1. Steps in writing a paper

Below is a standard protocol for paper writing, that should ease the process of getting a manuscript ready for submission (or of preparing a proposal). You will check in with relevant people in the group at every stage. (It is perfectly reasonable to also present a draft figure list or outline in group meeting). After each milestone is completed, you’ll update your paper status in the spreadsheet. Some part of the below is just good writing advice, some is workflow management. Writing is an unavoidable part of science, one of the few fields in which the practitioners must also communicate what they have done. It’s something you absolutely have to do well (and efficiently) to advance your career.

1. Define the boundaries of your paper – what subjects will you cover, what will you leave out? What is your conclusion? Who are your co-authors? Make sure each co-author knows that you are starting work. Identify who would be your main helper in writing the paper and get that person’s agreement to help (and enter them on the spreadsheet). Most co-authors will generally provide only individual analyses or figures and a quick read over once you have a draft. If you want someone to be involved as a helper during the whole process, you need to explicitly ask them to play a more significant role.

2. Choose an outside reader for your paper. This policy is a new experiment for the group, to help the writing process. Pick someone in the lab or climate groups who is NOT a co-author (or pick a generous friend). That person agrees to read various iterations of your paper draft and provide a reality check on whether it is understandable / convincing.

3. Do a literature search. You need to read all the background papers to understand how/whether your work contributes something new, to gather references you’ll use in the paper, and to figure out what kind of paper you want to write. Make sure the save the references that you’ll cite in EndNote or directly into a LaTeX.bib file. You probably want to download pdfs of each relevant paper you’ll be citing as well, and set up a directory where you can share those papers with co-authors (possibly in Dropbox).

4. Choose a target journal – identify the publication you’d like to submit to. Your literature search should help you decide what is the most appropriate journal. Once you have picked a journal, review the manuscript submission guidelines (usually in a web page called something like “For authors”). Identify the page limits and other restrictions (e.g. do they allow footnotes?), and figure out a reasonable number of figures. On the spreadsheet, enter the journal and your proposed n_figures and n_pages.
5. **Set up an svn archive** – a shared directory - for your paper. This will allow all co-authors to get all updates, and to make their own updates (editing the document, adding figures and references, etc.). Most journals will have a LaTeX template for manuscript preparation. Download this and place it in the archive. If there is no standard template, make one that fits the journal guidelines for submitted manuscripts.

6. **Make a figure list and figures.** Decide what figures convey information and tell a story, and then prune your list to match your target number. (All the others that may be relevant to some readers can go in Supplementary Online Material). If you can’t tell your story in the target number of figures, you have to back up a step and pick another target journal that allows longer papers. (But remember, shorter papers usually have higher impact). Make the figures – often you learn new complications when making figures, so it is worth making them well in advance of the writing. You probably want to write a draft figure caption for each figure at this time; it will only become harder and harder to remember what you’ve done as time goes on. (See Figure section for discussion of figure captions). Add your figures to the archive and insert them in the .tex file.

7. **Write an outline.** That means identifying the series of points you want to make, and assigning each of these points to a separate paragraph. (Then check the length – is the number of paragraphs consistent with your target paper length? If not, iterate). To ensure that your paper is a logical argument that flows naturally, every paragraph should have its own distinct point that is part of that argument. Check the flow of argument by stringing together all of your points: does it read like an abstract that tells a story? If not, reorganize until it does. Put your outline in the .tex file.

8. **Start filling in the outline.** You should not start with complete sentences. Almost all people find the writing process easiest if they simply concentrate on getting ideas and information down in the correct order rather than worrying about wording. That process actually helps generate short, clear, and informative sentences in the end rather than vague and wordy ones written before you’re sure what you’re actually trying to say. Note that as you continue filling in information, you may realize you need to change your outline – another reason not to obsess over perfect sentences at the beginning. Your “helper” should be reading the document as you work through this step.

9. **Smooth your filled-in outline into a readable draft**, in preparation for having others read it. If your structure is good, you should find linking ideas and information relatively easy at this point. Obviously put in all the references and results that you can, but it is OK to be missing the abstract at this point and to have a few blank spots or even a missing paragraph in your draft when you get your first readthrough. (I usually use XX’s for those: e.g. \cite{XX} means that I need a citation here, but haven’t found the right paper yet, or you may have XX’s for error bars at this stage, or a whole paragraph with just a lead-in sentence and
then “XX … need to fill in detail about the calibration here XX”). Use a consistent flag for missing information so that you can easily search for those places in the paper. Your co-authors would flag comments with the same symbol. It’s OK to give people a draft that is too long – generally your first effort will exceed the target page number – but estimate your required “compression factor” at this point. I.e. if you have a hard 10-page limit and your draft is 15 pages, then you need to compress by ~ 30%. Usually a 30% compression is fairly easy and actually makes the paper better. If you exceed your limit by a factor of 2 you may have to consider switching journals.

