

# Reproducible Research Is Really Freaking Hard

Patrick D. Schloss<sup>†</sup>

<sup>†</sup> To whom correspondence should be addressed: pschloss@umich.edu; Department of Microbiology and Immunology, University of Michigan, Ann Arbor, MI

## **Format:**

<http://mbio.asm.org/site/misc/authors.xhtml> \* Commentaries are short invited articles (1,000 words) that discuss mBio papers or issues of special interest. \* Perspectives are brief reviews (2,000 words) on a topic in which opinion and synthesis are encouraged.

**Counts:** ~XXXX words plus XX references, X figures, and a XXX word abstract

## **Abstract**

2 Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et  
3 dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea  
4 commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat  
5 nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim  
6 id est laborum.

## 7 Introduction

8 In 2011, Philip Bourne challenged those attending the *Beyond the PDF* workshop to reproduce the work  
9 performed in his group's 2010 article *The Mycobacterium tuberculosis Drugome and Its Polypharmacological*  
10 *Implications* [Garijo2013]. The response to that challenge resulted in a unique analysis, that would  
11 challenge concepts critical to the ability to reproduce another group's work when using the same data and  
12 methods (i.e. reproducibility). The investigators demonstrated that the value of reproducibility, the necessary  
13 degree to which research can be reproduced, the amount of effort required to reproduce the research, and  
14 who should be able to reproduce the research are not simple questions. On first blush, one might argue  
15 that any scientist should be able to reproduce another scientist's research with no friction. As Garijo and  
16 colleagues went on to demonstrate, even when a project led by thought leaders such as Philip Bourne,  
17 reproducibility is a thorny issue.

18 Their conclusion is in contrast to the report and accompanying editorial in *mBio* by the American Academy  
19 for Microbiology's (AAM's) 2015 colloquium, "Promoting Responsible Scientific Research". Although reports  
20 from this colloquium used the term reproducibility where others would use replicability (i.e. the ability to  
21 generate the same results after repeating the experiment independently of the first), the reports are a useful  
22 lens into how microbiologists view the reliability of research in the field. The colloquium identified "(i) sloppy  
23 science, (ii) selection and experimental bias, and (iii) misconduct" as the primary contributors to the ongoing  
24 problems with insuring the reliability of their research. Although the participants were quick to point out that  
25 misconduct was a relatively minor contributor to the problem, the four case studies that accompany the  
26 original report all concern misconduct. In addition, problems related to "sloppy science" are also prominent  
27 throughout the reports as well. Missing from these reports is any of the nuance or humility found in the earlier  
28 Garijo study. Few would suggest that Bourne's group was sloppy, yet the follow up reproduction analysis  
29 estimated that it would take a novice at least 280 hours to reproduce the work with at least 160 hours of that  
30 time spent trying to decipher the original methods descriptions.

31 When I ask other microbiologists how difficult it would be for me to reproduce a randomly selected figure  
32 from their latest paper, universally, they turn ashen. Does the inability to reproduce that figure indicate that  
33 their work is incorrect? No. However, the inability to reproduce a value significantly reduces the confidence  
34 that one might have in the work. In contrast, we might also ask, does the ability to reproduce a figure indicate  
35 that the work is correct? No. Investigators may release all of the the workflow to generate a figure from raw  
36 data to the final figure, but a coding error may result in accidentally switching the colors of the legend. Alas,  
37 the ability to interrogate the code, would quickly identify the error and allow the investigators to resolve the

issue. Finally, we might also ask whether research that is not reproducible is replicable. Again, if the methods used to run the study and analyze the data are not transparent, then it would be impossible to confidently understand why the inability to replicate a result was due to the use of a different study population, cell line, or analytical approach. Put simply, most problems with reproducibility and replicability of a study is not due to sloppy science, bias, or misconduct. While those are certainly issues, insuring that one's research design and methods are sufficiently clear is really freaking hard.

- Rehash AAM Report on Reproducibility

- <http://mbio.asm.org/content/7/4/e01256-16.full?sid=e0608d60-2133-4d24-a00c-54579527c544>
- [https://www.asm.org/images/Colloquia-report/Promoting\\_Responsible\\_Scientific\\_Research.pdf](https://www.asm.org/images/Colloquia-report/Promoting_Responsible_Scientific_Research.pdf)

- Goals for this commentary

- Provide better framework for thinking about reproducibility
- Give very reasonable justifications for why a study may not be reproducible
- Focus on analytical issues of reproducibility

## Threats to reproducibility

- Definitions

- Define reproducibility and replicability and introduce the four-celled grid
- Studies that are reproducible or replicable are not necessarily correct
- Point out there are many reasons why a study may fail upon repeating that have nothing to do with fraud
- File drawer effect

- Threats in microbiological data analysis and solutions

- Documentation
  - \* transparency (methods rabbit holes)
  - \* protocols.io
  - \* automation, dependency tracking
  - \* version control
- Link rot / email rot
  - \* ORCID
  - \* DOI
- Data availability

- 68                   \* Standard databases, FigShare, Data Dryad
- 69           – Code availability -> GitHub, Rmd, Jupyter notebooks
- 70                   \* Incorrect code is easier to correct if it is available
- 71           – Expertise
- 72                   \* Bourne article and ability of novices to reproduce workflows
- 73   • Other threats

## 74   **Need for training**

- 75   • Riffomonas project
- 76   • Software and Data Carpentry
- 77   • Coursera Data Analysis MOOC
- 78   • Case studies:
  - 79       – Paper airplane
  - 80       – How would I reproduce a random figure from your a paper in your CV?

## 81   **Reproducibility as Preventative Medicine**

- 82   • Sze case study
- 83   • Reproducible research as a positive force
  - 84       – Your most important collaborator is you 6 from months ago, and old you doesn't have email access
  - 85       – Your second most important collaborator is your PI
  - 86       – Reproducible research is better cited, allows others to build upon your work

## 87   **Conclusion**

88   Leewonhoek's lack of reproducibility anecdote

89   From Leek (<https://simplystatistics.org/2017/11/21/rr-sress/>) \* We can define what we mean by "reproduce"

90   and "replicate" \* We can remember that replication is statistical, not deterministic \* We can remember

91   that there is a difference between exploratory and confirmatory research \* We can remember that a failed

92   replication isn't always a scientific failure \* We can stop publicizing scientific results as solutions so quickly \*

<sup>93</sup> We can be persistent and private as long as possible \* We can make the realization that data is valuable but  
<sup>94</sup> in science you don't own it \* We should cut each other some slack