The Reproducibility Crisis and Academic Libraries

In recent years, evidence has emerged from disciplines ranging from biology to economics that many scientific studies are not reproducible. This evidence has led to declarations in both the scientific and lay press that science is experiencing a "reproducibility crisis" and that this crisis has significant impacts on both science and society, including misdirected effort, funding, and policy implemented on the basis of irreproducible research. In many cases, academic libraries are the natural organizations to lead efforts to implement recommendations from journals, funders, and societies to improve research reproducibility. In this editorial, we introduce the reproducibility crisis, define reproducibility and replicability, and then discusses how academic libraries can lead institutional support for reproducible research.

Background

Concerns about the reproducibility of research have become an issue in a number disciplines, including psychology,[1] biology,[2] biomedicine,[3] neuroscience,[4] drug development,[5] chemistry,[6] climate science,[7] economics,[8] and education,[9] among others. One recent study estimated the cost of funding irreproducible research at $28 billion a year in the United States alone.[10]

A number of landmark studies have garnered significant attention in both the scientific and popular press. The Open Science Collaboration's Reproducibility Project in psychology is probably the most well-known of these studies. Beginning in 2012, the project sought to estimate the reproducibility of psychological science by having 270 authors and 86 other volunteers across 41 institutions replicate a sample of 100 experiments.[11] In 2015, they reported the results, using p-values, effect sizes, and subjective assessments to determine if experiments had been replicated. They found that only 36 percent of replicated experiments had significant results, 47 percent of original effect sizes were in the 95 percent confidence interval of the replication effect size, and the researchers who conducted the replications subjectively rated only 39 percent of results as having successfully replicated the original results.[12] This project was covered prominently in the popular press and received considerable attention.[13]

Also significant was a 2005 article by John Ioannidis provocatively titled "Why Most Published Research Findings Are False." Ioannidis argued that "most research findings are false for most research designs and for most fields" due to a combination of biases in design, analysis, and reporting; testing by multiple independent teams leading to false-positive findings being published; and low-powered research designs. Admitting that there was no way to reach 100 percent certainty, Ioannidis called for higher-powered evidence, fixing publication bias, and addressing issues with other forms of bias.[14]

While the previously listed studies have focused on the reproducibility of disciplines, some individual studies that have failed to replicate have also received substantial attention. Examples include a retracted Wakefield et al. study on the MMR vaccine and autism and a LaCour and Green study on political canvassing and gay marriage that included fabricated data.[15] Another example is the power pose research of Carney, Cuddy, and Yap, which has received considerable attention partly due to a widely shared Ted Talk,[16] but, as noted in the *New York Times* and other venues, significant parts of the research have not been successfully replicated.[17]

In 2016, *Nature* conducted a survey of 1,576 researchers, which found that 52 percent believed there was a "replication crisis" in science, although less than 31 percent of respondents believed that failures to replicate meant that studies were necessarily wrong.[18] More than 70 percent of respondents reported having failed to reproduce another scientist's experiments, and more than 60 percent of researchers reported that pressure to publish and selective reporting were major factors that caused problems with reproducibility.[19] The survey also showed evidence of emerging responses to the crisis, with more than 30 percent reporting that they had taken steps to improve reproducibility in their research in the past five years.[20]

Definitions

The terms *reproducibility* and *replicability* are frequently misused, and their ambiguous and inconsistent use exists across different disciplines and contexts. Goodman, Fanelli, and Ioannidis have written about how reproducibility and replicability, along with *robustness* and *generalizability*, are not standardized and that "this diverse nomenclature has led to confusion, both conceptual and operational, about what kind of confirmation is needed to trust a given scientific result."[21] Patil, Peng, and Leek have also written about the general ambiguity of the terms, as well as how "the same words are used for different concepts by different people in different fields."[22]

One succinct definition was proposed by National Science Foundation's (NSF) Social, Behavioral and Economic (SBE) Division Subcommittee on Replicable Science. In a report on how to promote robust research practices, they define reproducibility as "the ability of a researcher to duplicate the results of a prior study using *the same materials and procedures* [emphasis added] as were used by the original investigator."[23] According to this definition, reproducibility uses the same methods *and data* to confirm the results of an experiment. The SBE defines replicability as "the ability of a researcher to duplicate the results of a prior study if the *same procedures are followed but new data are collected* [emphasis added]."[24] Thus, replicability goes a step further, gathering new data to confirm an earlier finding.

