AMSTATNEWS
The Membership Magazine of the American Statistical Association

Monthly Membership Magazine of the American Statistical Association

JASA Partners with Authors to Enhance Reproducibility

3 June 2024 No Comment

Article Authors
Julia Wrobel, *Emory University*
Emily C. Hector, *North Carolina State University*
Lorin Crawford, *Microsoft Research and Brown University*
Lucy D'Agostino McGowan, *Wake Forest University*
Natalia da Silva, *Universidad de la Republica*
Jeff Goldsmith, *Columbia University*
Stephanie Hicks, *Johns Hopkins University*
Michael Kane, *Yale University*
Youjin Lee, *Brown University*
Vinicius Mayrink, *Federal University of Minas Gerais*
Christopher J. Paciorek, *University of California at Berkeley*
Therri Usher, *US Food and Drug Administration*
Julian Wolfson, *University of Minnesota*

In 2016, the *Journal of the American Statistical Association* Applications and Case Studies introduced a reproducibility initiative to address the lack of standardized practices for reproducibility in scientific research. This initiative established minimum criteria for the inclusion of code, data, and workflow for *JASA* Applications and Case Studies papers and piloted a new editorial role—associate editor of reproducibility—to implement these standards. This initiative has since expanded to all original research manuscripts published in *JASA*.

The team of associate editors of reproducibility has also grown, the process has become more standardized, and the role of an associate editor of reproducibility in guiding reproducibility in statistical research has been refined. Additionally, *JASA* implemented a reproducibility award to recognize papers published in the journal with outstanding reproducibility materials.

The *JASA* initiative is part of a larger movement in the scientific community to address what is widely considered a 'replicability crisis' in science. However, there are many terms used interchangeably—including reproducibility, replicability, reliability, robustness, and generalizability—with potentially different definitions, depending on the field of research. Broadly, these terms address either *replicability/ generalizability*, which relates to whether research findings can "be shown in other data sets or populations," or *reproducibility*, which is defined as the ability of a researcher to "duplicate the results of a prior study using the same materials as were used by the original investigator."

These definitions derive from the experimental sciences. How they apply to the statistical methods–focused research featured in *JASA* is less clear but no less critical. Reproducibility in statistics supports the validity of the research presented in a paper and underpins the role of scientific research in advancing knowledge. This role is shared by three key stakeholders—journal, academic reader, and lay public—that benefit from and contribute to high reproducibility standards.

First, the reproducibility process fosters a culture of transparency and accountability that is critical to nurturing the trust placed by the lay public in scientific research. It is an essential safeguard against the dissemination of misinformation that can damage this trust and jeopardize the positive societal impact of future scientific discoveries.

In addition, the academic reader will frequently use published research to motivate future research. Reproducible research creates trustworthy cumulative knowledge on which researchers can confidently build to further advance their field.

Finally, the journal gains credibility from a system that holds researchers accountable for their work—credibility that can be leveraged to publish controversial and impactful research with the potential to change a field of study. As the mediator between researchers and lay public, the journal has substantial control of, and therefore substantial responsibility for, the dissemination of research results. *JASA* takes an active mediating role in upholding high reproducibility standards.

Here, reproducibility in statistical research is referred to as "methods reproducibility," which involves providing materials with the following two key elements:

1. They allow for the numbers in the paper to be directly reproduced
2. They are user-friendly enough for future readers to easily use and build on the proposed technique

This second element is the most critical of methods reproducibility and what associate editors of reproducibility at *JASA* strive to uphold as they leave the assessment of paper quality and mathematical accuracy to the traditional review process.

Reproducibility Effort

To clarify their mission, associate editors of reproducibility introduce the framework of "partners in" versus "officers of" reproducibility. Reproducibility officers would police the materials published by the author to ensure all manuscript numbers and tables are precisely recapitulated. *JASA* associate editors of reproducibility view their role as reproducibility partners, who work with and guide authors in providing resources that enable readers to test and build on their work.

To this end, and to ensure a baseline minimum set of standardized materials are provided by the authors of each paper, the associate editors of reproducibility developed a reproducibility assessment process centered on the author contributions checklist form, which is the central document of the reproducibility effort. The purpose of the form is to "document the artifacts associated with a manuscript (i.e., the data and code supporting the computational findings) and describe how to reproduce the findings." Its target audience is future readers of an accepted manuscript, and it is intended to be phrased in the present tense. The workflow of a 'prototypical' reproducibility reviewer is the following:

1. Read through the authors' completed author contributions checklist form for completeness and clarity.
2. Open any README documents provided and evaluate if they clarify the workflow of the attached code. READMEs are highly encouraged.
3. Look through the directory structure of the provided materials to understand their organization.
4. Open and read through attached code files to ensure the code is sufficiently documented for the use and understanding of future readers.
5. Depending on judgment of the associate editor of reproducibility, run examples or a few key pieces of code.
6. Write a review that comments on any issues found in the steps above with the intention of helping authors strengthen materials for future readers.

Author Dos and Don'ts

The reproducibility review strives to be specific and provide concrete action items to strengthen the reproducibility of the research. By following these guidelines, authors can avoid the step of having to undergo multiple rounds of reproducibility revisions.

