



Reproducible Research as Our New Default

Evolutionary Anthropology Society

Bret Beheim

The last decade has been a sobering one for science, marked by increasingly frank worries that “*something has gone fundamentally wrong with one of our greatest human creations*” to [quote Richard Horton](#), editor-in-chief of *The Lancet*. He continues:

Afflicted by studies with small sample sizes, tiny effects, invalid exploratory analyses, and flagrant conflicts of interest, together with an obsession for pursuing fashionable trends of dubious importance, science has taken a turn towards darkness.

[Multiple studies](#) in the social and life sciences have identified widespread use of “[questionable research practices](#)”, and many important results have failed to [replicate](#). Researchers working with the [Open Science Collaboration](#) recently attracted a great deal of attention when they were unable to replicate results from 30–60 percent of the 100 psychology studies they re-analyzed.

What does this “reproducibility crisis” mean for scientific anthropology, especially for those of us doing statistical analyses of field data?

From talking to colleagues, most first reactions—mine as well—are that reproducibility is a nonstarter in anthropology. Fieldwork is messy, and our data are observational rather than experimental. Moreover, field data are collected under circumstances so unique in time and location that it’s usually impossible to revisit that same data source again.

Yet there remains *another* kind of replication, one that can be applied to anthropology as it can to any other scientific endeavor. In fact, it’s among the most basic and fundamental components of all of science: if someone else looked at the *same* data you did, would they be able to reproduce your analysis and achieve the *same* results?

If someone else looked at the same data you did, would they be able to reproduce your analysis and achieve the same results?

Seems too simple, right? Last year two economists put this most humble of demands to a collection of recent studies, all peer-reviewed in [high-impact journals](#). Alarming, of 67 papers, they could reproduce the essential results of only 29 in this way. For the majority, the data were either unavailable (even in journals that required data be available), or upon re-analysis gave fundamentally different results. This was true even when running the original analysis script, and in a few cases, even with the direct help of the original authors.

Similar efforts in biology have been equally dismal. Vines, et al. attempted to assemble the datasets from 526 morphometry studies published between [1991 and 2011](#). They were unsuccessful in a shocking 80 percent of cases; half of contacted authors did not respond or were unwilling to share data, and another quarter had no working email. Moreover, in about a fifth of studies published before 1996, the data were not merely unavailable, but actually lost.

Much anthropological data was collected before 1996.

At this point, you might expect a “but”. There is no “but”. It is really that bad. If the above is any guide, we must confront the possibility that many (most?) quantitative results in anthropology cannot be reproduced, because the original datasets are lost to time, or upon reanalysis would not yield the same results.

If you have an existing analysis in print, consider the following as a personal challenge: delete all the intermediate steps, and also the final tables and figures. With only the original data, and the analysis script, can you recover everything else? (This challenge was inspired by Genzkow and Shapiro’s excellent [Code and Data for the Social Sciences: A Practitioner’s Guide](#).)

I recently tried this on [one of my own projects](#) that I thought had been properly archived, and was alarmed to find the answer was “No”. I had left behind multiple, cryptically-named versions of my files, none complete, and none by themselves producing the actual publication results. With much effort, I was able to properly reconstruct my original sequence of analysis steps, and (to my eternal relief) the published tables and figures re-appeared.

Yet the experience was a sobering one. If you cannot reproduce your own results, surely no one else will either, and our basic premise of scientific knowledge accumulation breaks down.

How is this phenomenon of non-reproducibility even possible, let alone apparently at pandemic levels? There's charlatanry, no doubt. And widespread reticence to make available the dataset upon which published claims rest. But often, I would guess, the explanation is simply sloppy documentation.

A statistical analysis in anthropology involves literally thousands of individual steps: table joins, cleaning, outlier removal, subsetting, transforming variables, constructing design matrices, fitting multiple model specifications, running tests, and visualizing results, all to be summarized in a brief Methods section. And that's not even including the actual fieldwork!

The ethic of reproducible work is part of the larger movement towards open source science

Herein lies the real value of scripting your work in a statistical computing language like R or Stata; the code is self-documenting. With everything in timeless, universal plaintext (think .csv, .txt, .r, .dta, or .dat), an analysis can be re-run automatically, and becomes available for review and cumulative improvement. It becomes that much harder to practice charlatany.

The ethic of reproducible work is part of the larger movement towards open source science, including practices such as pre-registering hypotheses, reporting null results, ensuring author availability through the Open Researcher and Contributor ID (ORCID), and using online repositories like GitHub to publicly enshrine code and data. The parallels with open-source software development are striking, and there are some very good resources available for academics eager to develop these skills (e.g., software-carpentry.org).

There are some recent commendable examples of open-source science in anthropology. The authors of the recent work on [Homo naledi](#) made almost the entire project available at publication, including downloadable 3D scans of the specimens. Venkataraman and colleagues included a link to their [data repository](#) as part of a recent PNAS submission, allowing reviewers to comment on and even modify the submitted analysis in their reviews. As Richard McElreath reminds us, "the data and its analysis are the scientific product. [The paper is just an advertisement](#)".

Open-source, reproducible science is [now the default approach](#) for many scientists; let it be so in anthropology as well.

[Bret Beheim](#) is currently a postdoctoral fellow at the University of New Mexico Department of Anthropology. He applies principles of evolutionary ecology to study cultural dynamics in human societies. He received his PhD from UC Davis and is an affiliate of the [Tsimane Health and Life History Project](#).

Melanie Martin and **Katie Starkweather** are contributing editors for EAS's AN column.

[Previous Article](#)

[Back to Top](#)

[Next Article](#)

[Related Categories](#) +

[Related Tags](#) +

◀ 36

Commenting Disclaimer +

© 2016 American Anthropological Association • 2300 Clarendon Blvd, Suite 1301 • Arlington, VA • 22201 • TEL (703) 528-1902 • FAX (703) 528-3546