Writing a scientific manuscript

Zachary C. Holman, *Arizona State University*

IT'S no coincidence that you don't find poorly written papers in the best journals. In science, presentation is as important as content. There are many excellent commentaries available to aid you in crafting accurate, clear, and concise scientific manuscripts.[1,2,3,4,5] This article complements the existing literature with select opinions on what makes a good paper. It is far from exhaustive. Though the lessons may be generalized, the focus here is on presenting experimental physics, chemistry, materials science, or engineering results.

The writing process

Writing reveals the holes in your argument. If you, like I, have ever tried to do all of the experiments for a paper before beginning the writing process, you likely discovered that several key experiments were missing when you sat down to write. It is often a hassle to go back and fill in these missing puzzle pieces, especially if your samples age. George Whitesides convincingly argues that you should start an outline for a manuscript as soon as you have your first idea, and that this outline should be updated regularly to reflect your new results and changing hypotheses.[5,6] Let this living outline, which is a messy agglomeration of previous work, questions, and results, illuminate the holes in your research early, and *direct* your next experiments. You will find that you arrive at a complete story quickly and efficiently. Moreover, you will avoid awkward explanations about, e.g., how Figs. 8 and 9 show that longer annealing times increase mobility, except that actually the samples in Fig. 8 were synthesized at 10 Torr and 200 °C while those in Fig. 9 were synthesized in 8 Torr and 190 °C—plus they sat out in the air for an extra day before measurement—so that conclusions based on comparisons between the data are unfortunately a bit speculative.

Structure

The figures (and tables) are the most important part of a manuscript. Think about how you first inspect a paper a colleague has left on your desk: you glance over the abstract, then flip through the pages and look at the figures. If they look interesting and you can understand something about the work from the figures alone, you'll probably read the paper. As your living outline progresses, let the figures become the backbone of your manuscript, around which the text is molded. In the best papers, the text provides motivation, experimental details, and nuanced discussion of the results and their implications for the field; however, an educated reader can glean the major results of the work from the figures and captions alone.

Introduction

I have written several papers, and read many more, with first paragraphs that can be summarized as: "Nanoparticles are interesting." For review papers and articles in journals with particularly broad audiences, this sort of introduction is appropriate. For all other manuscripts, save your words. Don't write fluff that you would skip over if you were the reader. Your readers probably know your subject well, else they wouldn't have made it past the abstract. (You can include references to seminal papers or reviews if you are concerned that readers may need more background.) State the problem straight away, describe relevant previous work, and end with a short description of the significance of the present contribution. For a full-length paper, it is acceptable but rarely necessary to include an outline at the end of the introduction (e.g., "We begin with the theory of Lambertian light scattering, then progress to simulations of random textures in Section II."). Letter-length papers do not have distinct sections and outlines should be omitted. Similarly, I believe that speakers should refrain from giving outlines in 15-minute conference talks (does any audience member not feel that the speaker has just wasted a minute?), but I may be unique in this opinion.

Figures

What do you think of Fig. 1? Do the data *look* believable? Does the figure inspire trust in the authors' methods and results? Where was the figure likely published? *Nature*? *Proceedings of Unknown Conference*? How about Fig. 2? The data in the figures are the same; only the formatting differs.

Prior to my first conference presentation as a PhD student, my advisor asked to see my slides. After looking them over, he announced that my figures were garbage: the frames were rectangular, the tick marks pointed outwards, and I had used Times New Roman font. I had worked hard on my figures and was upset that they were deemed inappropriate based on what I believed to be my professor's peculiar stylistic taste. Later, I came to realize that this was one of the most important lessons of my degree. Even if I didn't agree with his aesthetic, my advisor opened my eyes to an aspect of scientific presentation to which I was previously blind. He forced me to consider what makes a figure look professional—down to the minor tick

[1] G.D. Gopen and J.A. Swan, "The science of scientific writing," *American Scientist* **78**, 550 (1990).

[2] T. Spector, "Writing a scientific manuscript: Highlights for success," *Journal of Chemical Education* **71**, 47 (1994).

[3] K. W. Plaxco, "The art of writing science," *Protein Science* **19**, 2261 (2010).

[4] M. Rolandi, K. Cheng, and S. Pérez-Kriz, "A brief guide to designing effective figures for the scientific paper," *Advanced Materials* **23**, 4343 (2011).

[5] G.M. Whitesides, "Whitesides' Group: Writing a paper," *Advanced Materials* **16**, 1375 (2004).

