



Reproducibility in research: How a small field is pioneering a culture of sharing

RICHARD DE GRIJS^{1,2}

1. Kavli Institute for Astronomy and Astrophysics, Peking University, China
2. *The Astrophysical Journal Letters*

The ‘reproducibility crisis’ in science appears to be a widespread problem that may have its roots in the ‘publish or perish’ culture of the contemporary academy. Facilitated by a well-developed culture of data sharing, in astrophysics opportunities to reproduce or replicate published results have been part of the field’s fabric for many decades. The valuable lessons learned from this small discipline could easily be rolled out to other data-rich disciplines. This essay aims at triggering more extensive discussion of the numerous advantages of data sharing and responsible research attitudes.

READ REVIEWS

WRITE A REVIEW

CORRESPONDENCE:

grijs@pku.edu.cn

DATE RECEIVED:

May 28, 2016

DOI:

10.15200/winn.146449.94380

ARCHIVED:

May 28, 2016

KEYWORDS:

#LJAFreproducibility, open access, open data

CITATION:

Richard de Grijs,
Reproducibility in research:
How a small field is pioneering
a culture of sharing, *The
Winnower* 3:e146449.94380 ,
2016 , DOI:
[10.15200/winn.146449.94380](https://doi.org/10.15200/winn.146449.94380)

© de Grijs This article is distributed under the terms of the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and redistribution in any medium, provided that the original author and source are credited.

Psychology. Medicine. Economics. Disciplines that have all been rocked by allegations of driving a ‘reproducibility crisis.’ Awareness of a few high-profile cases—surely only part of what appears a widespread problem—was propelled into the spotlight by an eye-opening article in *The New Yorker* (Lehrer, 2010).

This perceived crisis is, however, just a symptom of the ‘publish or perish’ culture that pervades the academy. Publication bias toward statistically significant results, *p*-hacking—attempts at uncovering statistically significant signals irrespective of the nature of one’s data (Head *et al.*, 2015)—flawed significance testing, and even scientific fraud have all been suggested as the culprits. Avid readers of the *Retraction Watch* blog (<http://www.retractionwatch.com>) will be familiar with many of these problems. Meanwhile, the most prestigious journals and many scientific funding bodies tend to favor novel, flashy results. Many of those will be statistical outliers by their very nature. Psychologist Brian Nosek of the Center for Open Science concurs: “Incentives for achievement are similar across disciplines,” he says. “Publication is essential, and positive, novel, tidy results increase the likelihood of getting published everywhere” (Weir, 2015).

Reproducibility or, better yet, replication (Peng, 2015)—pursuing the same experiments independently rather than re-analysis of the same data sets—have so far not been considered of major importance across the scientific disciplines. But that attitude is changing. Psychology’s *Reproducibility Project* (Open Science Collaboration, 2015), biomedicine’s *Reproducibility Initiative* (Science Exchange, 2016), and the common *Principles and Guidelines in Reporting Preclinical Research* established by editors representing the leading biomedical journals (National Institutes of Health, 2014) suggest that change is imminent. Nosek agrees, “Innovation points out paths that are possible; replication points out



paths that are likely; progress relies on both” (Weir, 2015).

In my own discipline, astrophysics, opportunities to reproduce or replicate published results have been part of the field’s basic principles for many decades. Here I highlight some of the valuable lessons from a small field, ideas that could indeed be rolled out easily to other data-rich disciplines. I hope that this will trigger more extensive discussion of the numerous advantages of data sharing and responsible research attitudes. Attitudes and objectives that include transparency, independent verification, reproducibility, and replicability form the basis of rigorous scientific assessment. Reproducibility doesn’t make a result right; neither does failure to replicate a result make it necessarily wrong.

In astrophysics we have been operating under open-data policies for longer than my professional memory stretches. Our colleagues have been making primary as well as derived data freely available for decades—both observational results and, increasingly, the outcomes of large-scale numerical simulations. An excellent recent example was inspired by the discovery of a ‘supernova’—the explosive death of a star several times more massive than our Sun—which was serendipitously found to be ‘gravitationally lensed’ by a massive galaxy group located along our sightline. Similarly to optical lensing, this leads to the occurrence of multiple images of the supernova’s host galaxy. The main competing research teams agreed to share all relevant data to enable independent predictions of when this supernova might be seen in the other images (Treu *et al.*, 2016). In turn, this facilitated a true blind test of the models’ validity, leading to some of the most precise measurements ever achieved of a number of fundamental physical parameters.

