Theme Editor's Introduction to Reproducibility and Replicability in Science

Victoria Stodden^{1,2,3,4}

¹Department of Statistics, College of Liberal Arts and Sciences, University of Illinois Urbana-Champaign, Urbana, Illinois, United States of America,

²School of Informatics Sciences, University of Illinois Urbana-Champaign, Urbana, Illinois, United States of America,

³National Center for Supercomputing Applications, Office of the Vice Chancellor for Research and Innovation, University of Illinois Urbana-Champaign, Urbana, Illinois, United States of America,

⁴Coordinated Science Lab, Department of Computer Science, Grainger College of Engineering, University of Illinois Urbana-Champaign, Urbana, Illinois, United States of America

Published on: Dec 16, 2020

DOI: <https://doi.org/10.1162/99608f92.c46a02d4>

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

Since the inaugural publication of the *Philosophical Transactions of the Royal Society* in 1665 (Shapin & Shaffer, 1985), scientific research and dissemination standards have aimed to enable the independent verification of new findings. Since then, and particularly in the last several decades, scientific discovery has increasingly leveraged diverse new tools and technologies in a rapidly changing social and regulatory environment, raising corresponding questions about reproducibility and replication of results. The National Academies of Science, Engineering, and Medicine (NASEM) recently published a consensus report: [Reproducibility and Replicability in Science](#) (NASEM, 2019), a [summary](#) of which is included as a supplement to this introduction. The committee's Statement of Task reads as follows:

The National Academies of Sciences, Engineering, and Medicine will assess research and data reproducibility and replicability issues, with a focus on topics that cross disciplines.

The committee will

- 1. provide definitions of “reproducibility” and “replicability” accounting for the diversity of fields in science and engineering;*
- 2. assess what is known and, if necessary, identify areas that may need more information to ascertain the extent of the issues of replication and reproducibility in scientific and engineering research;*
- 3. consider if the lack of replicability and reproducibility impacts the overall health of science and engineering as well as the public's perception of these fields;*
- 4. review current activities to improve reproducibility and replicability;*
- 5. examine (a) factors that may affect reproducibility or replicability including incentives, roles and responsibilities within the scientific enterprise, methodology and experimental design, and intentional manipulation; (b) as well as studies of conditions or phenomena that are difficult to replicate or reproduce;*
- 6. consider a range of scientific methodologies as they explore research and data reproducibility and replicability issues; and*
- 7. draw conclusions and make recommendations for improving rigor and transparency in scientific and engineering research and will identify and highlight compelling examples of good practices.*

This special theme of *Harvard Data Science Review* presents a broad range of research and commentary from a variety of fields and stakeholders related to the report and the issues of reproducibility and replicability in research.

The twelve articles in this special theme fall into three categories. The first group comprises four articles that provide context and reflection on the NASEM report and reproducibility and replicability. The first of these is [an interview with Reproducibility and Replicability in Science Committee Chair Harvey V. Fineberg, President of the Gordon and Betty Moore Foundation and myself as committee member by HDSR Editor-in-Chief Xiao-](#)

[Li Meng](#). The interview probes the origin and goals of the committee. The second is [“Self-Correction by Design” by Marcia McNutt](#), President of the NASEM. This article calls for the longstanding attribute of self-correction to be built into the design of modern science. The third article in this section is called [“Leveraging the National Academies’ ‘Reproducibility and Replication in Science’ Report to Advance Reproducibility in Publishing” by Manish Parashar](#), Assistant Director for Strategic Computing at the Whitehouse Office of Science and Technology Policy, and Director of the Office of Advanced Cyberinfrastructure at the National Science Foundation. His article explores the implications of Report recommendations aimed toward publishers, and relates experience with the ongoing reproducibility pilot for the IEEE Transactions on Parallel and Distributed Systems. The last article is from three members of the Early Career Board of *HDSR*, Aleksandrina Goeva, Sara Stoudt, and Ana Trisovic: [“Toward Reproducible and Extensible Research: from Values to Action,”](#) which discusses the role of extensibility as a value in research dissemination.