Leek and Jager also reinforce this definition, noting that "[a] study is reproducible if all of the code and data used to generate the numbers and figures in the paper are available and exactly produce the published results."[25] Broman et al. point out that this means reproducibility "is the only thing that can be effectively guaranteed in a published study. Whether any claimed findings are indeed true or false can only be confirmed via additional studies, but reproducibility can be confirmed immediately."[26]

Generally, the term "reproducibility" has stood in for both the concepts of reproducibility and replicability. Even the Open Science Collaboration's Reproducibility Project, perhaps the best known attempt at *replication*, used the term *reproducibility* in the title. This article will focus primarily on reproducibility as defined above, with the understanding that measures that improve reproducibility are a prerequisite for improving replicability.

Guidelines

To address the reproducibility crisis, many stakeholders, including funders, journals, scientific societies, institutions, and individual researchers, have developed reproducibility guidelines and recommendations. A review of the recommendations in these guidelines demonstrates that academic libraries have a considerable role in making research more reproducible.

One prominent example is the Transparency and Openness Promotion (TOP) Guidelines, a set of eight standards for academic journals created in 2014 by the Center for Open Science (COS), disciplinary leaders, journal editors, funding agency representa-

tives, and disciplinary experts.[27] The Center for Open Science is a nonprofit started in 2013 with a goal "to increase openness, integrity, and reproducibility of research."[28] The guidelines are meant to be flexible enough to be applied across disciplines, and each standard can be implemented at one of three levels: 1) disclosure, 2) mandating, or 3) verification. For example, for the "data transparency" standard, the first level would require that authors state whether the underlying data are available, the second level would mandate that data be shared, and the third level would include the journal verifying that the data have been made available. This modular framework recognizes that there are different disciplinary norms for transparency and abilities to implement change across disciplines. As of October 2017, more than 5,000 journals and societies have signed on to the TOP Guidelines.[29] In September 2017, Elsevier announced new data guidelines across 1,800 journals that align with the TOP Guidelines.[30]

The TOP Guidelines contain many recommendations that academic librarians who support research, scholarly communication, and research data management will find familiar:[31]

- **Citation:** Proper citing of data, code, and materials and the recognition of these products as legitimate intellectual contributions to science. **(Standard 1)**
- **Data Transparency, Analytic Methods (code) Transparency, Research Materials Transparency:** These three distinct standards relate to the degree to which data, code, and research materials are made available to other researchers to enable reproducibility and replication. **(Standard 2, 3, 4)**
- **Design and Analysis Transparency:** This standard encourages authors to follow explicit guidelines for disclosing key aspects of research design and analysis. For example, the PRISMA guidelines outline explicit standards for reporting systematic review research, and the ARRIVE Guidelines outline similar standards for reporting animal research. **(Standard 5)**
- **Preregistration of Studies, Preregistration of Analysis Plans:** Preregistration of studies involves publicly declaring the research you are conducting in advance, thereby increasing discovery of research that was not published (thus addressing publication bias). Preregistration of analysis plans goes a step further by including details about planned analysis, preventing problems like p-hacking as well as certifying the distinction between confirmatory and exploratory research. **(Standard 6, 7)**
- **Replication:** This standard relates to the journal's willingness to publish direct replications of studies it previously published. **(Standard 8)**

The American Statistical Association (ASA) has developed guidelines for funding agencies, aimed at improving reproducibility.[32] While these guidelines address funders, they contain many measures familiar from the TOP Guidelines. The ASA's recommendations include, among other things, increasing support for the methodological training, with particular emphasis on data management skills and adding code management plans to existing requirements of Data Management Plans (DMPs).