This culture of sharing is not driven by altruism—competition for hot scientific results and high-profile publications is as fierce as anywhere in science—but determined by demographics. Astrophysicists don’t usually run their own laboratories, but most rely on shared international facilities. After all, we need access to dark skies, a rare commodity near leading institutions in major population centers. Instead, we need to travel to isolated islands or remote mountains, sites that enjoy unimpeded air flows—all to avoid image blurring caused by turbulence.

Construction, operation, and maintenance of cutting-edge astronomical observatories is expensive; our research funders understandably want to achieve the best value for their money. Ever since the Internet became a mainstream tool, most publicly funded observatories have provided open data archives. Principal investigators retain proprietary data rights for a year, sometimes longer, but once that period has passed, anyone with an Internet connection can download, analyze, and eventually publish the observational data and their own analysis. All major data archives employ well-documented software pipelines, so that one can go back to the original observations and retrace the steps leading to the final products. Current open-data policies have gone significantly beyond initial efforts to provide access on a per-facility basis. For more than a decade, data managers at many public observatories have been working on implementing common standards in archiving in the framework of the International Virtual Observatory Alliance (Quinn *et al.*, 2004), pursuing full interoperability among its member organizations.

Nevertheless, few articles only reproduce previous work because of the journals’ focus on novelty. In many fields, replicating previous studies before pursuing the novelty factor carries an implicit penalty: reproduction takes time and often significant resources, while it doesn’t lead to tenure or promotions. Not yet, at least, although current developments are promising.

Common fears and objections regarding open-data policies (e.g., Smith & Roberts, 2016) have not been borne out. Sure, our competitors also have access to the same data we are interested in, and on rare occasions we might even get scooped. Fortunately, such instances do not happen regularly, not even in hotly contested subfields. There simply is an overwhelming body of data ‘out there,’ so encroaching on someone else’s territory is not a priority. And indeed, some colleagues will use and publish publicly available data and gain promotions using data they didn’t obtain themselves. This, I believe, is laudable: after all, secondary use of data languishing in archives enhances their impact.

Not everyone agrees. In a recent editorial in the *New England Journal of Medicine* (Longo & Drazen, 2016), scientists engaging in secondary data analysis were called “research parasites”. Fortunately, this thoughtless generalization triggered a significant backlash and led to the establishment of annual awards for rigorous secondary data analysis, tongue-in-cheek referred to as ‘*The Parasites*’ (Oransky & Marcus, 2016).

But providing access to data on its own does not guarantee replicable science. Human fallibility plays a major role in the publication of incorrect results. Scientists are often poorly trained in data curation, computational version control, or the latest statistical advances (Fidler & Gordon, 2013; Peng, 2015). I am deputy editor of one of the family of journals owned by the American Astronomical Society (AAS). Unique among our competitors, we employ a dedicated statistics editor who scans most submissions for potential issues related to the statistical treatments used by their authors—publication bias, use of inappropriate statistical techniques, and p -hacking come to mind.

Now that we have entered the era of Big Data, new and evolving data policies must be considered. One of the impending major developments in astronomy is the construction of the *Large Synoptic Survey Telescope (LSST)*, which will observe the entire sky visible from its base in Chile every three days. The data deluge that is due to pour into the observatory’s computer servers will amount to many *petabytes* every night—too much for a single human to comprehend, and too much to store long-term as well. The prospects of making a movie of the night sky—which is, in essence, the *LSST*’s *raison d’être*—is exciting, but it will fast forward us into a truly new era of data sharing.

Our journals are gearing up for the task already. The AAS journals recently established a specific workflow for computational articles. Peer-reviewed computer codes and the associated data sets can now be published, and thus cited. In turn, this policy serves to encourage authors to publish the computational codes underlying their advanced data analyses, all in the interests of reproducibility. Some of our sister journals have been publishing similar material for a few years already. This approach is also adopted by a number of journals in fields like psychology, which issue open-data badges to articles that provide clear information on their data provenance.