10. **Check the “Common Errors” listed below** and eliminate all you find.

11. **Do a structure check**: take the first sentence of each paragraph, and string them together into a single meta-paragraph. Is that meta-paragraph comprehensible? Does it flow well, and read like an abstract? If it does, the structure of the paper is good. If it doesn’t read well, you may have serious structure problems. Sometimes the problem is mostly solved by reordering. Sometimes it is requires deeper cuts and changing. If you can’t identify what point a paragraph is making, try eliminating the paragraph and see if nothing in it was actually necessary to your argument. If you find that something is now missing, then that something must have been the point of the paragraph. It can help to write a separate sentence explaining what are you trying to convey. Sometimes you will find that that sentence should be the first line of the paragraph. Identify and fix issues till you can pass the structure check.

12. **Begin iterating – get your outside reader or readers to comment.** When your reader is done reading the draft, ask what the main point is, and what advance you made on the previous state of the field. Assess: did they understand what you meant to say? Iterate till everyone agrees the paper is readable.

13. **Compress.** Even if you are within the page limits, take a pass through your paper to compress it. Remove repetitions. Identify any personal tics or idiosyncrasies you have and fix all instances. (Most people have at least one word that they overuse). Delete unnecessary sections. Finally, go through sentence by sentence and prune unnecessary words. Your outside reader can be invaluable in helping you cut and compress the paper. Usually the author gets attached to his or her words and believes each word is vital to the paper, but compressing usually improves the paper.

14. **Write your acknowledgments.** You want to acknowledge anyone who is not a co-author who gave you significant help on the analysis or the writing (including your reader). Acknowledge the projects or people who helped provide data. Acknowledge your funding sources (often giving the grant numbers).

15. **Pass the paper up to me (or Michael Stein) at this point.** It’s definitely OK to get input from me earlier, and I’ll always check the outline and figures to make sure you’re on the right track. If all has gone well and the paper is in reasonable
shape, the final stage won’t involve major reorganization or rewrites but will be more a polish with some minor additions.

16. **Review the journal submission guidelines** again. Identify suggested reviewers. (Most journals let you recommend a few people, and often let you also recommend that certain people NOT be reviewers). Write a draft cover letter to the editor. (Ask other students for examples). The cover letter briefly synthesizes the paper and tells the editor why the manuscript is important to your field and is relevant to this particular journal.

17. **Once you get the go-ahead, start the submission process.** Even if you’re not the corresponding author (generally the group leader is the corresponding author), you’ll log in as such for the purpose of submission. Some journals have easy submissions, but for others the total time to complete the forms can take several hours. Many journals want you to submit not a pdf but a .tex file and all the component files that go into building the pdf, including figures, .bib and .sty files, and have particular rules about naming of figure files.

2. **Steps and components of proposals**

The steps above are for papers; note that proposals have a *lot* of other pieces that need to be managed and completed: budgets, budget justifications, biosketches, current and pending statements, conflict of interest forms, work agreements from outside collaborators, internal documents needed for subawards to outside institutions, postdoctoral mentoring plans, broader impact statements, outreach components, etc. The actual science proposal in a proposal is usually 15 or 20 pages but the total package submitted can be 75+ pages.

Proposal submission generally involves two separate offices at the University of Chicago, the Physical Science Division (PSD) Local Business Center (LBC) and the University Research Administration (URA), as well as two computer systems, U. Chicago’s internal system (aura.uchicago.edu) and the system of whatever agency you’re applying to (many are now consolidated in research.gov). If the grant application is to a private foundation, there is also a separate U. Chicago Foundations Office.

Proposals that are in response to a particular agency Request for Proposals (RFP) will have an inviolable due date. Miss the due date, and it’s as though the proposal was never submitted. URA usually wants a proposal finished a full week before its due date. If pressed, they will reluctantly do a “contingent submission” after that time, which means that they transmit the proposal to the agency but do not guarantee that you actually followed the rules to make the proposal acceptable. Getting a font or margin or page limit wrong can result in a proposal being returned without review, so URA is not wrong to be worried.
3. Structure

Getting the structure right is the single most important part of writing a paper. Both papers and proposals have a kind of canonical structure. Begin with convincing the reader why he/she should care about your work, before you start talking about what you did. Because your abstract is a kind of shorthand version of the entire paper, that structure is reflected in the abstract as well.