[6] G.M. Whitesides, "Publishing Your Research 101: How to write a paper to communicate your research," ACS Publications video series, http://pubs.acs.org/page/publish-research/episode-1.html.

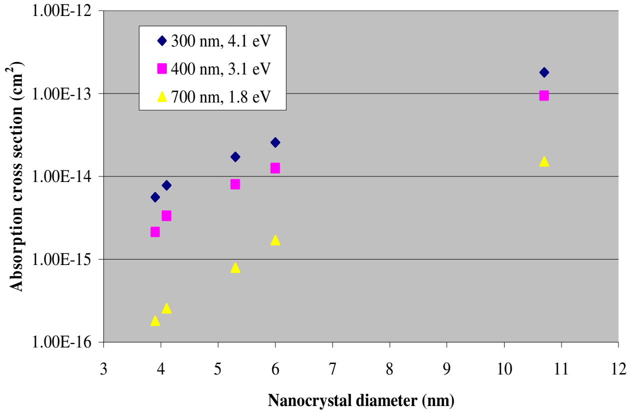

Figure 1. A poorly formatted, amateur figure made in Excel.

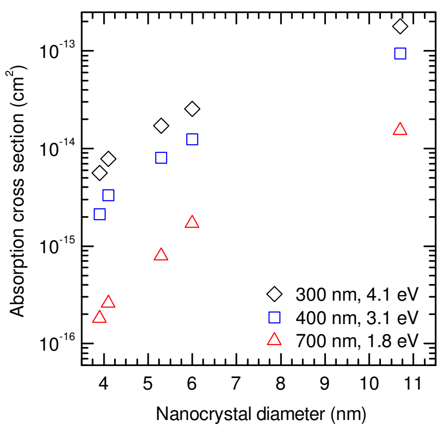

Figure 2. An aesthetically pleasing, professional figure made in Origin. The data are the same as those in Fig. 1.

marks. And, now I *do* agree with him: I prefer figures with square frames, tick marks that point inward, fonts that are sans-serif, no gridlines or background fill, primary colors (if colors must be used) instead of Excel's default colors, and data sets that are distinguishable even if printed in black and white. Figure 1 is displeasing exactly because it violates these guidelines, and more (e.g., the cumbersome number formatting on the y-axis). To make professional figures, don't waste your time with Excel, which gives you little formatting control. Invest the time to learn Origin, Igor, or SigmaPlot as soon as you start your research—it will more than pay off.

Good figures aren't just formatted well; they're also easy to understand. Consider Fig. 3—take a look now, before reading further. It displays the calculated reflectance and absorbance of a three-layer optical stack. Although abbreviations for incident angle, rear internal reflectance, absorbance in ITO, and absorbance in silver were introduced in equations in the text (θ, r_r, A_{Ag}, and A_{ITO}, respectively), the full names of the variables are retained in the axes titles so that the reader can comprehend the figure without referring back to the text. The simulated structure is drawn schematically so that, again, the reader can understand the results immediately. Furthermore, each part of the figure (a, b, or c) shows the reflectance or absorbance in one layer of the structure, and the plots are arranged in the same order as the layers in the structure.

The data in Fig. 3 could have been displayed in one plot, instead of three, but the figure would then have been too complicated. The plots could also have been displayed in three separate figures, but this would have de-emphasized their relatedness. Combine plots in a single figure if and only if they are complementary and the reader is intended to view them together. Tempting as it is, do not put two unrelated plots in the same figure to save space.

The three plots in Fig. 3 have the same x- and y-scales to facilitate comparison, and a blow-up inset is provided in Fig. 3c to provide detail that cannot be seen on the common y-scale. This inset retains the x-axis of its host figure so that additional labels don't need to be crammed in. The orange arrows and symbols at the top of the figure mark ranges of incident angles introduced in the associated text, thus anchoring the figure to the discussion.

The key to the figure is provided in the figure itself instead of in the caption. This way, the reader doesn't have to link the words "green, dashed" to the line that is green and dashed; he can just look and see. The data labels in the key are easy to understand, and the labels denote the parameter that distinguishes each simulated structure (in this case, the carrier density in the ITO layer). Avoid using the sample names you use in the lab in figure keys, even if you explain the corresponding conditions in the text or in a table. You don't want the reader to have to jump back and forth between the figures and the text. Finally, not only is the key in Fig. 3a used for all three figure parts, but the colors and line styles introduced in it are used to represent the same samples throughout *all* the figures in the associated manuscript.

