



A manuscript checklist for improving science

JONATHAN PELLE¹

1. Washington University in Saint Louis

Summary: To encourage each other to do better science, my lab has put together a checklist of goals and guidelines for scientific papers. Our focus is on encouraging best practices for open science and statistical rigor.

➤ READ REVIEWS

✍ WRITE A REVIEW

CORRESPONDENCE:
peellej@ent.wustl.edu

DATE RECEIVED:
March 23, 2016

DOI:
10.15200/winn.145877.77781

ARCHIVED:
March 23, 2016

KEYWORDS:
open science, publishing,
statistics

CITATION:
Jonathan Pelle, A manuscript
checklist for improving science,
The Winnower
3:e145877.77781, 2016, DOI:
10.15200/winn.145877.77781

© Pelle This article is
distributed under the terms of
the [Creative Commons
Attribution 4.0 International
License](#), which permits
unrestricted use, distribution,
and redistribution in any
medium, provided that the
original author and source are
credited.



I've been thinking a lot recently about how to do better science. Over the years I've been inspired by many researchers on topics related to good scientific practice, and how to increase the quality and reproducibility of scientific research.

One particularly appealing initiative is the [Peer Reviewers' Openness Initiative](#) (PRO), which I've [written about previously](#) (Pelle, 2016). The logic behind the PRO is that researchers respond to incentives, and as reviewers we can push authors towards more open (and thus hopefully better, and more reproducible) science. I like this approach both because I'm generally convinced open science is better science, and because it helps me focus on things I can control (for example, the content of my peer review).

The PRO provides [guidelines for openness](#) that we, as reviewers, can ask (or require) authors to follow. I found this list very helpful—and I realized that, as a lab, we haven't had any documented best practices for scientific publishing. I fear this has meant that we haven't incorporated as many advancements as we could have in our own work, both in terms of openness, and in terms of statistical best practices.

To begin to remedy this, I've put together a checklist for scientific papers ([pdf](#)) that we'll start using whenever we submit a manuscript. Our checklist covers both scientific openness and statistical rigor: each item has a "yes" or "no" checkbox, with some thoughts and readings that have shaped my own thinking on the following pages.

From my perspective, it would be ideal if all of the items on the checklist were checked "yes" for every manuscript coming out of my lab. In practice, some will not be, and the idea is to then use the checklist to help educate (or remind) us why something might be a good idea, and identify reasons for not following a guideline. The aim is to focus on discussion and improvement rather than simply ticking boxes.

In anticipating reasons why a guideline might not be followed, I've come up with some possibilities:

- **We don't know how to do it.** Sure, putting all of my code online seems like a good idea, but how do I organize it? What website do I use? How do I link to it in a manuscript? The total picture can be daunting for anyone, and may be particularly intimidating for researchers not used to command-line tools or repositories. I think lack of know-how is a very real reason best practices are not followed. Of course, we can all learn how to do new things, so I don't think lack of knowledge is a particularly *good* reason not to do something. (If it were, we all should have dropped out early in our PhDs...).

- **It's too hard.** Related to not knowing how to do something is the perception that it is too difficult or too time-consuming. It's true that many things are difficult or time-consuming the first time we did them, but they typically got easier with practice and a community of colleagues to ask for assistance. Within reason, the overall "difficulty" of something does not seem to be a good reason not to do it. (It took Charles Darwin some [20 years to publish his ideas on evolution and natural selection](#).)
- **It's scary.** Although there are a lot of different ways fear can come into play, fear of looking foolish is probably near the top of the list: publishing my analyses, for example, means that someone else may discover an error I've made, or see that I'm not such an amazing programmer. Although this would be good for Science and Truth, it would be bad for my ego, as I would have a mistake pointed out (perhaps publicly). We can also be afraid of being criticized (even in the absence of any mistake), reduced productivity (because we may take longer to publish), or the opinions of other scientists who may not share the same values or priorities. All of these are very *real* reasons for not doing something, but in my estimation not particularly *good* ones. [An important caveat is that for researchers in earlier career stages, even well-intended criticism could have a damaging effect—I think about this a lot and is something I'll cover more in future posts.]
- **Ethical constraints make it impossible.** This is an easy one, in the sense that *of course* no one is advocating sharing any private or protected data in the name of openness. However, I've intentionally used the word "impossible", because quite often with some forethought of both experimental design and IRB planning, many types of data can safely be shared. So, in practice, this is not a common barrier (at least in my lab).
- **Following a guideline would be bad for science.** The final reason is somewhat tongue-in-cheek: doing something scientifically "bad" would be an extremely good reason not to follow a guideline, but (at least in our lab) I don't think this would happen very often. I have this on the list to emphasize how rarely it happens. (If I'm wrong and this reason crops up a lot, that's great too, because we'll have talked about it and I will have learned something.)
- **Other.** I'm sure there are reasons I haven't thought of, and it is useful to document these, too. Again, the point is for us as authors to engage in meaningful discussions about the best way to practice science.

