Paper formatting advice

Jack W. Baker

August 13, 2020

This document lists issues that often arise when writing papers. I developed this list as a reminder of items to check in my own writing, and as a guide for my advisees. The underlying concept is that your document should be clear and easy for your readers to follow. Sloppy writing and formatting will obscure your ideas and makes you seem less competent. So following the rules below will help your ideas shine through clearly.

These rules are essential for polished documents, but don't worry about them while writing your early drafts. I produce rough material while getting my draft organized, and go back later to clean all of these details up.

With most documents, I focus on a few key ideas, but this one is different because I try to be exhaustive. The point of a checklist like this is that you don't have to memorize it–you can always come back to this list as a reminder. There are many items to consider, so take one pass through your paper to check for each individual item. That way, you can concentrate on just one issue at a time, and don't risk missing anything.

A good strategy for finding solutions to situations not covered in this document is to look at good papers and see how they did it. Try to find papers that you found easy to read, and from the field or journal that you are targeting (because some conventions vary by journal). Then when you are trying to figure out how to do something in your paper, you can see how those papers did it.

1 Software tools

This section is a little different than the later ones, in that there are no correct or incorrect choices, but I include them because proper tools will aid your editing and formatting work. My personal tool choices are below.

1. I use Google Docs or Microsoft Word when drafting my outline or copy-and-pasting some ideas around. Their visual format is good for outlines, and Docs is good with collaborators, but I don't do any formatting there.

2. I switch to Latex when I am ready to write text. It handles the formatting, reference, and equation issues below very well. I think it is the clear winner for scientific writing at present. Overleaf is an easy Latex tool and excellent for collaborative work. I use TexPad and Atom for more complex documents if I am not collaborating, but there are many good alternatives.

3. I use Zotero to store reference information and PDFs, and the Better BibTex plugin to output a .bib file that I can use for citations in Latex. You should use a reference manager of some sort. Manually compilation of reference lists is inefficient and prone to errors.

4. I use Grammarly to help identify grammar issues. It doesn't fix everything, but it is quite good. I have a paid account and use it on everything I write. (I copy-and-paste my Latex

text into the Grammarly app and copy-and-paste it back when I'm done editing. Grammarly ignores most Latex markup and focuses on the text.)

2   Formatting for review comments

5. Double-space all text and leave big margins on the sides (~ 1.25"), to leave space for written comments by your advisor and coauthors.

6. Add page numbers, and ideally line numbers, to aid referencing specific parts of the document.

7. Put a date at the front of your document. This will help your collaborators make sure they are looking at the latest draft, and help you find the source document that your collaborators marked up.

3   Writing, formatting, and notation

8. Check for consistent notation and terminology throughout the paper. Did you always use the same variable and the same name to refer to a given quantity or concept? It may seem a little boring to keep using the same word, but it is beneficial to the reader if you always use the same word or variable to refer to a given concept. It is easy to accidentally invent new notation partway through a draft that doesn't match earlier content, so make sure to clean that up. Even minor inconsistencies such as 'damage grade' versus 'damage state,' or $Sa(T)$ versus $S_a(T)$ versus $SA(T)$, should be fixed to avoid confusion.

9. Use consistent notational conventions as much as possible. I prefer to use capital letters for random variables, lower-case letters for constants, bold symbols for vectors, and non-bold for scalars. Sometimes I run into problems that force the use of other conventions, but I only do it as a last resort and with some trepidation. And I err on the side of explaining my conventions as well (e.g., 'where boldface indicates a vector').

10. Consistently use either present or past tense verbs. Either can work, but don't switch between one and the other absentmindedly.

11. Use consistent capitalization in the paper title and heading titles. You can use sentence case (where only the first word is capitalized, like in this document) or title case (where each word is capitalized), but be consistent throughout the document. Some journal paper templates will also use an ALL-UPPERCASE style for titles. That's fine if the template uses that, but sentence or title case is ideal in most cases.

12. For each equation, define every variable in the equation. Look at each variable in each equation, and ask 'will a reader be able to figure out what this is?' If not, you need more documentation.

13. Use precise cross-references as much as possible. Instead of 'as discussed above,' write 'as seen in Figure 2.' Instead of 'using the above procedure,' write 'using Equation 3.' It is obvious to you what you are referring to, but don't make your reader guess. I usually do a word search on 'above' and 'below' as one technique for identifying and eliminating these cases.

4   Figures

Do the following for each figure:

14. Think about the message you are trying to communicate and what is needed to convey this message. Have you chosen the appropriate figure type and the necessary data to include?

15. Remove any markings that distract from the information you are trying to convey. Make sure that unimportant markings are shown using thin lines and do not take up too much space.

16. Make sure that every figure is cited and discussed in the paper. If a figure is not discussed, it does not belong in the paper.

17. Check that all items on the figure are labeled or explained.

18. If there is a legend, make the order of entries (from top to bottom) match the order of lines or symbols in the plot. Or even better, label the data directly rather than using a legend.

19. Choose appropriate numerical ranges for the axes.

20. Choose an appropriate color palette for the figure. `https://colorbrewer2.org` and `https://coolors.co/` are useful tools for choosing color schemes, and they have options to make sure the colors work well in greyscale or for colorblind people. Maintain consistent color schemes throughout the paper, to the extent possible.

21. Label all axes, including units. Ideally, give the text name of the property being plotted, the variable name, and the units. The text name makes it easier for the reader to understand what is being plotted without having to go back to the paper for a definition. The variable name (if there is one) adds an explicit connection to any related equations. See Figure 1 for a demonstration.