*Paper abstract structure*
1. A is an important scientific problem that needs to be solved
2. It hasn't been solved before because of B
3. We now had the ability to solve it because of C
4. We did D, E, and F to solve it.
5. We conclude the solution is G
6. The solution has these important broader implications for science

For proposals there is only a slight modification, because the work has not yet been done.

*Proposal abstract structure*
1. A is an important scientific problem that needs to be solved
2. It hasn't been solved before because of B
3. We now have the ability to solve it because of C
4. We propose to do D, E, and F to solve it.
5. We may find either solution G or H, either of which is interesting
6. The solution will have these important broader implications for science

Even an instrument paper can follow this structure to some extent. The reader should know, for example, what scientific problems your instrument is intended to solve (1), and what allows you to advance measurement ability (2,3).

In the actual paper, points 1 and 2 and possibly 3 are part of your introduction. The introduction must also serve the purpose of describing the state of scientific knowledge in the field and reviewing previous work. (That review is part of both 1 and 2).

3. Overarching principles

*Define a science question that you are addressing.* You should be trying to answer a question that can be clearly articulated (and that ends in a question mark). You have to explain why your work is important if you want readers to invest time reading your paper. Even if your results are only suggestive, not definitive, still start with the question you’re trying to answer. (And even in instrument papers, you need to articulate the kind of questions that your instrument is designed to address.)
Write for your reader and respect your reader. You aren’t writing the paper for yourself. A paper is not a monument to yourself or a record of all your labors. It is a story written for other people. The only way to be a good writer is to put yourself in the position of those readers, so you can see how to give them the information they want in the order that it’s needed. You need to imagine your readers and see the paper through their eyes, to understand how someone not familiar with the details of your work would react to what you’re writing. You will almost certainly need to consciously practice this kind of identification. Initially, ask (beg) someone else to be an outside reader of your material, and then quiz them to see what they have understood. The results can be sobering.

Most science papers are badly written. Do not model your writing on the average science paper, since the average paper is bad. Model your writing on a paper that is especially clear and compelling to you. Often you will notice that the people who write clear and compelling papers seem to be the ones who rise to prestigious positions in science. That is not a coincidence.

Do not rest until you can explain something clearly. If you can’t explain an idea clearly, probably you don’t understand it fully, and that means your science is not complete. If you don’t understand it, your reader won’t either. Often the uncleanness is a scientific signal, that something is wrong with your mental framework. Scientific advances often come from worrying over a point of confusion until you find an underlying fallacy or misplaced assumption.

Your paper should be intelligible if someone reads only the abstract, the figure captions, and the conclusion. In reality most of your readers may not get much further than that. The abstract needs to tell the reader what science question you’re addressing, why it’s an important question, and what your answer is. (See Structure section above). The figure captions need to contain all the information needed to understand each figure. (See Figures section below). And the conclusions need to tell the reader what the wider implications of your research are, and what next steps would likely follow.

Write the minimal amount needed to tell your story. People have limited time to read; don’t waste it. For each detail you include, ask yourself – would the reader’s comprehension suffer if this detail were not included? If not, delete it. Is the detail relevant only to a subset of your readers? If so, put it in Supplementary Online Material (SOM). You need to at least identify all the analysis methods you used, and document every relevant assumption you made, but most readers are not going to want to go through your methods exhaustively. The few that do should be OK about reading it in SOM. (Or some journals have separate methods sections). On the micro-level, use simple sentence structure and choose short rather than long words where possible.
**Be thorough on your literature search.** You need to show both readers and reviewers that you totally understand the field and how your work fits into it. Be generous with citations. Unless you are really hemmed in by a page limit, there’s no reason to try to limit your bibliography. Remember that your reviewers will be looking to see if their papers are cited...I’d estimate that 50% of all problems during review come from not having read the literature as thoroughly as is needed. Even an initial literature search can take weeks. If the literature is too deep to cite everything, make sure you are citing the *earliest* paper that finds a given result, or that you’re citing a review paper and clearly marking it as review paper (“e.g. Smith et al 2007 and references therein.”).