Also interesting for librarians are the "Principles" and "Observations" preamble to the ASA's recommendations. Specifically, the principles recognize the importance of using open data whenever possible, publishing data and code in open repositories, making data and code citable and recognized as real research contributions, and using computational practices during data analysis and processing.[33]

Major research funders have also started to address reproducibility. In 2014, the National Institute of Health (NIH) hosted a joint workshop with the Nature Publishing Group, the journal *Science*, and journal editors representing the 30 journals where NIH-funded research is most frequently published.[34] These guidelines call for improved rigor in statistical analysis, reporting transparency, data and material sharing,

consideration of refutations, considering the establishment of best practice guidelines for certain images and descriptions of biological materials.

There are a number of other guidelines by funders, professional scientific societies, individual scientists, institutions, and conferences, and we recommend that librarians find guidelines associated with the disciplines they support. For example, the Federation of American Societies for Experimental Biology has a broad set of recommendations regarding research using animal models or antibodies, intended for research organizations, individual investigators, and publishers.[35] The Society for Neuroscience's "Research Practices for Scientific Rigor: A Resource for Discussion, Training, and Practice" has discussion points and recommendations covering methodological issues, data analysis, and transparency.[36] These guidelines are invaluable for liaisons and functional specialists working within these areas. A future publication by the authors will address a comprehensive review of reproducibility guidelines.

Recommendations to improve the reproducibility of research can be conceptualized as falling into two realms: those that belong to the practice of science itself and those that belong to the packaging and reporting of science. For example, recommendations relating to study design and statistics clearly belong to the practice of science; recommendations related to these areas will differ depending on the discipline, and librarians will likely have a limited role to play. On the other hand, recommendations related to the packaging of science such as preregistration of studies and analysis plans or improvements to transparency such as data sharing have clear implications for the services and expertise that academic librarians provide. Indeed, in most cases, services and expertise related to these issues are already being offered by academic libraries; extending them to support reproducibility will require only that we effectively frame them as supporting reproducible research. In a future publication, the authors will set out an explicit model of services and expertise related to reproducibility.

Discussion

The systems needed to promote reproducible research must come from institutions—scientists, funders and journals cannot build them on their own. These kinds of changes will require additional money, infrastructure, personnel and paperwork. The load on institutions and investigators will be real, but so is the burden of irreproducible research.[37]

—C. Glenn Begley, Alastair M Buchan, and Ulrich Dirnagl

Research institutions have a responsibility to work with researchers, funders, journals, and professional societies to address issues with reproducibility. The academic library is one of the few organizations within an academic institution that has the expertise and infrastructure to broadly support reproducible research. Academic librarians have extensive experience with finding and evaluating scientific literature, scholarly communication, research metrics, and data management and sharing. We support computational and data-intensive research such as GIS, bibliometrics, data mining, and digital humanities. Most academic libraries have subject specialists assigned to every discipline, whose job it is to build and maintain relationships and support students, faculty, and researchers and understand issues within the discipline. Academic libraries are thus well placed to lead support for many aspects of reproducible research.

The role that academic librarians can play in supporting reproducibility is beginning to be recognized. Stodden et al. highlighted the role libraries could play in "supporting a culture change toward reproducible research," including assistance with data management plans and archiving.[38] In the same article, Stodden et al. wrote that "co-

ordination between departments and the institute's library system could help provide the support and resources necessary to manage and maintain digital scholarly output, including datasets and code."[39] In a report to the Director of NIH about the future of the National Library of Medicine (NLM), the advisory committee recommended that NLM "lead efforts to support and catalyze open science, data sharing, and research reproducibility, striving to promote the concept that biomedical information and its transparent analysis are public goods."[40] The *Journal of Visualized Experiments* (JoVE) has recently started giving librarians travel awards and highlighting the work they are doing to support reproducibility.[41]