THE CHECKLIST

You can [view the source code](#) or [download a PDF](#) from GitHub. A bare-bones version of the current version of the checklist appears in the table below. I would encourage you to make your own checklist, tailored to the goals and needs of your own research and lab.

Item	YesNoWhy not?
OPENNESS	
Made stimuli and materials publicly available?	
Made data publicly available?	
Made data analyses publicly available?	
Submitted a preprint?	
Published in an open access journal?	
Published original figures CC-by prior to journal?	
STATISTICS	
Preregistered the experiments?	
Documented a rule for stopping data collection (e.g., number of participants, Bayes Factor) before starting?	
Explained all conducted tests? I.e., included a statement such as "We report how we	

determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.”
Collected an informative data set?

Plotted the data informatively?

Avoided p values?

Included effect sizes?

Included confidence intervals (preferably Bayesian)?

REAL-WORLD EXAMPLES

Here are some examples of things we've done recently in the lab that would fit on the checklist (although we didn't have the checklist yet):

1. **Sharing experimental stimuli.** With Ward et al. (2016) we put our first set of experimental stimuli online, hosted on a private webserver (in part for convenience, in part due to the size of the video files). We have since posted our [first set of auditory sentence stimuli to the Open Science Framework](#), which I feel better about as it's hosted by a third party and thus more likely to persist.
2. **Sharing unthresholded fMRI results.** For Lee et al. (2016) we [posted our results images on neurovault.org](#), providing a link to this in our paper. The process was straightforward and we're planning on posting all future group level fMRI results to neurovault.
3. **Preregistration.** Although [there may be reasons to avoid preregistration](#), in general I'm convinced it's more often beneficial than harmful. The main barrier to be not doing this has been a lack of knowledge about the best way to do so. Of course, creating a text document with the date and some hypotheses may have been a good start, but I've wanted something more formal. So, my reason for not preregistering was “I don't know how to do it” (and maybe that it's too hard to learn, or that I'm scared of looking foolish). With my recent discovery of [aspredicted.org](#), I now have an easy of preregistering a study, which I plan to try out in the coming months.

MOVING FORWARD

One of the things I like about the checklist format is it rewards us for each item we are able to accomplish. Good science is never “all or nothing”—separating out various goals encourages us that doing something is better than nothing, and push our work in a positive direction.

I'll be the first to admit that what I've started is just an initial attempt. Please feel free to comment below (or by email) about topics or resources that I've overlooked. I'd also love to hear whether you use something similar (formal or informal) in your own work, and whether it's changed the way you do science.

REFERENCES

Lee Y-S, Min NE, Wingfield A, Grossman M, Peelle JE (2016) Acoustic richness modulates the neural networks supporting intelligible speech processing. *Hearing Research* 333:108–117.

[doi:10.1016/j.heares.2015.12.008](#)

Peelle JE (2016) Reviewing for open science. *The Winnower* 3:e145877.76807.

[doi:10.15200/winn.145877.76807](#)

Ward CM, Rogers CS, Van Engen KJ, Peelle JE (2016) Effects of age, acoustic challenge, and verbal working memory on recall of narrative speech. *Experimental Aging Research* 42:126–144.

[doi:10.1080/0361073X.2016.1108785](#)