**Show why your conclusion is valid.** Show how your results allow discriminating between competing hypotheses. (If you can’t discriminate between hypotheses, you probably shouldn’t be writing a paper.) It’s not enough to say “the model results look like observations”. You have to answer, how closely do your results resemble observations, and what does that similarity tell us? For an instrument paper or proposal, show what level of measurement precision and accuracy you *would need* to make a scientific finding, and then demonstrate that you can achieve that precision and accuracy.

**No surprises.** Don’t try to lead your reader through the same intellectual steps that you went through hoping that he / she will independently have an “aha” moment and reach the same conclusion that you did. Instead, deliver your message multiple times: in the abstract, probably at the end of the introduction, again in the body of your work, and expand on it in the discussion. Repetition is OK.

**Make sure the paper flows.** Examine how you are ordering ideas on all levels. The structure check helps you ensure flow from paragraph to paragraph, but check lower as well. If the paper seems to stumble, consider how you’re structuring individual sentences. When in doubt, start a sentence by echoing the previous sentence, and then proceed to something new: “A major scientific question is whether spacetime is curved or flat. Flat spacetime would be implied by $g > 1$.” Use parallel construction whenever possible – that is, if you’re introducing a second variant of an idea, use the same sentence construction that you did in when introducing the first variant. (“The first theory would predict X. .... The second theory would predict Y”)

**Listen to your readers.** Your reader may complain about something in your paper and suggest a change that you don’t like. You may feel that the reader didn’t understand, and want to disregard their specific suggestion. That’s OK, but even if you don’t accept the suggestion, you need to accept that there was a problem. If the reader had trouble understanding what you meant, that’s a problem that requires a solution. So work on finding a solution that satisfies both of you.

**Use consistent terminology.** It’s tempting to try to avoid repetition by changing up the words you use to describe the same phenomenon/approach/technique, but changes only confuse the reader. Minimize any changes that can produce confusion.
4. Figures

It’s important to sort out what information belongs in a figure or table caption and what information belongs in the text. In general

- The combined figure caption and legend should contain ALL information needed to understand what is in the figure: what the data is, how it’s been manipulated, what the units are, what each line or symbol means. Don’t make the reader hunt through the text for information. (If however you state e.g. the units in the axis labels, you don’t need to repeat them in the figure caption.)

- The text should contain NO statements that would be meaningless without looking at the figure. That is, if a piece of information is only useful to the reader while he/she looks at the figure, it belongs in the figure caption and not in the text (“The blue line shows...”).

- If information is already in the figure caption, it need not be repeated in the text. In some cases you may want to repeat a little, but overlap isn’t necessary.

Those rules push a lot of information into the captions and out of the text. Can you get away with such minimal discussion of your figures in the text? Yes, and you want to organize the paper that way. As a general principle, you want to be writing about your science, not about your figures. Make the text of your paper about the actual science, and only refer to figures as supportive of your findings. For example, it’s bad to start a sentence saying “Figure 5 top panel, blue line, shows two months of measurements of g...”, as the reader will most likely confront that sentence without having seen Figure 5. It’s better to just talk about your science: “We find that g > 1 in 90% of our experiments, suggesting that existing theories are incorrect (Figure 5)”. Very occasionally you’ll find that you have to discuss a figure explicitly in the text, but generally you don’t need to.

On the graphics themselves: good graphics are extremely important to reader understanding and so worth spending time on. Try to use minimal “data-ink” to make your point: no clutter. The books of Edward Tufte (in the Liberatory) have good advice about graphics. Keep enough whitespace that your figure is not confusing. (If making a graphic elegant is difficult in your plotting language, you can make a pdf and then modify the graph further in Adobe Illustrator.) The best font size for readability is always much larger than you think it would be – keep increasing the font size til it’s definitely too big, and then back off just a little. Use bright color and clearly differentiable symbols, and filled instead of open symbols wherever possible. Try to make lines/symbol groupings that help the reader intuitively classify the data that you’re showing. For example, if you’re comparing a large number of different experimental and modeling predictions, you could use red lines for all experimental data, with different dashes for the different experiments, and black lines for all modeling results. Or you could use solid lines for model results and dashed for experimental results, with different color codes for individual studies. No reader wants to keep referring back to a legend with 30 different elements... try to provide a simple code so that the reader can intuitively understand what you’re presenting.
5. Classic mistakes to check for:

