

Introducing the Social Science Reproduction Platform, a resource for teaching and improving computational reproducibility

Aug. 12th, 2021

 Send to Kindle

By Aleksandar Bogdanoski (Project Manager, BITSS), Fernando Hoces de la Guardia (Project Scientist, BITSS), Edward Miguel (Faculty Director, CEGA/BITSS), and Lars Vilhuber (Executive Director, Labor Dynamics Institute, Cornell University, and Data Editor, American Economic Association). This post is also posted on the CEGA blog.

Do you teach, or participate in, an empirical course where students reproduce research papers? Would you like the results of such assignments to be credited as real, citable scientific contributions? Learn how the Social Science Reproduction Platform can be of help!

Computational reproducibility, or the ability to reproduce results, tables, and other figures using the available data, code, and materials, through a process of reproduction, is necessary for instilling trust in science. Sharing data and code—the building blocks for reproducibility—allows researchers to build on top of each other’s work and illuminates tacit knowledge linked to the scientific process, thereby making science more inclusive.

Despite these benefits, reproducibility in the social sciences is surprisingly low. For example, in economics (where we are most familiar with existing evidence) reproduction audits of samples of papers in top journals suggest that less than half of articles published before 2019 were fully computationally reproducible. At

the same time, incentives for large-scale reproductions audits are limited because journals rarely publish such work. Moreover, reproducing others' work can lead to disagreements with the authors whose work is analyzed—especially in “failed” attempts. Addressing these challenges requires innovative approaches in teaching reproducibility and in how we conduct, reward, and communicate the outcomes of reproductions.



Web banner for the SSRP depicting a branching tree diagram, which is meant to illustrate the “garden of forking paths” in analyzing data and the importance of transparency in this process.

That is why we are excited to announce the official launch of the **Social Science Reproduction Platform (SSRP)**, a new resource to facilitate reproductions of **published social science research**. Developed by CEGA’s Berkeley Initiative for Transparency in the Social Sciences (BITSS) in collaboration with Dr. Lars Vilhuber, the SSRP streamlines the process of assessing and improving the computational reproducibility of published research. It can also be used for facilitating reproductions as class assignments in applied social science courses, allowing students to learn about fundamental concepts, research methods, and tools for reproducible research.

How the SSRP works

The SSRP allows its users to:

The platform leads reproducers through a standardized five-stage reproduction process, from selecting research articles to identifying their claims (*Scoping*); evaluating their current reproducibility (*Assessment*); implementing changes (e.g., debugging code, securing data files) to enhance their reproducibility (*Improvements*); and testing the validity of the results using alternative conditions of analysis (*Robustness*). We provide detailed guidance and resources for each stage in the **Guide for Accelerating Computational Reproducibility (ACRe Guide)**, which is meant to be used in combination with the SSRP.



The stages of reproductions conducted based on the SSRP approach.

The SSRP is *constructive* in its approach to reproducibility. Rather than designating entire research articles as (ir)reproducible, SSRP reproductions focus on *specific research claims* and their associated display items (figures and tables) and identify concrete steps to improve their reproducibility. Reproducibility assessments are based on a 10-point scale, which accounts for different levels of availability of data and materials, and whether the available materials faithfully reproduce the item of interest. For example, if one reproducer identifies a result to be reproducible at Level 4 because of issues in running the analysis code, then they (or future reproducers) can debug the code and upgrade its reproducibility to Level 5 (and get credit for it!).

The reproduction process results in *reproduction reports*, which are transparent and publicly document their analyses to facilitate collaboration, discussion, and reuse. *These reproduction reports can be easily shared* (using view-only anonymous links generated by the reproducer), *cited, and included as scientific contributions in research portfolios*. In instances where reproducers are uncomfortable with

sharing their identities publicly (e.g., early career researchers fearing retaliation after finding evidence that weakens the original claims of a paper), we allow them to remain anonymous or keep their reproductions private for a period of time.

Get inspired with some examples here!

- *Provide and receive constructive feedback from peers and original authors.*

Constructive communication is crucial to fostering productive collaborations, as adversarial exchanges damage professional relationships and are counterproductive, especially in collaborative enterprises like scientific research. To facilitate constructive communication between reproducers and original authors, the ACRE Guide includes recommendations and template language for common interactions, such as requesting missing data files or sharing the results of reproductions. We hope to actualize the golden rule of reproductions: *treat others and their work as you would like others to treat you and your work.*

Beyond interacting directly with original authors, SSRP users can provide and receive feedback from peers in the SSRP Forum, a moderated online discussion board for all things related to the SSRP. For example, SSRP users can visit the Forum to get help refining code files as part of the *Improvement* reproduction stage or get feedback on their reproductions before submitting. We hope the Forum will facilitate a community of diverse reproducibility-minded users interested in sharing ideas and learning from each other. **Check out the code of conduct and join the discussion!**

- *Access and contribute to the creation of meaningful metrics of reproducibility of social science research.*

Beyond learning about the reproducibility of individual papers, the SSRP's standardized approach allows the community to learn about the current reproducibility of *bodies of literature* and their evolution. A forthcoming *Reproducibility Dashboard* will summarize and display the data recorded by

reproducers. In the spirit of openness, we will release (anonymized) data with all reproductions to allow the research community to further investigate the reproducibility of different fields.

“By crowdsourcing and sharing reproductions that are conducted in classrooms worldwide but may otherwise go unpublished, we hope that the SSRP will help strengthen norms for reproducible research.”

Who the SSRP is for

The SSRP is free and open to all social science researchers interested in advancing the reproducibility of research. It may be especially useful for instructors and their students involved in reproducing published research articles as part of class assignments in applied social science courses. We have already tested the platform in several courses with overwhelmingly positive feedback and have developed guidance for instructors with tips and resources for teaching and grading.

Beyond the classroom, the SSRP can be used by researchers interested in auditing the reproducibility of their own or other’s work. Before submitting to a journal, researchers can go through the *Scoping* and *Assessment* stages to identify (and remedy) potential challenges in reproducing their work. Similarly, in studies that involve data that cannot be publicly shared, researchers can use reproduction reports generated by independent reproducers to certify the reproducibility of their work.

Finally, with leading social science journals like the AEA journals and the American Journal of Political Science requiring pre-publication verification of computational reproducibility, we are exploring the SSRP’s role in the journal review process.

What’s next for the SSRP

Though the SSRP is now open and ready for public use, we will continue developing new and exciting features based on input from the community. Several features that we will release soon include Digital Object Identifiers (DOI) for all completed reproductions; improved workflows for instructors that include the ability to provide feedback directly on the platform; and a tool for mapping reproduction packages (collections of data, code, and instructions to reproduce research articles). **Want to get involved?** If you're an instructor, check out our guidance and use the SSRP in your class. **Have thoughts about how to improve the SSRP?** Share feedback or request a feature, or even better, contribute directly! We always welcome (and credit) contributions to the ACRé Guide (learn more here) and the SSRP itself, so get in touch!

<https://www.bitss.org/introducing-the-social-science-reproduction-platform-a-resource-for-teaching-and-improving-computational-reproducibility/>