two principles in yellow (bottom) are those that occur first, during the figure design phase. The six principles in green (middle) are generally considerations and decisions while making a figure. The two principles in blue (top) are final steps often considered after a figure has been drafted. While the general flow of the principles follows from bottom to top, there is no specific or required order, and the development of individual figures may require more or less consideration of different principles in a unique order.

for new readers of the information. The reality for the vast majority of figures is that you need to make your point in a few seconds. The longer someone looks at a figure and doesn't understand the message, the more likely they are to gain nothing from the figure and possibly even lose some understanding of your larger work. Following a set of guidelines and recommendations—summarized here and building on others—can help to build robust visuals that avoid many common pitfalls of ineffective figures (Figure 2).

All scientists seek to share their message as effectively as possible, and a better understanding of figure design and representation is undoubtedly a step toward better information dissemination and fewer errors in interpretation. Right now, much of the responsibility for effective figures lies with the authors, and learning best practices from literature, workshops, and other resources should be undertaken. Along with authors, journals play a gatekeeper role in figure quality. Journal editorial teams are in a position to adopt recommendations for more effective figures (and reject ineffective figures) and then translate those recommendations into submission requirements. However, due to the qualitative nature of design elements, it is difficult to imagine strict visual guidelines being enforced across scientific sectors. In the absence of such guidelines and with seemingly endless design choices available to figure authors, it remains important that a set of aesthetic criteria emerge to guide the efficient conveyance of visual information.

ACKNOWLEDGMENTS

Thanks go to the numerous students with whom I have had fun, creative, and productive conversations about displaying information. Danielle DiIullo was extremely helpful in technical advice on software. Finally, Ron McKernan provided guidance on several principles.

AUTHOR CONTRIBUTIONS

S.R.M. conceived the review topic, conducted the review, developed the principles, and wrote the manuscript.

REFERENCES

1. Stirling, P. (1987). Power lines. NZ Listener, 13–15.
2. Fleming, N.D., and Mills, C. (1992). Not another inventory, rather a catalyst for reflection. To Improve the Academy *11*, 137–155.
3. Lane, S., Karatsolis, A., and Bui, L. (2015). Graphical abstracts: a taxonomy and critique of an emerging genre. In Proceedings of the 33rd Annual International Conference on the Design of Communication, K. Gossett, ed.. https://doi.org/10.1145/2775441.2775465.
4. Csuri, C. (1974). Computer graphics and art. Proc. IEEE *62*, 503–515.
5. Wang, G., Gregory, J., Cheng, X., and Yao, Y. (2017). Cover stories: an emerging aesthetic of prestige science. Public Underst. Sci. *26*, 925–936.
6. Cleveland, W.S. (1984). Graphs in scientific publications. Am. Stat. *38*, 261–269.
7. Tufte, E.R. (2001). The Visual Display of Quantitative Information, *vol. 2* (Graphics Press).
8. Cooper, R.J., Schriger, D.L., and Close, R.J. (2002). Graphical literacy: the quality of graphs in a large-circulation journal. Ann. Emerg. Med. *40*, 317–322.
9. Weissgerber, T.L., Milic, N.M., Winham, S.J., and Garovic, V.D. (2015). Beyond bar and line graphs: time for a new data presentation paradigm. PLoS Biol. *13*, e1002128.

10. Nature Communications (2015). Nature Collections: Visual Strategies for Biological Data, the Collected Points of View (2010–2015) (Macmillan Publishers Ltd.).

11. Wilkinson, L. (2013). The Grammar of Graphics (Springer Science & Business Media).

12. Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (Springer).

13. R Core Team (2020). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).

14. O'Donoghue, S.I., Baldi, B.F., Clark, S.J., Darling, A.E., Hogan, J.M., Kaur, S., Maier-Hein, L., McCarthy, D.J., Moore, W.J., Stenau, E., et al. (2018). Visualization of biomedical data. Annu. Rev. Biomed. Data Sci. *1*, 275–304.

15. Wilke, C.O. (2019). Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures (O'Reilly Media).

16. Lane, D.M., and Sándor, A. (2009). Designing better graphs by including distributional information and integrating words, numbers, and images. Psychol. Methods *14*, 239–257.

17. Streit, M., and Gehlenborg, N. (2014). Bar charts and box plots. Nat. Methods *11*, 117.

18. Annesley, T.M. (2010). Bars and pies make better desserts than figures. Clin. Chem. *56*, 1394–1400.

19. Tukey, J.W. (1977). Exploratory Data Analysis, *vol. 2* (Addison-Wesley).

20. Borkin, M.A., Vo, A.A., Bylinskii, Z., Isola, P., Sunkavalli, S., Oliva, A., and Pfister, H. (2013). What makes a visualization memorable? IEEE Trans. Vis. Comput. Graph. *19*, 2306–2315.

21. Spence, I., Kutlesa, N., and Rose, D.L. (1999). Using color to code quantity in spatial displays. J. Exp. Psychol. Appl. *5*, 393–412.

22. Harrower, M., and Brewer, C.A. (2003). Colorbrewer.org: an online tool for selecting colour schemes for maps. Cartogr. J. *40*, 27–37.

23. Krzywinski, M., and Altman, N.S. (2013). Points of significance: error bars. Nat. Methods *10*, 921–922.

24. Tufte, E.R. (2006). Beautiful Evidence (Graphics Press).

25. Gelman, A., Pasarica, C., and Dodhia, R. (2002). Let's practice what we preach: turning tables into graphs. Am. Stat. *56*, 121–130.

26. Spence, I., and Lewandowsky, S. (1991). Displaying proportions and percentages. Appl. Cogn. Psychol. *5*, 61–77.

Steve Midway is an assistant professor in the Department of Oceanography and Coastal Sciences at Louisiana State University. His work broadly lies in fisheries ecology and how sound science can be applied to management and conservation issues. He teaches a number of quantitative courses in ecology, all of which include data visualization.