Some academic libraries have begun hiring librarians specifically to support reproducibility. New York University's Division of Librarians and the Center for Data Sciences hired a dual appointment Librarian for Research Data Management and Reproducibility.[42] The University of Utah's Spencer S. Eccles Health Sciences Library (EHSL) has posted a job opening for a faculty librarian with reproducibility job components including outreach and education. Eligible candidates are required to demonstrate knowledge of "research integrity, research reproducibility, and open science" while also developing and teaching workshops on various tools like Open Science Framework, and specialized software like R.[43] Other institutions have started adding reproducibility to job descriptions along with research services and data management.

Libraries have also begun hosting events and partnering with others to address reproducibility. NYU, EHSL, and the Medical Library Association have hosted reproducibility symposiums.[44] The NYU Reproducibility Symposium included faculty, doctoral candidates, and others with the largest disciplinary representation from psychology and data sciences.[45] The University of Utah EHSL hosted a two-day conference in 2016 that was nationally and internationally attended and positioned the libraries to lead change at that institution.[46]

The University of Minnesota has been building capacity to support reproducibility. In 2016, the University of Minnesota Libraries cohosted a lecture by Brian Nosek, cofounder of the Center for Open Science. Following this talk, we cohosted a Reproducibility in Research event that featured lightning talks on various topics related to transparent, open, and reproducible research, including numerous speakers who highlighted library expertise and services.[47] This internal event served as an opportunity to showcase how library services connect to reproducibility and connect our staff with staff from other research support units. Following this event, we partnered with other research support units on campus to develop a web portal (https://www.lib.umn.edu/researchsupport/reproducibility) with the goal of bringing together campus stakeholders to help improve research practices.[48] These activities have led to numerous invitations to speak to research groups about reproducibility.

To support reproducibility, librarians will have to develop expertise in how the disciplines they support define reproducibility, the disciplinary norms regarding transparency, and the methods used in these disciplines. Librarians will also need to consider gaining the expertise needed to support computational research and workflow technology (such as OSF, R, Python). Finally, academic libraries will need to reframe existing services and expertise as impacting reproducibility and strongly advocate for our role in addressing these issues.

What we can do individually and collectively to support reproducibility will depend on many factors, including professional development. Future iterations of reproducibility symposiums such as those offered by MLA 2017 Annual Conference and at the University of Utah may be an option for professional development. Librarians could also attend disciplinary conferences and symposiums about reproducibility. For librarians in the social sciences, Research Transparency and Reproducibility Trainings are

hosted by the Berkeley Initiative for Transparency in the Social Sciences.[49] One iteration in June 2017 at University of California, Berkeley included academic librarians in attendance.

We believe that academic institutions have a clear responsibility to address issues with research reproducibility and that academic libraries are positioned to be natural leaders on these issues. Many of the measures recommended to improve reproducibility represent core areas of academic librarianship, including data management, scholarly communication, and support for data and computationally intensive research. By increasing our knowledge of disciplinary, journal, funder, and society perspectives on reproducibility, and reframing existing librarian expertise and services, academic librarians will be well positioned to support reproducible research.

Acknowledgements

Franklin Sayre
Pharmacy Librarian at University of Minnesota

Amy Riegelman
Social Sciences Librarian at the University of Minnesota

Notes

1. Open Science Collaboration, "Estimating the Reproducibility of Psychological Science," *Science* 349 (2015).

2. Keith A. Baggerly and Kevin R. Coombes, "Deriving Chemosensitivity from Cell Lines: Forensic Bioinformatics and Reproducible Research in High-Throughput Biology," *Annals of Applied Statistics* 3 (2009).

3. John P.A. Ioannidis, "Why Most Published Research Findings Are False," *PLoS Medicine* 2 (2005); Iain Chalmers et al., "How to Increase Value and Reduce Waste When Research Priorities Are Set," *Lancet* 383 (2014).

