What's wrong with scientific publishing and how we can fix it

Jinkinson Smith[1]

[1]Oglethorpe University

February 26, 2018

There are a lot of problems with scientific publishing that have become more obvious to scientists in recent years. The most important of these is undoubtedly the so-called "replication crisis" in many scientific disciplines, most prominently in psychology. This term refers to the fact that a surprisingly large number of studies have not been reproduced by other researchers in the years (if not decades) since they were first conducted. Though some authors argue that these failures may themselves be spurious, and that more studies are needed to determine whether the effect is a real one, (Maxwell et al., 2015) it is clear that most studies published in respected peer-reviewed journals do not stand the test of time. Ioannidis famously documented this fact in a 2005 paper,(Ioannidis, 2005) and further evidence of it has been provided by more studies in the years since then (e.g. (201, 2015),(Begley and Ellis, 2012)(Prinz et al., 2011)). Here, I propose 5 measures that I think the scientific community should adopt if they want to try to address this problem, as they undoubtedly do.

My proposed solutions are as follows:

1. End the practice of issuing press releases for primary studies.

By this I mean to suggest that research centers/universities should completely stop issuing press releases for new research studies that are **not review articles/meta-analyses**. This is because we know that most published studies reporting a finding are wrong,(Ioannidis, 2005) and that this phenomenon, at least in medicine, is due in large part to the desire to "publish or perish".(Begley and Ioannidis, 2014) In psychology, a major contributor to the so-called "replication crisis" is grant culture, which incentivizes questionable research practices and discourages direct replication attempts,(Lilienfeld, 2017) as well as publication bias.(Etz and Vandekerckhove, 2016) The media shouldn't be exacerbating this problem by breathing down scientists' necks and eagerly asking them what their latest positive study is so they can run a news story about it. Science should be done by scientists for other scientists, and should not be displayed or packaged in an entertaining way until the findings regarding the phenomenon being studied are definitive. Unfortunately, this means that the general public will have to wait rather longer before being able to repeat the findings of a single study as though they were conclusively proven. But this is undoubtedly a worthy price to pay for having the integrity of science protected from a media that cares only about getting as much publicity for its reporting as possible.

2. Make peer review an automated process.

Peer review is way too important a process to let fallible, easily error- and bias-prone humans be controlling it. Research has shown that reviewers do not detect the majority of major errors in papers submitted to journals, even at prestigious ones like *BMJ*.(Schroter et al., 2008) There needs to be an objective way to judge the quality of a manuscript so that whether it is or isn't worthy of publication can be determined by an algorithm. This might involve adapting a tool like that of Bartoli et al., which was originally designed to create fake reviews of real papers,(Bartoli et al., 2016) to create real reviews of such papers according to

established criteria for assessing study quality. These criteria include the Jadad scale(1) and Pedro scale(de, 2009) (both for clinical trials). For systematic reviews (SRs) and meta-analyses (MAs) one may compare a submitted manuscript to the PRISMA guidelines.(pri)

3. Create a registry for all planned research studies in all academic disciplines.

Such a registry would include not only all the clinical trials found at websites like clinicaltrials.gov, or the AEA's registry of economics trials at socialscienceregistry.org, but also studies in all other disciplines–not just medicine or economics. It would also include systematic reviews/meta-analyses that a researcher(s) was planning on conducting, meaning that it would also include listings on SR/MA databases like PROSPERO, DARE, and the Cochrane Library's protocols. This would be beneficial because it would encourage scientists to freely share their ideas for research, which would help advance the progress of science as a whole. I imagine that this registry would also not allow anyone to delete a submission, which would make it obvious if they had decided not to pursue one of their ideas past a certain point (eg writing up or submitting the manuscript), thereby cutting down on publication bias. This would be expected to be beneficial because it is known that publication bias in the social sciences results from researchers not writing up or submitting non-positive results.(Franco et al., 2014)

4. Share all the data.

This is similar to a recent recommendation by Shrout & Rodgers, who recommended that psychologists "... adopt open science conventions of preregistration and full disclosure and that replication efforts be based on multiple studies rather than on a single replication attempt."(Shrout and Rodgers, 2018) This touches on the point made by Maxwell et al. (2015) regarding the question of whether a contradictory result is really a "failure to replicate". You can't really be sure of whether it is or not until more studies have been done to examine whether the effect reported in the first study is real, and when these studies' methodological quality has become apparent. This does not happen nearly as quickly as many in this field would undoubtedly like.

