Writing a clear and engaging paper for all astronomers

Abstract: Scientists usually receive no formal training in how to communicate effectively scientific information. What little training we do get comes from our PhD supervisors, who may or may not be good communicators themselves. Moreover, too many scientists seem to feel that the goal of scientific writing is to impress others with the author's intelligence, and most of the rest forget that even people in closely related fields may not be aware of the jargon, background and technical details specific to each subfield. Yet the principles of clear writing are easily grasped, and with a little practice will become natural to implement. Even in a technical journal the audience is not restricted simply to a few direct competitors, so you need to explain why the general topic is interesting, what problems there are in the field, what you have done and how it has helped advance us towards the resolution of one or more of the problems.

It is a sad commentary on our abilities as scientists to communicate effectively that many of the "crank" papers submitted to Nature are actually better written – from a purely stylistic point of view – than many of the professional papers. The most common and most easily fixed error by professional scientists is to assume too much knowledge of their audience. At a minimum, every paper should explain in clear and simple language the context within which the work was done.

You should never underestimate the work involved in writing a good paper. Bad papers are easy to write, and almost inevitably a bad paper will be longer than a good one. In order to write a good paper, you need to look carefully at what you want to accomplish: what important message do you want readers to take away from the paper? Once you have made that decision, it is easier to write with a tight focus.

My advice below on how to write a good paper is inevitably influenced by the fact that I work for Nature; I happen to think that the 'formula' that has evolved at Nature over the last 130+ years is the best and most efficient way of presenting information to an audience. My own ApJ papers now are written more in the form of a Nature paper; not in the specific layout, but in the overall philosophy of how best to convey the information.

Abstracts are even more important on the web than in print

Many people write an abstract almost as an afterthought to the paper, without recognizing its importance to the goal of communicating effectively. But the abstract really is the invitation you give to readers, hoping to entice them to read further.

To see the importance of the abstract, you need to assess the human psychology of reading papers. For example, how do you read the literature? You probably follow the general practice of skimming the abstract to see if you want to read further. If so, then you'll read the introduction and the conclusions. Only if the work is immediately and directly relevant to your current research will you take the time to read the entire paper, and even that might be deferred until you absolutely need

to, such as when you are writing your own paper and comparing your results to those published earlier.

The web is emphasizing these characteristics, rather than mitigating them. Now, all you see in front of you when you look at the new submissions to astro-ph or do an ADS search are the abstracts. The potential barrier to reading further has been increased over what it was in print journals, because sometimes a remark, equation or reference would catch your eye lower on the page and you would continue. That is no longer possible, so you must engage the reader immediately with an attractive and comprehensible abstract.

Let's make the discussion more concrete by using some examples. A typical abstract in an astronomy journal will look like this:

"We used [telescope x] to measure the [technical property] of source(s) [y]. The [technical property] differs from that [measured by, or predicted by, z]. This has implications for our understanding of [a]."

While this is certainly abbreviated, and many abstracts contain a lot of jargon and technical details, I believe that I've captured the usual essence. The problem with such a structure is that no clue is given as to why the source is interesting, or what in fact has been learned. We frequently see in the literature the phrase "has implications for …", but it conveys no information.

Here is what an abstract should look like.

"Sources such as [y] are interesting/important because [provide an explanation]. Particularly crucial to our physical understanding is a measurement/calculation of [z], because that will tell us [b]. In the past, it has been difficult/impossible to accomplish this, because [generally, equipment was inadequate]. Now we have measured/calculated [z] and find that it is/is not as expected. In the light of this result, we can now determine that our understanding of the physical processes underlying [b] is/is not complete. We have accordingly determined that [relate your discussion back to why the source is interesting, to give your readers a sense of progress towards a goal].

With an abstract structured like that, even if the result is not immediately relevant to most readers, they have in fact learned something new about your subject even without reading anything else in the paper. You might even get them to read further.

Do not try to make your abstract too detailed – that isn't the point of an abstract. Too many technical details – and too much jargon – will simply obscure the "big picture" that you want to convey. Notice that I include telescope name as a technical detail – it is omitted in my abstract. The reason is that in general people simply don't care where the data came from; what they want to

know is what the data mean for our broader physical understanding. There are of course some exceptions to this, but “yet more HST data” is not one of them.

Scientific papers have evolved over time into a fairly standard structure of introduction, observations/computations, results, and discussion. There are nuances and variations, but for many reasons this is the generally accepted format. But most authors do not make good use of the acceptance of this standard format to make their main points stand out clearly in the text.

Section headings should make a point

As every scientist knows the general format of a paper, you’re relieved of the burden of identifying your sections with the usual boring headings. Instead, you should take the opportunity to make a point, or highlight something from the section to come. Rather than “Observations”, you could say something like “VLA map(s) of [y]”, or even “Observations of [y] with the VLA”. Instead of “Results”, you could say “Detection of the first multipole peak”, and in place of “Discussion” you could proudly state “A gamma-ray burst at a cosmological distance”.

The issue of pride in your work is, I believe, both unnecessarily contentious and misunderstood by many scientists. There are no books anywhere on technical writing that encourage you to write in the third person impassive! In this chapter I am using the active mode as a means to engage you – the reader – more fully in what I am saying.

Many people believe – with considerable justification – that the egos of some scientists are so large that the only way to keep them under control is to force them to use the third person. In fact, I think this fails miserably in its aim, and along the way produces turgid papers. So, forget about it. You are justifiably proud of what you have accomplished as a scientist, so write in the first person and let some of your excitement show in the paper.

Caveats and qualifications

Most scientists feel compelled to qualify their conclusions with caveats and possible reasons why they might be wrong. To a large extent this is simple intellectual honesty – we often do not understand well the phenomena about which we are writing, so it is only fair to point out to readers potential weaknesses in our arguments.