A figure caption should provide a one-sentence summary of the figure's content, followed by important notes that are unique to the figure but too detailed for the text. A caption may include experimental parameters that apply only to the associated figure (default parameters should appear only in the experimental section), an explanation of markings in the figure that are not clear from the key (e.g., the orange arrows and symbols atop Fig. 3), or the statistical meaning of data or error bars.

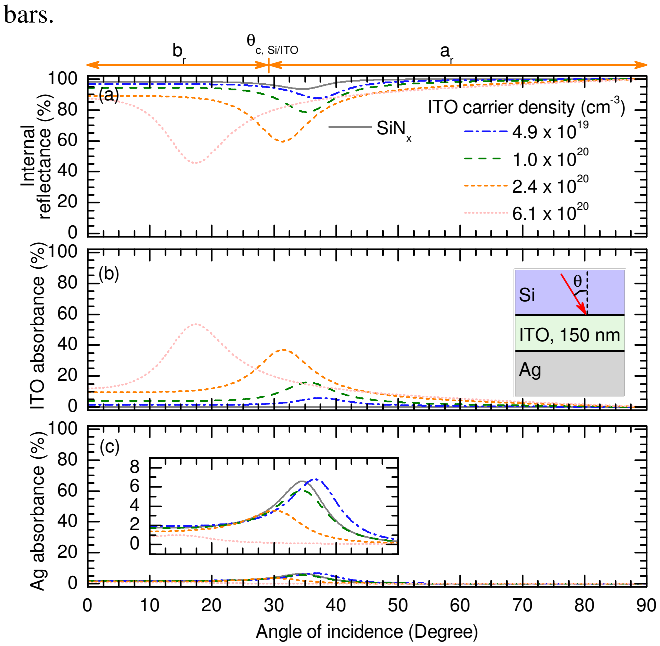

Figure 3. Can you deduce what this figure shows before reading the text?

SYMBOLS AND ABBREVIATIONS

When should you use symbols and abbreviations in your scientific manuscript? I constantly struggle with this question when writing. Remember this hierarchy as a guide: accuracy trumps clarity, and clarity trumps conciseness. Aim for all three attributes, but do not use abbreviations (which make sentences more concise) at the expense of clarity or accuracy. Sentences that are too thick with abbreviations are difficult to read, even for experts in your field. When you do introduce abbreviations, use them consistently throughout the manuscript, except in cases where doing so jeopardizes the hierarchy.

Symbols and abbreviations represent either quantities or lengthy terms of art. In the first case, introduce a symbol just before or after the equation in which it first appears. Use the symbol in subsequent equations, but not as a word replacement in the text. That is, do not replace every "temperature" with T, but consider "For $T > 300$ K" in place of "For temperatures greater than 300 K." In the second case—cumbersome terms of art—stick to abbreviations that are common (e.g., a-Si:H for hydrogenated amorphous silicon) or intuitive (e.g., MO-ATR for metal-overlayer attenuated total reflection). Avoid unnecessary abbreviations and abbreviations that do not suggest an immediate connection to the corresponding term of art (e.g., i.i. for ionized impurity).

CONCLUSION

It's too easy to throw away the conclusion. You've just finished writing the "real" part of the manuscript and you're ready for the process to be over. Why not just write a second, reworded abstract? Many authors do. Some papers deserve summaries in their concluding paragraphs—those, for example, that are more than ten pages long and have multiple, distinct arguments or sections. For all others, use the conclusion to situate your new contribution within the broader field, extrapolate to future experiments, or discuss implications for applications. If the last paragraph of a paper starts, "In conclusion…" or "In summary…," I skim or skip it. Keep the conclusion fresh and interesting; you want the end of the paper to sneak up on the reader so that she is engaged in the concluding paragraph before she realizes that it's the last.

REFERENCES

Citations serve two purposes: they either acknowledge specific previous contributions and give due credit to the authors, or they direct readers to sources with information outside the scope of the manuscript. The citations to George Whitesides's work in the second paragraph of this article are of the specific-reference, credit-giving variety. This kind of reference is easy: when you invoke an argument, claim, or result from a previous publication, cite it. The other kind of citation—the sort that guides readers to an optional, elaborating reference, and in so doing defines the content boundaries of your manuscript—is more difficult. Often, citations of this variety are used because the referenced material provides a specific example of a general statement. (See, for example, the first paragraph of this article.) In this case, include a wide range of relevant seminal contributions. Your citation list for a particular statement does not need to be exhaustive, however. The goal of "elaborating citations" is just to provide the interested reader with directions with which he may efficiently explore the larger literary landscape relating to your claim.