

How to Write Your First Paper

Tandy Warnow

December 7, 2015

1 Introduction

This document provides some very simple guidelines for writing papers, focusing on papers involving data (experiments, simulations, etc.), most likely for a conference, but potentially for a journal, where there is a real deadline. The target audience is students and postdocs who will be the first author of the paper, and so have the most responsibility for writing. If the paper has co-authors, then every step of this has to be done with substantial communication with your co-authors, and with input from your advisor(s).

2 Planning the paper

Planning is essential, especially if you have a deadline you want or need to meet. You need to know the deadline, the page limit, the word count (if any), and have read papers that appeared in the conference or journal to understand the style and level of detail that is expected.

2.1 Why are you doing this?

The first thing to focus on is *why*: what questions are you are trying to answer, and why are they important? What answers might you find and what difference will those answers make to the world? Even if what you are trying to do is to develop a new method and show how good it is, think in terms of the advantage (if any) provided by the method you develop, compared to others. Is this advantage going to be important? If the method won't provide any real advantage over the best current methods, then there is no point to the study (except maybe to learn that what you did doesn't work). So think about the motivation. Understanding the motivation and identifying the question will help you design the study appropriately.

2.2 What will you do?

Make sure you have figured out in detail *what* you will do - what datasets you will analyze, what methods will you use to analyze the datasets, and what

measurements you will make - so that you are answering the questions you have identified. Figuring this out well in advance of the deadline for any paper is essential, and you should discuss this with other researchers in your group (your advisor if you are a student, your collaborators if any). Make sure you are picking the best methods for your question, and appropriate datasets. Think about how long each analysis will do, and plan accordingly. Think also about statistical significance - will you have enough data for statistically significant results to be obtained?

2.3 Timing for papers that have data

At least 3 months before due date. Plan your paper - list the datasets, methods, and questions you will answer. This should be written in a detailed document that could be in your paper you will be submitting, even if only as the appendix. Specify enough detail that the experiment you are describing is reproducible¹. If you are doing a paper in which you are designing and testing a new method, make sure that any training data are disjoint from the test data. Get feedback on this document.

How early you need to do this depends on how computationally intensive your experiments will be, and what your computational resources are. You might need to start this earlier if you will need more than one month to generate your initial data, and more than two months to generate all your data. Start early, since data generation and analysis can take more time than you expect.

2 months before due date. Have an initial set of data to learn from at least two months before the paper is due. The initial set of data will help you decide what other data you need to generate and analyze, and that's why you need these data two months before the paper is due.

5 weeks before due date. The final set of data on which you will write your paper should be complete, so that after this point you will not need to generate and analyze data. You will still need to make figures and tables, do statistical analyses, etc., but basically henceforth you should just be writing the paper.

4 weeks before due date. Prepare figures and tables to show your results, and present your data to your research colleagues and get feedback in a 20 minute (or longer) talk. Have you missed some important cases? Do you need to generate more data? Hopefully the answers to these questions are "no", and you have all the data you need. Otherwise, prepare to get more data. But also, use the comments to help you think about how to understand the results you have obtained, and how they fit into the bigger research picture.

¹See Chapter 15 in "An Introduction to Computational Phylogenetics", by this author

3 weeks before due date. Have an initial draft of the paper, with the abstract, introduction, materials and methods, results, and acknowledgments written (i.e., everything but conclusions and bibliography). Include an initial set of figures and/or tables. Make sure this is in the style of the journal or conference you are targeting. Get feedback on your initial draft.

2 weeks before due date. Have revised the initial draft, and included the conclusions and bibliography. Revise figures. Get feedback on this draft.

1 week before due date. Have a submission ready version of the paper, without any typos, grammar problems, etc. Try to make the writing clear and elegant. Get feedback on this draft.

3 days before due date. Polish again based on feedback. Send out the new version.

1 day before due date. Submit. Do not wait until the last day! For conferences, you will often have the opportunity to resubmit, so just make sure you get the initial submission in on time.

3 Some objectives to follow

Rule 1: Make it easy to read and understand for the non-expert reader. You may think the purpose of your paper is to convince the expert that you've discovered something important, but your paper won't get published unless you can convince the editors and the reviewers - and they may not be experts. So, make your paper accessible to non-experts as well as to experts. In particular, be careful about where you go into detail; too much detail too early can lose your reader. On the other hand, too little detail is also a problem! So think carefully about where you need to provide more detail, and where the detail is not necessary. Remember, though - a paper is not a textbook.

Rule 2: Know your audience, and tailor your writing style for the audience. Understanding your audience is essential, and different audiences respond differently. For example, your paper may be of interest to both biologists and computer scientists, but you wouldn't explain things the same way to a biologist and a mathematician. But more importantly, think about what will motivate your reader - what will make him or her interested in your results, and able to appreciate what you are doing?