22. Legends, axis labels, and other text should have a similar font size to the rest of the text in the paper. Some plotting program's default font sizes are too small, so this requires a small effort to fix. See Figure 1 for a demonstration.

23. Use one or at most two font sizes in a figure. Ideally, all figures should have identical font sizes.

24. Use acronyms and abbreviations in figures with great caution. The reader may have to search back through your paper to find the acronym definition before interpreting the figure. You will often be surprised at how little space it takes to spell out your terms.

25. Make sure any variable names you use are consistent with definitions in the text.

26. Place a caption below every figure and omit titles at the top of the figure image. The caption should describe the figure. I put comments interpreting the figure in the body text, but some fields and journals include some in the caption; if you do this, think about consistency across the figures in your paper.

27. If there are subfigures, describe or identify each subfigure in the caption.

28. Write the caption in complete sentences.
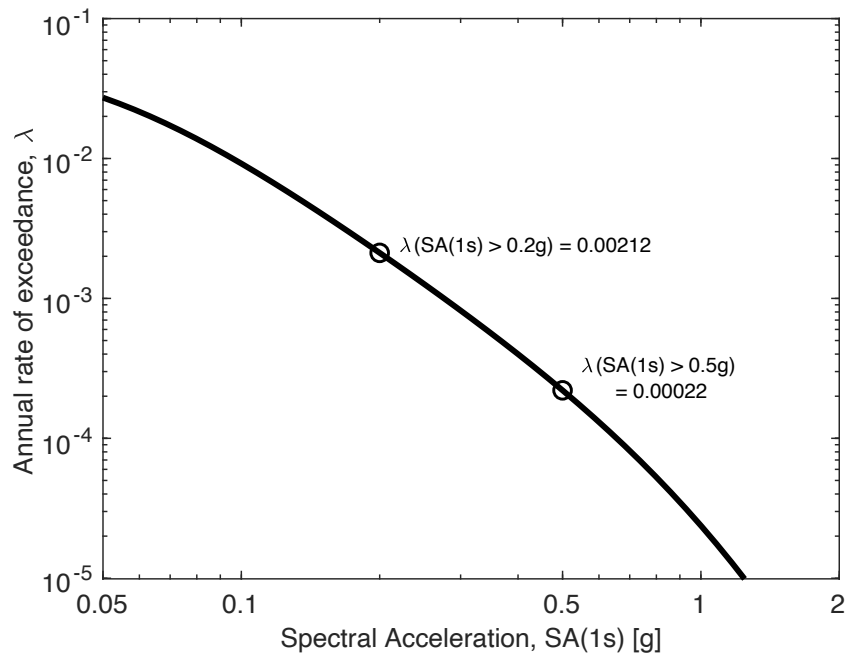
29. End the caption text with a period.

Figure 1: An example figure that illustrates axis labels and units. There are also only two font sizes in the figure, and they are similar to the text font size.

5   Tables

Do the following for each table:

30. Confirm that a table is the best format for the data and idea you are communicating. Often a figure is preferable. If the table is small, with only a few numbers presented, it may be preferable to list them in a sentence.

31. Consider the appropriate number of significant figures to report. Usually, only a few digits are appropriate. If the numbers are results from experiments or statistical analysis, consider how many digits are signals as opposed to estimation noise. And keep in mind that the more digits you report, the more difficult it is for a reader to skim the table and discern patterns.

32. Consider the appropriate alignment of each column. Columns of text are easier to read when left-aligned. Columns of floating-point numbers should generally be right-aligned or aligned at the decimal point. This will differentiate numbers of differing magnitudes.

33. Put a caption above the table.

34. The caption should be written in complete sentences.

35. The caption text should end with a period.

36. Avoid excess border lines in the table. A good practice is generally to have a horizontal line at the top of the table, a line following the header labels, and a line at the bottom (i.e., no vertical lines, and no horizontal lines between rows of data).

See Table 1 for an illustration of several of the above items.

Table 1: An example table illustrating several table formatting items. A caption, with complete sentences, is placed above the table. Numerical values are right-aligned, with an appropriate number of digits, to help the reader compare values.

Analysis Category	Result 1	Result 2	Result 3
Ordinary Least Squares	10.3	7.3	3.7
Weighted Least Squares	4.5	1.2	-1.2
Fisher Transform	24.2	9.1	0.3
Linear Regression	10.6	4.8	3.0

6   Citations

37. Use one of the following formats when you cite a reference:

    (a) Mention the author by the last name in the sentence and then give the year of the publication in parenthesis:

    > *According to Rodgers (2003), the 1906 San Andreas earthquake likely included super-shear rupture velocities.*

    (b) Give the facts or ideas mentioned by the author and then attribute these facts or ideas by putting their last name and the date in parenthesis:

    > *The 1906 San Andreas earthquake likely included super-shear rupture velocities (Rodgers, 2003).*

    In both cases, the author's first name is not used, the date is always in parentheses, and there are never two sets of nested parentheses. If you are mentioning an idea that is described in several places, and you are providing one example reference, precede the citation by an 'e.g.':

    > *Probabilistic seismic hazard analysis (e.g., Kramer, 1996) was used to obtain a ground motion hazard curve for the example site.*

    You can include a prefix in a Latex citation with the command `\cite[e.g.,][]{kramer_1996}`.

    Some journals use numbered references instead of the 'author-date' references shown above. In that case, the above reference text would look like '*According to Rodgers [1]...*' and '*super-shear rupture velocities [1].*'

38. Check that the citation information in your reference list is complete. Common sources of errors are reports with report numbers missing, and papers that were initially cited as 'in press' but are now published.