doi:10.5281/zenodo.3265454

Enabling scientific reproducibility in plasma research

Nicholas A. Murphy,[1, a)] B. L. Alterman,[2, 3] D. Stansby,[4] Arturo Dominguez,[5] and Dominik Stańczak[6]

[1)] *Center for Astrophysics | Harvard & Smithsonian*

[2)] *Climate & Space Science, University of Michigan*

[3)] *Applied Physics, University of Michigan*

[4)] *Physics Department, Imperial College London*

[5)] *Princeton Plasma Physics Laboratory*

[6)] *University of Warsaw*

I. HOW REPRODUCIBLE IS PLASMA RESEARCH?

Reproducibility is a cornerstone of the scientific method. Experimental results should be able to be reproduced before those results are accepted by the broader scientific community. In 2015, the Open Science Collaborative announced astonishing results from the Reproducibility Project: Psychology.[1] Researchers worldwide attempted to reproduce results from 100 studies published in three top psychology journals in 2008. **Only 39% of the published effects were able to be replicated.** Low replication rates have also been reported for fields ranging from oncology[2] to economics.[3] The frequent inability to replicate published results has been dubbed the **reproducibility crisis** of modern science.

Physicists are reported to have high confidence that physics research is reproducible.[4] However, no large-scale studies of the reproducibility of physics or astronomy research have been performed to back up or refute this claim. While we lack quantitative information on the reproducibility of plasma research, we are able to assess the extent to which plasma research satisfies scientific reproducibility conditions and best practices.

Stodden et al.[5] describe necessary conditions for research to be scientifically reproducible. Research is *reviewable* if the methodology can be independently evaluated and the results can be judged for credibility, such as through peer review and journal publications. Research is *confirmable* if the results and conclusions can be obtained independently without using the equipment or software used for the original project. Research is *replicable* when tools are available for others to duplicate results, including data sets and the software used to analyze them. Research is *auditable* when the research materials are organized and archived so that the research can be evaluated later if necessary. Research that is replicable, auditable, and publicly available is deemed *open* and *reproducible*. A built-in requirement to this condition is that the data and source code used to reach scientific conclusions be findable, accessible, interoperable, and reusable (FAIR).[6]

In order for plasma science research to be reviewable and confirmable, the experimental, computational, or theoretical methodology must be fully reported. Research publications require a detailed description of the methodology followed during a project. In principle, the description should contain all of the information necessary to recreate the results. In

[a)] Email: namurphy@cfa.harvard.edu

practice, the methodology contains most of the necessary information but may gloss over details that are not deemed important or do not fit within space constraints. This condition is not fully met for at least some fraction of plasma physics research, though the value of this fraction is hard to evaluate without performing a Reproducibility Project: Plasma Science. **Plasma science research is generally reviewable and largely confirmable.**

All of the data and files used for the analysis must be available in order for a research project to be replicable. To place some bounds on the fraction of plasma research that is replicable, we can estimate the fraction of articles that cite either openly available source code or open access data as a proxy. Of the first fifty articles of the December 2018 issue of *Physics of Plasmas*, we found one citation of source code from the Astrophysics Source Code Library and zero citations of data sets. **Plasma research is generally not replicable, auditable, or open. Plasma research therefore usually does not meet the conditions of being reproducible.**

The cultural aspects of research and academia that contribute to the reproducibility crisis[5,7] are present in the plasma physics community too. Of particular importance is the incentive structure that is currently in place. Making research reproducible is usually not highly valued by prospective employers, funding agencies, or journals. Positive results are more valued than negative results, so negative results are much less likely to be submitted or accepted for publication. Such publication bias can increase the number of false positives beyond what is expected. Because the number of peer-reviewed publications is one of the most commonly used metrics to judge scientific output, early career scientists are likely to be at a disadvantage in the job search if they invest extra time making their data sets open access or writing documentation for their code. Improving the reproducibility of research is unlikely to help with getting tenure. The culture and incentive structure across the scientific community must change in order for reproducible research to be valued.