To date, reproducibility scandals haven’t hit astrophysics as a field (but see Chang 2016; NASA 2016); instead, our highly valued culture of data sharing, driven by necessity, has enabled our science to move forward much faster and more decisively than if we had selfishly held on to our data. The positive aspects of this approach have proven to outweigh the disadvantages by far. Physics as a whole has been at the forefront of public data provision. Perhaps other fields should seize the opportunities as well. Replicability doesn’t need to depend on the goodwill of a few individual scientists, but fears of

losing control of one's data are deep-rooted.

"Reproducibility is important, hard, and improvable," says Nosek. "We can nudge the incentives driving our behavior, so that researchers are rewarded for more transparent and reproducible research." Indeed. Now tell your research administrators. And your colleagues.

REFERENCES

- Chang K. "How big are those killer asteroids? A critic says NASA doesn't know." *The New York Times*, 24 May 2016, p. D1. <http://www.nytimes.com/2016/05/24/science/asteroids-nathan-myhrvold-nasa.html>
- Fidler F., Gordon A. "Science is in a reproducibility crisis – How do we solve it?" *The Conversation*, 19 September 2013. <https://theconversation.com/science-is-in-a-reproducibility-crisis-how-do-we-resolve-it-16998> (Accessed 11 May 2016)
- Head M. L., Holman L., Lanfear R., Kahn A. T., Jennions M. D. "The Extent and Consequences of P-Hacking in Science." *PLoS Biol.* 13(3), e1002106 (2015). doi:10.1371/journal.pbio.1002106
- Lehrer J. "The Truth Wears Off: Is there something wrong with the scientific method?" *The New Yorker*, 13 December 2010, pp. 52–57. <http://www.newyorker.com/magazine/2010/12/13/the-truth-wears-off>
- Longo D. L., Drazen J. F. "Data sharing." *N. Engl. J. Med.*, **374**, 276–277 (2016). doi: 10.1056/NEJMe1516564
- NASA. "NASA response to recent paper on NEOWISE asteroid size results." (2016) <http://www.nasa.gov/feature/nasa-response-to-recent-paper-on-neowise-asteroid-size-results> (Accessed 29 May 2016)
- National Institutes of Health. "Principles and Guidelines in Reporting Preclinical Research." (2014) <https://www.nih.gov/research-training/rigor-reproducibility/principles-guidelines-reporting-preclinical-research> (Accessed 9 May 2016)
- Open Science Collaboration. "Estimating the reproducibility of psychological science." *Science*, **349** (6251) (2015). doi: 10.1126/science.aac4716
- Oransky I., Marcus A. "For science to improve, let's put the right prizes on offer." (2016) <https://www.statnews.com/2016/05/05/incentives-science/> (Accessed 11 May 2016)
- Peng R. "The reproducibility crisis in science: A statistical counterattack." *Significance*, **12**, 30–32 (2015). doi: 10.1111/j.1740-9713.2015.00827.x
- Quinn P. J., Barnes D. G., Csabai I., Cui C. Z., Genova F., Hanisch B., Kembhavi A., Kim S. C., Lawrence A., Malkov O., Ohishi M., Pasian F., Schade D., Voges W. "The International Virtual Observatory Alliance: recent technical developments and the road ahead." *Proc. SPIE*, **5493**, 137–145 (2004). doi: 10.1117/12.551247
- Science Exchange. "The Reproducibility Initiative." (2016) <http://validation.scienceexchange.com/#/reproducibility-initiative> (Accessed 9 May 2016)
- Smith R., Roberts I. "Time for sharing data to become routine: the seven excuses for not doing so are all invalid (v. 1)." *F1000Research*, **5**, 781 (2016). doi: 10.12688/f1000research.8422.1
- Treu T., Brammer G., Diego J. M., et al. "Refsdal Meets Popper: Comparing Predictions of the Re-appearance of the Multiply Imaged Supernova Behind MACSJ1149.5+2223." *Astrophys. J.*, **817**, 60 (2016). doi: 10.3847/0004-637X/817/1/60

Weir K. "A reproducibility crisis? The headlines were hard to miss: Psychology, they claimed, is in crisis." *Monitor on Psychology*, **46**(9), 39 (2015).