**Writing issues**

- **Sentences with no subjects: passive tense.** Sentences should have subjects (e.g. say “We adjusted the laser” instead of “The laser was adjusted”). Generally one can avoid passive tense altogether. If you find that you’re writing too many sentences beginning with “We” then find a different way of stating these ideas rather than removing yourself from the sentence. Or, consider whether the overuse of “we” is a clue that you are spending too many sentences describing your experimental procedure.¹

- **Sentences with “This” or “That” as a subject.** Don’t make the reader have to guess what your subject is. (e.g. “This suggests that existing theories are incorrect”. Better to use a subject: “Our measurement of $g > 1$ suggests that existing theories are incorrect.”) If you feel that adding in a subject would make your writing repetitive, the problem may be that you’ve broken an idea that really belongs in one sentence into two sentences, that should be consolidated.

- **Overuse of adverbs.** If a word ends in “ly”, be wary. Look at your sentence and ask: how would the sentence read without that word? Is the adverb adding meaning, or is it just decorative? Or worse, is the adverb substituting for a persuasive argument? It’s easier to pump up your sentence and claim a result is “hugely significant” than to demonstrate to the reader that the result is in fact hugely significant. Claims made with a dramatic word are not proof, and the reader will know it, and will be suspicious.

- **Lists instead of persuasive arguments.** A list may seem efficient, but usually doesn’t communicate much to the reader.

- **Undefined acronyms.** The first time you use an acronym in a paper, it should be defined, even if you are pretty sure your readers would understand it. (“The fourth Intergovernmental Panel on Climate Change (IPCC) report (AR4)...”). Remember that even when people think that know an acronym, they may still be confused. SOA means secondary organic aerosol to some. To others it means “state of the art”.

- **Undefined terms.** If you’re going to use a symbol, you have to define it.

- **Unnecessary jargon or detail.** Don’t try to impress the reader by using technical jargon that he/she is not likely to understand. The point of the paper is not to assert that you’re smart, it’s to convey information and insight. (This rule

¹ Note: in exceptional cases, when a scientific issue is deeply controversial, you may strategically deploy passive tense to make your writing vague and inoffensive and to disavow responsibility for a discovery. That is, you can selectively use bad writing to minimize the shock of the new theory you are advancing.
holds for proposals as well). Include equations only if they are necessary to the argument, and only include them if you’re going to meaningfully refer to them again in the paper. Your default should be to write as simply and clearly as possible. Only use long technical words if they are absolutely necessary (and even then you may want to define those technical terms with some more evocative and intuitive explanation.) The more jargon-y your paper is, the smaller your potential readership, and the less compelling the paper even to those who understand it.

- **Sentences beginning with “Figure X shows….”** See Figures section above.

**Mistakes on order and location of information**

The abstract (and paper) begin with detail on the subject of the investigation rather than on the science question that your investigation is trying to answer. Don’t launch into detail before you have shown the reader why he should care about that detail. For example, opening a paper in a climate journal with “*Histonia* is a rare type of slime mold that reproduces only when temperatures exceed 30 C for a two-week duration” leaves the reader unsure why he should care or read further. If however you begin with “*The exact sensitivity of the climate to radiative forcing remains an open question. Recently several investigators have theorized that the behavior of slime molds contains fundamental clues to climate sensitivity*”, then he knows why he should care.

Another example, drawn from real student writing, as the first sentence of an abstract: “*The bulk properties of cirrus clouds, including radiative impact, are determined by the microphysical processes (heterogeneous nucleation, homogeneous nucleation, growth, evaporation, sedimentation and aggregation) that occur during cloud formation and evolution.*” A better first sentence would be something like “*Possible changes in the distribution of cirrus clouds and their radiative properties are one of the largest sources of uncertainty in climate forecasts. Both the distribution and radiative properties are affected strongly by the microphysics of growing ice particles*” (The example also stands as an example of one of the writing issues, substituting lists for explanations.)

**The work is not motivated by a science question.** Frequently I read proposals (and increasingly also papers) in which the sole justification for the work is that “it will improve parameterizations in models”. In other words, the authors did some technical work that they expect/hope will eventually be used by someone else to do some science, but they can’t articulate what that science is. Why would you trust that the author knows what an “improvement” is, if he/she doesn’t know what it would be used for?

**Chronological writing.** “First we did X, then we did Y, then we did Z”. The reader doesn’t care what order you did things in. The reader cares about why your problem is important, and what you learned, and wants to see just enough procedure to be confident that you didn’t mess up the analysis. Think of how to tell the story of what you learned, not the story of what you did.

Copyright E. Moyer 2014