Threats to reproducible science include cognitive biases, inadequate methodology, and incorrect application of statistics.[7] Practices such as hypothesizing after the results are known (HARKing)[8] and data dredging (also known as P-hacking)[9] can substantially increase the rate of false positives. These dangers can be mitigated through pre-registration of research protocols, improved statistical/methodological training, and independent statistical/methodological support.[7]

Open access to data from laboratory experiments and numerical simulations will improve reproducibility, enable data science studies, and increase scientific return on investment. Open data policies are extremely common in astronomy and solar physics. One need only browse through an issue of *Solar Physics* or the *Astrophysical Journal* to find dozens of articles that make use of one or more open data sets. Many of these articles would not have been written if not for open access policies. A concern that has been raised in our community is that making research products openly available will lead to people using them incorrectly. Experience from the astronomy community shows this is not a significant problem. Moreover, there are times when researchers misinterpret the data they obtained,[10,11] and having new researchers investigate data sets could yield alternate interpretations. There are challenges to making data available from plasma experiments and simulations, including that data sets for plasma physics are more heterogeneous than in the astrophysical sciences. However, these challenges are surmountable.

II. MAKING OUR RESEARCH REPRODUCIBLE

We provide recommendations to funding agencies, publishers, and institutions for enabling scientific reproducibility in a paper submitted to the Plasma 2020 decadal review.[12] These recommendations include teaching best practices for scientific reproducibility, code development, and statistics when educating the next generation of plasma scientists; investing in research intended to reproduce important previously published results; adopting editorial policies that support scientific reproducibility; favorably viewing reproducible research practices in hiring and tenure decisions; making plasma data sets openly available; and funding efforts to openly archive previously acquired plasma data sets. In a companion paper, we advocate for an open source software ecosystem for plasma research and education.[13] In this paper (which is adapted from Ref. 12), we describe possibilities for how we can make plasma research reproducible as members of the plasma physics community. We encourage authors and advocates of different initiatives to incorporate any of these aspects into their initiatives.

A. Create an online portal to access plasma data sets

Open access to plasma data is necessary for scientific reproducibility. The plasma physics community should create a portal to access different plasma data sets. This portal would be similar to the Virtual Solar Observatory and enable greater findability, accessibility, and reusability of data. The portal could access different open plasma data archives. This portal should have a well-defined application programming interface to enable the community to create search and download tools (for example, as a component of an open source software ecosystem for plasma physics).

B. Develop open metadata standards

Plasma data sets produced by different groups are often stored using different metadata conventions. Shared metadata standards will make data sets more interoperable and reusable, and simplify cross-device studies (including for data science initiatives). Example metadata standards for plasma physics and closely related fields include openPMD[14] and the SPASE metadata model.[15]

C. Educate current and future researchers on reproducibility best practices

All members of the plasma physics community should be provided opportunities to learn scientific reproducibility best practices. These practices may be worked into plasma physics curricula by either creating a new course or by coherently incorporating these practices into existing courses. New educational materials should be made openly available online so that they can be used in courses at other institutions to avoid duplication of effort. Funding agencies should support workshops or summer schools that help students and researchers learn the skills necessary to make research reproducible.

D. Establish a national committee on scientific reproducibility

A committee on scientific reproducibility would be responsible for tasks such as:

- Coordinating community education on statistical best practices and threats to reproducibility such as *P*-hacking and HARKing;
- Working with publishers to adopt policies to reduce publication bias, such as encouraging publication of negative results and replication attempts and awarding badges to articles that share data or code
- Implementing metadata standards describing experimental data, simulation results, research software, and experimental equipment;
- Creating sample workflows;
- Writing reproducibility best practices guides;
- Devising and implementing strategies for openly archiving data from experiments and simulations;
- Encouraging and coordinating replication attempts; and
- Developing a roadmap for making plasma science reproducible based on the recommendations from the recent National Academy of Sciences report on *Reproducibility and replicability in science*.[16]

E. Establishing a national committee on software sustainability

A committee on software sustainability would be responsible for tasks such as:

- Coordinating software development activities across the community;
- Making sure that critical research software is being supported;
- Identifying gaps in functionality that need to be filled;
- Identifying duplication of functionality;
- Determining which software projects are most important to sustain;
- Improving findability, accessibility, interoperability, and reusability of software;
- Working with publishers to publicize how to acknowledge/cite software in articles;
- Providing expertise and trainings to the community on software sustainability;
- Coordinating Software Carpentry workshops; and
- Educating the community about open licensing of software and data.