4. Rick O. Gilmore et al., "Progress toward Openness, Transparency, and Reproducibility in Cognitive Neuroscience," *Annals of the New York Academy of Sciences* 1396 (2017).

5. C. Glenn Begley and Lee M. Ellis, "Drug Development: Raise Standards for Preclinical Cancer Research," *Nature* 483 (2012).

6. Bruce C. Gibb, "Reproducibility," *Nature Chemistry* 6 (2014).

7. Rasmus E. Benestad et al., "Learning from Mistakes in Climate Research," *Theoretical and Applied Climatology* 126 (2016).

8. Thomas Herndon, Michael Ash, and Robert Pollin, "Does High Public Debt Consistently Stifle Economic Growth? A Critique of Reinhart and Rogoff," *Political Economy Research Institute Working Paper Series* (2013).

9. Matthew C. Makel and Jonathan A. Plucker, "Facts Are More Important Than Novelty: Replication in the Education Sciences," *Educational Researcher* 43 (2014).

10. Leonard P. Freedman, Iain M. Cockburn, and Timothy S. Simcoe, "The Economics of Reproducibility in Preclinical Research," *PLoS Biology* 13 (2015).

11. Open Science Collaboration, "An Open, Large-Scale, Collaborative Effort to Estimate the Reproducibility of Psychological Science," *Perspectives on Psychological Science* 7 (2012).

12. Open Science Collaboration, "Estimating the Reproducibility of Psychological Science."

13. Benedict Carey, "Psychology's Fears Confirmed: Rechecked Studies Don't Hold Up," in *New York Times* (2015); "Analysis Casting Doubt on Their Work Is Still Welcomed by Many Psychologists," in *New York Times* (2015); "Report Questioning Psychology Studies Is Criticized," in *New York Times* (2016).

14. Ioannidis, "Why Most Published Research Findings Are False," e124.

15. A.J. Wakefield et al., "Retracted: Ileal-Lymphoid-Nodular Hyperplasia, Non-Specific Colitis, and Pervasive Developmental Disorder in Children," *Lancet* 351, no. 9103 (1998); M.J. LaCour and

D.P. Green, "Political Science. When Contact Changes Minds: An Experiment on Transmission of Support for Gay Equality," *Science* 346, no. 6215 (2014).

16. D.R. Carney, A.J. Cuddy, and A.J. Yap, "Power Posing: Brief Nonverbal Displays Affect Neuroendocrine Levels and Risk Tolerance," *Psychological Science* 21, no. 10 (2010); Amy Cuddy, "Your Body Language May Shape Who You Are," *TED* (2012).

17. Susan Dominus, "When the Revolution Came for Amy Cuddy," *New York Times*, October 18, 2017; Andrew Gelman and Kaiser Fung, "Amy Cuddy's Power Pose Research Is the Latest Example of Scientific Overreach," in *Slate,* January 19, 2016.

18. Monya Baker, "Is There a Reproducibility Crisis?" *Nature* 533 (2016).

19. Ibid.

20. Ibid.

21. Steven N. Goodman, Daniele Fanelli, and John P.A. Ioannidis, "What Does Research Reproducibility Mean?" *Science Translational Medicine* 8, no. (2016): 1.

22. Prasad Patil, Roger D. Peng, and Jeffrey T. Leek, "A Statistical Definition for Reproducibility and Replicability," *BioRxiV*, https://www.biorxiv.org/content/early/2016/07/29/066803 [accessed September 13, 2017].

23. Kenneth Bollen et al., "Social, Behavioral, and Economic Sciences Perspectives on Robust and Reliable Science" (paper presented at the Report of the Subcommittee on Replicability in Science Advisory Committee to the National Science Foundation Directorate for Social, Behavioral, and Economic Sciences, 2015), 3.

24. Ibid., 4.

25. Jeffrey T. Leek and Leah R. Jager, "Is Most Published Research Really False?" *Annual Review of Statistics and Its Application* 4 (2017): 111.