	Do	Don't
ACC Form	✓Fully complete the author contributions checklist form	✗ Leave template comments in the form
	✓Thoroughly explain materials included	✗ Provide one-word answers in sections
Code	✓Add comments explaining the purpose of functions	✗ Provide uncommented code blocks
	✓Explain inputs and outputs user-facing functions	
Data	✓Include detailed instructions about data access	✗ Forget to include a data dictionary
	✓When feasible, provide the data in the reproducibility materials or a public data repository	✗ Forget code for processing data
	✓When data is not shared, submit a realistic facsimile	
Workflow	✓Choose an easy-to-navigate directory structure	✗ Forget a comprehensive README file
	✓Provide a README file	✗ Use uninformative file names
	✓Include the file sequence to follow in the README file	

Table 1. Dos and Don'ts Based on Common Reasons Reproducibility Revisions Are Requested

Frequently Asked Questions

Q. ***Do the associate editors of reproducibility run all the code and check that all figures and tables are numerically reproducible? If you don't run all the code, how are you even assessing reproducibility?***
A. As a rule, associate editors of reproducibility do not run all code provided by the authors. Their role is to ensure the author contributions checklist form and code are sufficiently documented such that future readers of the paper can implement and build on results. Further, they frequently review papers with large data sets and computationally demanding methods that cannot be easily reproduced numerically but still have a rightful place in *JASA*; they do not want to discourage authors from submitting such manuscripts or hold more computationally tractable methods to a higher bar. As partners in reproducibility with the authors, the associate editors of reproducibility trust the authors have put in a good faith effort to ensure the materials they provide will produce results consistent with their published work.

Q. ***Making my code reproducible will take substantial effort and time. Do I really have to provide reproducibility materials?***
A. For reasons previously described, reproducibility plays a crucial role in published research. While time-consuming, preparing reproducibility materials can be greatly facilitated by doing so throughout the research process rather than at the time of submission. Authors should be encouraged to carefully prepare reproducibility materials by the prospect of publication in top-tier journals such as *JASA* and increased likelihood others will cite their work if the method is easy to reproduce.

Q. ***Will my paper be rejected based on the reproducibility review?***
A. No, papers at *JASA* will not be rejected based on the reproducibility materials. However, if the paper has been accepted but the authors do not comply with the associate editors of reproducibility requests, the paper will be returned to the authors for additional revisions.

Q. ***Do I fill out the author contributions checklist form when I initially submit my paper to* JASA?**
A. No, the reproducibility review process starts after a paper has already gone through one round of review and been flagged for major or minor revisions. When revising and resubmitting their paper in this second round, authors must provide reproducibility materials.

Q. ***If I have a GitHub repository with my code/software package, do I still need to fill out the author contributions checklist form?***
A. Yes. GitHub repositories and software packages are useful supplements to a paper. However, it is still important to fill out the author contributions checklist form to ensure standard materials are included and to explain the workflow. In addition, while software packages can help others use the method, associate editors of reproducibility ask that authors also contribute code from their manuscript typically not included in a software package. This includes data processing code, code for running simulations, and code for reproducing tables and figures.

Q. ***If I provide a link to a GitHub repository, doesn't that unblind the process?***
A. Blinding is intended to keep bias out of the acceptance and rejection decision. As associate editors of reproducibility do not make decisions to reject or accept papers, an unblinded GitHub repository does not affect the reproducibility review process. Authors may alternatively choose to anonymize their GitHub repository.

Q. ***How can I strive to make my materials good enough to be considered for the reproducibility award?***
A. Authors should place themselves in a reader's shoes by imagining what they would need when attempting to reproduce a paper's results or use the paper's methods.

Q. ***I have a complicated, large data set with code that is too computationally intensive to run on a standard laptop. Can I still be considered for the reproducibility award?***
A. Yes. All papers accepted by *JASA* Theory and Methods or Applications and Case Studies in the previous calendar year will be considered for the reproducibility award.

Final Words

While the authors put forth *JASA*'s current best efforts at guiding this type of reproducibility, they acknowledge methods reproducibility is not a static target and this process will likely change and improve over time.

They believe a wider discussion about methods reproducibility—including adoption of common standards—by editors across statistical journals would benefit the field and hope describing their process can facilitate discussion. In addition, they welcome the adoption of the author contributions checklist form as a template or starting point for other statistics and data science journals initiating their own reproducibility review process.

Visit the *JASA* website for the complete version of this article and author contact information.

(No Ratings Yet)

Leave your response!

Add your comment below, or trackback from your own site. You can also subscribe to these comments via RSS.

Be nice. Keep it clean. Stay on topic. No spam.

Name

Mail (will not be published)

Website (optional)

You can use these tags:
<a href="" title=""> <abbr title=""> <acronym title=""> <b> <blockquote cite=""> <cite> <code> <del datetime=""> <i> <q cite=""> <s> <strike>

This is a Gravatar-enabled weblog. To get your own globally-recognized-avatar, please register at Gravatar.

Submit Comment

Welcome!

Amstat News is the monthly membership magazine of the American Statistical Association, bringing you news and notices of the ASA, its chapters, its sections, and its members. Other departments in the magazine include announcements and news of upcoming meetings, continuing education courses, and statistics awards.

ASA HOME

- The American Statistical Association
- JobWeb
- Calendar of Events
- Awards and Recognition

Departments

Select Category

Archives

Select Month

ADVERTISERS

PROFESSIONAL OPPORTUNITIES
FDA
US Census Bureau

QUOTABLE

" Let's tell our stories, so others will know the value of calling on professional statisticians and data scientists to drive discovery and inform decisions."

— Bonnie Ghosh-Dastidar, ASA President

Editorial Staff

Contact us

Amstat News
American Statistical Association
732 North Washington Street
Alexandria, VA 22314-1904
(703) 684-1221
www.amstat.org

Address Changes

Amstat News Advertising