F. Perform a Reproducibility Project: Plasma Physics

The first section of this paper asks: how reproducible is plasma physics research? Because no large-scale studies have been performed to test the replicability of published results in plasma physics, we do not really know the answer to this question. The Reproducibility Project: Psychology was a transformative event in the psychological sciences. Psychology is now one of the leading fields in the credibility revolution of modern science.

What if we perform a Reproducibility Project: Plasma Physics too? Such an effort could entail choosing ∼25 articles that were published in a leading journal within a particular year and attempting to independently replicate the results. The effort could include attempts to replicate experimental, observational, and numerical published results. If most research

results can be replicated to within statistical uncertainties, then we can have confidence that plasma physics research is replicable. If a large fraction of the results are not able to be replicated, then we will know that research practices in plasma physics will need to change and that major decisions should not be made based on results from only one paper or one research group.

A Reproducibility Project: Plasma Physics would require significant participation from the broader plasma physics community and financial support from funding agencies. The outcome could effect a paradigm shift in plasma physics research.

III. FINAL THOUGHTS

Reproducibility is a critical yet severely undervalued component of the scientific method. Reproducibility is especially important in a field like magnetic fusion energy where major decisions on large expensive projects are based on results that may or may not be reproducible. Incorporating reproducibility best practices into our discipline will require policy, educational, and cultural changes to our established practices. Adapting these practices into our field will result in research results that are based on a more solid foundation.

REFERENCES

[1]Open Science Collaboration, "Estimating the reproducibility of psychological science," Science **349** (2015), 10.1126/science.aac4716.

[2]G. C. Begley and L. M. Ellis, "Raise standards for preclinical cancer research," Nature **483** (2012), 10.1038/483531a.

[3]G. S. Christensen and E. Miguel, "Transparency, reproducibility, and the credibility of economics research," Working Paper 22989 (National Bureau of Economic Research, 2016).

[4]M Baker, "1,500 scientists lift the lid on reproducibility," Nature **533**, 452 (2016).

[5]V. Stodden, D. H. Bailey, J. Borwein, R. J. LeVeque, W. Rider, and W. Stein, "Setting the Default to Reproducible," (2013).

[6]M. D. Wilkinson et al., "The FAIR Guiding Principles for scientific data management and stewardship," Scientific Data **3** (2016), 10.1038/sdata.2016.18.

[7]M. R. Munafò et al., "A manifesto for reproducible science," Nature Human Behaviour **1**, 0021 (2017), 10.1038/s41562-016-0021.

[8]N. L. Kerr, "HARKing: Hypothesizing After the Results are Known," Personality and Social Psychology Review **2**, 196 (1998).

[9]L. K. John, G. Loewenstein, and D. Prelec, "Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling," Psychological Science **23**, 524 (2012).

[10]R. Cowen, "Gravitational waves discovery now officially dead," Nature (2015), 10.1038/nature.2015.16830.

[11]G. Miller, "A Scientist's Nightmare: Software Problem Leads to Five Retractions," Science **314**, 1856 (2006).

[12]N. A. Murphy, B. L. Alterman, and D. Stansby, "Making plasma research reproducible," 10.5281/zenodo.2578291 (2019), Zenodo.

[13]N. A. Murphy et al., "Building an open source software ecosystem for cross-disciplinary plasma research and education," 10.5281/zenodo.2578277 (2019).

[14]A. Huebl, R. Lehe, J.-L. Vay, D. P. Grote, I. Sbalzarini, S. Kuschel, D. Sagan, F. Prez, F. Koller, and M. Bussmann, "openPMD 1.1.0: Base paths for mesh- and particle- only files and updated attributes," (2018).

[15]D. A. Roberts, J. Thieman, V. Gnot, T. King, M. Gangloff, C. Perry, C. Wiegand, D. De Zeeuw, S. F. Fung, B. Cecconi, and S. Hess, "The SPASE Data Model: A Metadata Standard for Registering, Finding, Accessing, and Using Heliophysics Data Obtained From Observations and Modeling," Space Weather **16**, 1899–1911 (2018).

[16]National Academies of Sciences, Engineering, and Medicine, *Reproducibility and Replicability in Science* (National Academies Press, 2019).