As it happens, some researchers have already begun an effort in pursuit of the laudable goals of openness and transparency, so that all scientists may have an equal opportunity to try to reproduce someone else's findings. This effort is known as the Transparency and Openness Promotion (TOP) Guidelines, published by the Center for Open Science.(gui) Over 5,000 journals and organizations have already signed the guidelines, and I encourage more of them to do so.

5. Quality, not quantity

Psychologist Scott O. Lilienfeld (2017) recently criticized his fellow psychologists' "grant culture", which, he argues, is promoted by institutional factors throughout the discipline. These factors include, according to Lilienfeld, "incentives for engaging in questionable research practices" and "researchers promising more than they can delive". These problems are not unique to psychology, as Begley & Ioannidis (2014) noted. They result in part from powerful incentives that encourage researchers to publish new studies, not replicate old ones, and to publish positive results, not neutral or negative ones. This is often summed up in the phrase "publish or perish", and Lilienfeld argues that it exerts its negative effects through an " ...emphasis on external funding as an expectation or de facto requirement for faculty tenure and promotion ".

Moreover, the increasing number of high-impact journals publishing studies that get lots of media attention has also helped encourage scientists to want to compete for this attention. To quote Begley & Ioannidis: "...the sheer magnitude of high-profile studies creates a challenge for any investigator to remain current... In the preclinical arena, there seems to be a wide-spread conscious or unconscious belief that a rigorous research process, that follows what most would consider standard scientific methodology (blinding,

repeating experiments, inclusion of positive and negative controls, use of validated reagents, etc.), may stifle the creative, innovative act of discovery. That is clearly not the case."

What, then, is to be done to encourage researchers to produce better studies rather than more that just happen to be published on high-profile topics in high-impact journals? A universal set of guidelines regarding how to conduct a specific type of investigation should not only exist, as many such guidelines already do, but should be compulsory if someone wants to publish something in a reputable journal. Then decisions to reward scientists (e.g. with tenure or other academic promotions, or with more funding) should be made on both whether their studies are high-quality and on whether they can be replicated using similarly high-quality methodologies. As others have already noted, the current system too often rewards scientists with higher citation counts for their papers, a practice with little empirical support behind it.(Michalska-Smith and Allesina, 2017)

Concluding remarks

In short, the way science currently operates incentivizes scientists to conduct sloppy research quickly and, if they get a positive result, submit it to a respected journal, which will then want to publish it because of its positive result (assuming it's about a topic that is interesting to the public). There are few fields where a meritocracy, where the question of whether one's research stands the test of time, should be a greater factor in employment decisions than in academia. Again, this will take longer, much like waiting to publicize research until its results are definitive will also require a much longer time to wait. Why do it? Because this will improve the confidence scientists can have in their results, and reduce the amount of time (not to mention money) that they spend following red herrings originating from flawed, attention-grabbing studies.

References

Appendix: Jadad Scale for Reporting Randomized Controlled Trials. In *Evidence-based Obstetric Anesthesia*, pages 237–238. Blackwell Publishing Ltd. doi: 10.1002/9780470988343.app1. URL https://doi.org/10.1002%2F9780470988343.app1.

TOP Guidelines. URL valhttps://cos.io/our-services/top-guidelines/uehere.

PRISMA. http://www.prisma-statement.org/PRISMAStatement/Default.aspx. URL http://www.prisma-statement.org/PRISMAStatement/Default.aspx. Accessed on Fri, February 23, 2018.