Sometimes, though, it is carried to extremes. While this is less true in astronomy than in many other sciences – where public policy decisions may well be made based on research in biology, medicine, or climate control, to name a few prominent examples – it is not uncommon to see a significant fraction of the volume of a paper devoted to explaining the ways in which the authors might be wrong. The upshot of such an extended discussion of possible alternatives may well be that the main message of the paper is obscured, and the reader will be left wondering just what to believe.

In astronomy, this is rarely justified. I see it most often when people are simply trying to mark out a bit of territory for themselves, but with sufficient wiggle room that it would be difficult to say that they are wrong if they are eventually shown to be incorrect. While this serves the author's immediate selfish purpose, it is hard to see how such a practice benefits science as a whole.

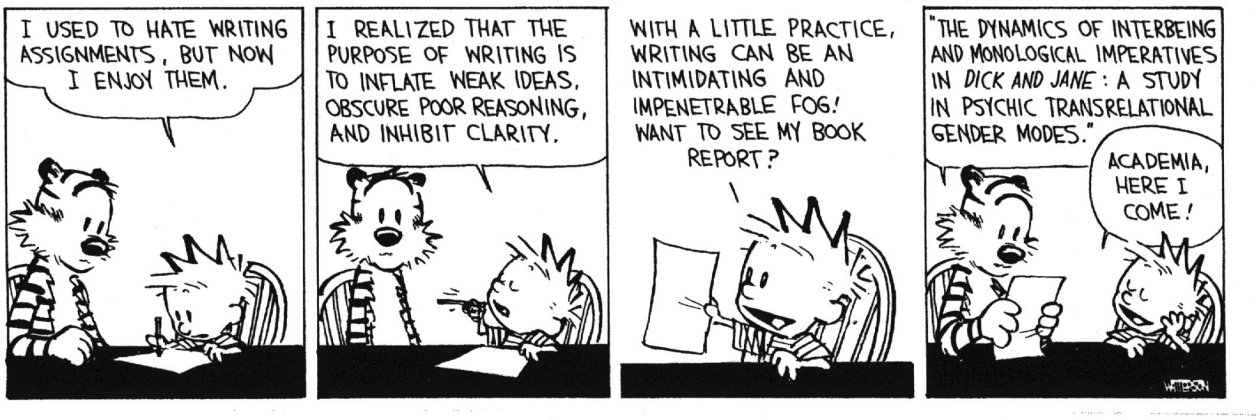
On the other hand, it also is a disservice to the community to claim too much of the data, though because science is self-correcting, any grandiose claim will immediately come under scrutiny.

Say what you mean and say it clearly

As scientists, we generally receive no training in how to write a good paper. We learn by example when we read the literature, but as most papers are badly written that simply perpetuates the problem. Or we learn from PhD advisors, who even if they can write well, sometimes do not really know why they are doing so, and therefore have trouble teaching us why they are successful. Some scientists who are appallingly bad writers even believe that they are good.

Finally, there is a culture of being intentionally obscure. This is most obvious at rather senior levels – I have had arguments with people who firmly believe that their papers should be comprehensible only to a few other people in the world – but younger people tend to adopt the positions of their seniors. While I am generally gratified that attitudes are changing slowly, and that more value is being placed on effective communication, the sad fact is that there still are too few people with the tools and knowledge to communicate well.

Let me illustrate my point with one of my favourite cartoons. The little boy, Calvin, is a bright and perceptive observer of humanity. His companion, the tiger Hobbes, comes alive only when they are alone. The reference to "Dick and Jane" brings to mind the first books that people of my generation and older used in the first year of school in North America to introduce reading: "See Dick. See Jane. See Dick run. See Jane run…"



Most scientific papers are written in the style of Calvin's book report. They use ten words when two would be enough. Rather than speaking plainly and to the point, authors try to be erudite, but often achieve only obscurity.

The best way to write well is to put yourself in the place of a reader, and ask what they most care about. Any general reader will want to know why the topic is important – otherwise, why should he or she devote any of their time to it? Next they will need to know what problems or issues are important to the field. This will set the stage for your own work, when you tell readers what you have done. Then the paper should end with a clear description of how your work relates back to the problems and issues you mentioned earlier. Have you solved any of the problems, or has the mystery deepened? Or perhaps you have discovered something startling or unexpected. Whatever you have achieved, it has to be placed firmly in the wider context of the field, so that readers will get a clear sense of progress towards a goal.

I am often asked if papers authored by people for whom English is not their first language are at a disadvantage in the peer review process. The emphatic answer is "no" – it is exceedingly rare that an author's weak grasp of English is relevant to either referees' assessments of the science or to an editorial decision. In fact, two of the worst written papers I have seen in my time at Nature came from native-English speakers based at a major UK university.

Summaries often serve little purpose

"Tell them what you are going tell them, tell them, then tell them what you told them." This often-repeated phrase is used to justify some of the worst atrocities in scientific writing and oral presentations. Many papers or talks will begin with an outline that goes something like "the observations are discussed in section II, the results in section III, and we discuss what they mean in section IV". Unless a paper goes beyond 15 or so pages, there is no need for such an outline, and even in long papers only the different and exceptional should be highlighted.

Summaries often restate – in the same or very similar words – what has already gone before. There generally is little point in that (though it can occasionally serve a purpose in very long papers). What is much more interesting is to express your thoughts on what specific important problems remain, and what must be done to resolve the issues. Vague statements such as "this has implications for …" are useless. What are the implications, and why are they important? Similarly, vague appeals for more data also are a waste of space. What observer does not need more data? The relevant issue is what specific data are needed to solve the specific remaining problems.

Once you have said what you wanted to say, stop.