Rule 3: Motivate your reader! The most important part of writing is to engage the reader to care about what you are doing, and want to read the paper. So explain to the reader why the question you are addressing is important - not just at a very high level but in the specifics of what you are doing. For example,

if your work is about a new cancer therapy, the motivation you give the reader cannot just be that cancer is bad and new treatments are good, but rather about your specific new idea, and why it's likely to be substantially better than other treatments. Motivate, motivate, motivate! But also, be honest and don't inflate the importance of the work beyond where it belongs – the reader will notice this, and your reputation will suffer.

Rule 4: Organize the material carefully. Typically each journal provides a default structure for the organization of their papers. Follow this structure. Most of these organizations have sections labelled Introduction, Methods, Results, and Discussion; sometimes they also have other sections (like Background and Summary). Make sure your material is divided carefully and appropriately into these sections. In general, the Introduction needs to provide a very high-level context, and not too much detail, but it has to give enough information so that the reader understands the big question you are aiming at, and the specific questions you are addressing in the paper. The Introduction is key to most of the paper – getting it right makes the rest of the paper usually work well – so work on this carefully. The Methods section has to give enough information for the results to be reproducible; if you can't put it all in the Methods section, create an appendix and put the remaining material there. The Results section is where you put your actual results – not before. You may combine Results and Discussion in some cases, but if not then the Results section is mostly low-level observations rather than high-level inferences based on the individual results. To understand what goes where, read papers and see how they organize their content.

Rule 5: Follow journal or conference proceedings instructions rigorously. This is an easy one to do, but often ignored. Journals have rules (as do conference proceedings); just follow them. If they have a required bibliography format, follow it rigorously. They may have word count limits for the abstract, capitalization rules for section headings and subheadings; find out what they are and follow them. These instructions also sometimes specify how to organize the paper (e.g., Introduction, Background, Results, Discussion, and then Methods). Sometimes these rules can be broken, but sometimes not; check carefully before breaking them.

Rule 6: Proofread your paper over and over. Some people can proofread well from the screen, but others (like me) do best with hardcopy. Proofread over and over until you can't catch any more mistakes. Read it aloud to make sure you find awkward sentences. Don't forget to read the figure and table captions, and check the bibliography, too. Use a grammar checker and spell checker if you have access to one. Don't assume that the spellchecker found all the problems; you may sometimes have words that are incorrect but are real words (such as using "its" instead of "it's", or "their" instead of "there"). Read the paper carefully, over and over.

Rule 7: Find good examples of writing and try to emulate the best aspects of these papers. This is particularly useful for submission to journals or conference proceedings that have low acceptance rates. What is it you like about the writing? Think about it, and try to emulate it, if you can. Think about both low-level aspects of the writing (e.g., individual phrases, or how they transition between sentences) and high-level aspects (e.g., how the paper is organized, what is put in the introduction, how related work is introduced and discussed, how the limitations of the study are presented, how controversies are handled). Think in particular about how the author convinces you that the problem they are addressing is important, and that what they have discovered or done is fascinating. (For that matter, find poor examples of writing, and learn to avoid the mistakes those authors did in their papers.) Have a high standard for writing, and work to be better at it.

Rule 8: Never plagiarize - even unintentionally. Plagiarism is often unintentional; learn the rules and make sure you follow them. In particular, *never copy text without proper attribution*. For that matter, don't even paraphrase without attribution!

Rule 9: Use figures and tables appropriately and well. Figures and tables are often essential for communicating well; think about how to present the data to make the most important trends easy to detect. If space is not an issue, use multiple figures rather than one, so that each figure makes only a few important points. When presenting numeric values, think carefully about how many digits to present. Think about whether to present results in rows or columns – comparisons between numbers are easier when scanning vertically rather than horizontally. Use colors to your advantage, if allowed. If you are preparing your document in black and white, print it and see what it looks like printed – differences that are obvious on the screen may not be obvious when printed. Similarly, on the screen, figures can be enlarged and made easy to understand - this is not true for printed documents. Check font sizes for captions, legends, etc., in your figures, and make sure everything is big enough for your readers to be able to understand.

Rule 10: Pick your references carefully. It's best if you've actually understood the papers and books you cite! At a minimum, have read them, and select them for a purpose. Also make sure all your points are substantiated with appropriate references.

Rule 11: Be a good co-author. If you are a co-author, be a responsible one. Don't miss agreed on deadlines, and don't send text that has writing errors in it. If you are writing a latex document, make sure it compiles before you send your document.

Rule 12: Don't forget the acknowledgments. Don't forget to acknowledge your funding agencies (and give the grant numbers), and any other sources that contributed to your work (e.g., other people, time on supercomputers, etc.).

Rule 13: Read other advice about writing. Look for advice on how to write papers and how to write grants - it's basically the same advice, because the issues are the same. Here are some urls I found doing a google search for advice for writing papers:

- <https://contemplativemammoth.wordpress.com/2012/11/08/crowd-sourced-advice-for-writing-your-firstgrant/>
- <https://www.sci.sdsu.edu/~smaloy/MicrobialGenetics/topics/scientific-writing.pdf>
- <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3178846/>