Estimating the reproducibility of psychological science. *Science*, 349(6251):aac4716–aac4716, aug 2015. doi: 10.1126/science.aac4716. URL https://doi.org/10.1126%2Fscience.aac4716.

Alberto Bartoli, Andrea De Lorenzo, Eric Medvet, and Fabiano Tarlao. Your Paper has been Accepted Rejected, or Whatever: Automatic Generation of Scientific Paper Reviews. In *Lecture Notes in Computer Science*, pages 19–28. Springer International Publishing, 2016. doi: 10.1007/978-3-319-45507-5_2. URL https://doi.org/10.1007%2F978-3-319-45507-5_2.

C. G. Begley and J. P. A. Ioannidis. Reproducibility in Science: Improving the Standard for Basic and Preclinical Research. *Circulation Research*, 116(1):116–126, dec 2014. doi: 10.1161/circresaha.114.303819. URL https://doi.org/10.1161%2Fcircresaha.114.303819.

CG Begley and LM Ellis. Drug development: Raise standards for preclinical cancer research. *Nature*, 483: 531–3, Mar 2012.

Morton NA de. The PEDro scale is a valid measure of the methodological quality of clinical trials: a demographic study. *Aust J Physiother*, 55:129–33, 2009.

Alexander Etz and Joachim Vandekerckhove. A Bayesian Perspective on the Reproducibility Project: Psychology. *PLOS ONE*, 11(2):e0149794, feb 2016. doi: 10.1371/journal.pone.0149794. URL https://doi.org/10.1371%2Fjournal.pone.0149794.

A. Franco, N. Malhotra, and G. Simonovits. Publication bias in the social sciences: Unlocking the file drawer. *Science*, 345(6203):1502–1505, aug 2014. doi: 10.1126/science.1255484. URL https://doi.org/10.1126%2Fscience.1255484.

John P. A. Ioannidis. Why Most Published Research Findings Are False. *PLoS Medicine*, 2(8):e124, aug 2005. doi: 10.1371/journal.pmed.0020124. URL https://doi.org/10.1371%2Fjournal.pmed.0020124.

Scott O. Lilienfeld. Psychology's Replication Crisis and the Grant Culture: Righting the Ship. *Perspectives on Psychological Science*, 12(4):660–664, jul 2017. doi: 10.1177/1745691616687745. URL https://doi.org/10.1177%2F1745691616687745.

SE Maxwell, MY Lau, and GS Howard. Is psychology suffering from a replication crisis? What does failure to replicate really mean? *Am Psychol*, 70:487–98, Sep 2015.

Matthew J. Michalska-Smith and Stefano Allesina. And not or: Quality, quantity in scientific publishing. *PLOS ONE*, 12(6):e0178074, jun 2017. doi: 10.1371/journal.pone.0178074. URL https://doi.org/10.1371%2Fjournal.pone.0178074.

Florian Prinz, Thomas Schlange, and Khusru Asadullah. Believe it or not: how much can we rely on published data on potential drug targets? *Nature Reviews Drug Discovery*, 10(9):712–712, sep 2011. doi: 10.1038/nrd3439-c1. URL https://doi.org/10.1038%2Fnrd3439-c1.

Sara Schroter, Nick Black, Stephen Evans, Fiona Godlee, Lyda Osorio, and Richard Smith. What errors do peer reviewers detect and does training improve their ability to detect them? *Journal of the Royal Society of Medicine*, 101(10):507–514, oct 2008. doi: 10.1258/jrsm.2008.080062. URL https://doi.org/10.1258%2Fjrsm.2008.080062.

Patrick E. Shrout and Joseph L. Rodgers. Psychology Science, and Knowledge Construction: Broadening Perspectives from the Replication Crisis. *Annual Review of Psychology*, 69(1):487–510, jan 2018. doi: 10.1146/annurev-psych-122216-011845. URL https://doi.org/10.1146%2Fannurev-psych-122216-011845.