The next set of articles began life as commissioned papers by the NASEM Reproducibility and Replicability in Science committee. Each of these five articles informed the committee’s deliberations and was subsequently reviewed and revised (some rather substantially) for publication in *HDSR*. The first article, [“Reproducibility and Replicability in Economics” by Lars Vilhuber](#), discusses the history and current state of reproducibility and replicability in the academic field of economics, including the many innovations taking place to advance both. The next article, [“Reproducibility and Replicability in Science, A Metrology Perspective” by Anne L. Plant and Robert J. Hanisch](#), proposes a focus on reporting sources of uncertainty with new results as an indicator of good science. The third article in this set, [“Perspectives on Data Reproducibility and Replicability in Paleoclimate and Climate Science” by Rosemary T. Bush, Andrea Dutton, Michael N. Evans, Rich Loft, and Gavin A. Schmidt](#), discusses the current state of reproducibility and replicability in climate and paleoclimate science, along with new and recent approaches towards improvement of reproducibility and replicability, challenges, and recommendations. The fourth article [“Science Communication in the Context of Reproducibility and Replicability: How Non-Scientists Navigate Scientific Uncertainty” by Emily Howell](#) discusses the public facing aspects of science communication, specifically how media coverage of reproducibility and replicability in science affects public perceptions of and trust in scientific results. [Xihong Lin’s article “Learning Lessons on Reproducibility and Replicability in Large Scale Genome-Wide Association Studies”](#) then motivates several recommendations grounded in an analysis of replication and reproducibility efforts by the GWAS community.

The final group of articles are research articles that extend or analyze the report recommendations or conclusions. The first article is [“Selective Inference: The Silent Killer of Replicability” by Yoav Benjamini](#). This article outlines the reproducibility and replicability discussion across stakeholders in the scientific community and argues for an increased focus on the role of selective inference in enhancing replicability. The next article [“Trust but Verify: How to Leverage Policies, Workflows, and Infrastructure to Ensure Computational Reproducibility in Publication” by Craig Willis and myself](#), evaluates publisher policy responses to reproducibility, and distills key points for policy implementation and responsiveness to report

recommendations.¹ The last article in this group [“Reproducibility and Replication of Experimental Particle Physics Results” by Thomas R. Junk and Louis Lyons](#) examines how reproducibility and replicability is approached when experiments are prohibitively capital intensive, including procedures used to ensure computational reproducibility of results and methods that maximize their reliability.

On behalf of *HDSR* and its Editor-in-Chief, I thank the NASEM and their outstanding staff who expertly shepherded this work, especially our Study Director Jenny Heimberg and Program Director Tom Arrison, which made this special theme possible. I hope readers will find this set of articles both informative and thought-provoking, because it provides context for the complexity of the topics of reproducibility and replicability and its broad span of issues from research training, experimental design, statistical methodology, data collection and archiving, and dissemination and communication to the public. The discussion in this special theme conveys a sense of optimism, with reflections on initiatives to enable reproducibility and replicability and to develop recommendations for improving sound and rigorous scientific practices.

Acknowledgments

The preview image for this editorial was provided by photographer Igor Siwanowicz.

Disclosure Statement

Victoria Stodden has no financial or non-financial disclosures to share for this editorial.

References

Shapin, S., & Schaffer, S. (1985). *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*. Princeton, NJ: Princeton University Press.

National Academies of Sciences, Engineering, and Medicine. (2019). *Reproducibility and replicability in science*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/25303>

©2020 Victoria Stodden. This editorial is licensed under a Creative Commons Attribution (CC BY 4.0) [International license](#), except where otherwise indicated with respect to particular material included in the editorial.

Footnotes

1. The Willis and Stodden article was handled entirely by any anonymous HDSR editor. We are very grateful to the editor and the reviewers for their constructive and critical comments, which allowed this article to be considered for inclusion in this special issue. ↵