26. Karl Broman et al., "Recommendations to Funding Agencies for Supporting Reproducible Research" (paper presented at the American Statistical Association, 2017), 2.

27. B.A. Nosek et al., "Promoting an Open Research Culture," *Science* 348, no. 6242 (2015).

28. "A Brief History of COS," https://cos.io/about/brief-history-cos-2013-2017 [accessed October 19, 2017].

29. "Top Guidelines," accessed September 13, 2017. https://cos.io/our-services/top-guidelines/

30. "Center for Open Science Announces Elsevier as New Signatory to Top Guidelines," https://cos.io/about/news/centre-open-science-announces-elsevier-new-signatory-top-guidelines/ [accessed September 30, 2017].

31. Center for Open Science, "Guidelines for Transparency and Openness Promotion (Top) in Journal Policies and Practices" (paper presented at the COS, 2015).

32. Broman et al., "Recommendations to Funding Agencies for Supporting Reproducible Research," 3–4.

33. Ibid.

34. NIH, "Principles and Guidelines for Reporting Preclinical Research," *National Institute of Health* (2014).

35. Federation of American Societies for Experimental Biology, "Enhancing Research Reproducibility: Recommendations from the Federation of American Societies for Experimental Biology," http://faseb.org/Science-Policy--Advocacy-and-Communications/Science-Policy-and-Research-Issues/Research-Reproducibility.aspx [accessed May 25, 2017].

36. Society for Neuroscience, "Research Practices for Scientific Rigor: A Resource for Discussion, Training, and Practice," http://www.sfn.org/advocacy/policy-positions/research-practices-for-scientific-rigor [accessed September 13, 2017].

37. C. Glenn Begley, Alastair M Buchan, and Ulrich Dirnagl, "Robust Research: Institutions Must Do Their Part for Reproducibility," *Nature* 525 (2015): 27.

38. Victoria Stodden et al., "Setting the Default to Reproducible Reproducibility in Computational and Experimental Mathematics," (2013), http://stodden.net/icerm_report.pdf [accessed July 26, 2017].

39. Stodden et al., "Setting the Default," 7.

40. Eric Green et al., "National Institutes of Health Advisory Committee to the Director National Library of Medicine Working Group Final Report," (2015), https://acd.od.nih.gov/documents/reports/Report-NLM-06112015-ACD.pdf [accessed July 22, 2017].

41. "JoVE Librarian Travel Award," https://www.jove.com/blog/2017/08/21/jove-librarian-travel-award/ [accessed September 20, 2017].

42. Vicky Steeves, "Reproducibility Librarianship," *Collaborative Librarianship* 9 (2017).

43. "Assistant or Associate Librarian," University of Utah job posting since deleted. Find similar postings at https://utah.peopleadmin.com/postings/ [accessed 11 December 2017].

44. "Research Reproducibility Conference," accessed August 11, 2017. http://campusguides.lib.utah.edu/utahrr16; "The Librarian's Role in Reproducibility of Research Symposium," http://mlasymposium.libguides.com/2017 [accessed August 11, 2017].

45. Steeves, "Reproducibility Librarianship."

46. Melissa L. Rethlefsen, Mellanye Lackey, and Shirley Zhao, "Building Capacity to Encourage Research Reproducibility and #Makeresearchtrue," *Journal of the Medical Library Association (forthcoming article).*

47. "Research Data Management (Rdm) iCoP," https://sites.google.com/a/umn.edu/rdm-cop/home [accessed August 29, 2017].

48. Jon Jeffryes, "Improving Research through Reproducibility," in *Continuum* (2017), https://www.continuum.umn.edu/2017/05/improving-research-reproducibility/ [accessed August 11, 2017].

49. "Institutes," Berkeley Initiative for Transparency in the Social Sciences, https://www.bitss.org/event-types/institute/ [accessed October 30, 2017].