By **Lorena A. Barba**

The hard road to reproducibility

E arly in my Ph.D. studies, my supervisor assigned me the task of running computer code written by a previous student who was graduated and gone. It was hell. I had to sort through many different versions of the code, saved in folders with a mysterious numbering scheme. There was no documentation and scarcely an explanatory comment in the code itself. It took me at least a year to run the code reliably, and more to get results that reproduced those in my predecessor's thesis. Now that I run my own lab, I make sure that my students don't have to go through that.

In 2012, I wrote a manifesto in which I committed to best practices for reproducibility. Today, a new student arriving in my group finds all of our research code in tidy repositories, where every change is recorded automatically. Version control is our essential technology for record keeping and collaboration. Whenever we publish a paper, we create a "reproducibility package," deposited online, which includes the data sets and all the code that is needed to recreate the analyses and figures. These are the practices that work for us as computational scientists, but the principles behind them apply regardless of discipline.

It takes new students some time to learn how to work to these standards, but we have documentation and training materials to make it as painless as possible. My students don't resent investing their time in this. They know that practices like ours are crucial for the integrity of the scientific endeavor. They also appreciate that our approach will help them show potential future employers that they are careful, conscientious researchers.

I am pleased when our group is recognized for our high standards in other people's writings, and when we are invited to speak about these practices at meetings. But we've found we still have a lot to learn about what it takes for research, even when done to high standards of reproducibility, to be replicated. A couple years ago, we published a paper applying computational fluid dynamics to the aerodynamics of flying snakes. More recently, I asked a new student to replicate the findings of that paper, both as a training opportunity and to help us choose which code to use in future research. Replicating a published study is always difficult—there are just so many conditions that need to be matched and details that can't be overlooked—but I thought this case was relatively straightforward. The data were available. The whole analysis was open for inspection. The

"My students and I continuously discuss and perfect our standards."

additional details were documented in the supplementary materials. It was the very definition of reproducible research.

Three years of work and hundreds of runs with four different codes taught us just how many ways there are to go wrong! Failing to record the version of any piece of software or hardware, overlooking a single parameter, or glossing over a restriction on how to use another researcher's code can lead you astray.

We've found that we can only achieve the necessary level of reliability and transparency by automating every step. Manual actions are replaced by scripts or logged into files. Plots are made only via code, not with a graphical user interface. Every result, including those from failed experiments, is documented. Every step of the way, we want to anticipate what another researcher might need to either reproduce our results (run our code with our data) or replicate them (independently arrive at the same findings).

About 150 years ago, Louis Pasteur demonstrated how experiments can be conducted reproducibly—and the value of doing so. His research had many skeptics at first, but they were persuaded by his claims after they reproduced his results, using the methods he had recorded in keen detail. In computational science, we are still learning to be in his league. My students and I continuously discuss and perfect our standards, and we share our reproducibility practices with our community in the hopes that others will adopt similar ideals. Yes, conducting our research to these standards takes time and effort—and maybe our papers are slower to be published. But they're less likely to be wrong. ∎

Lorena A. Barba is an associate professor at The George Washington University in Washington, D.C. Send your story to SciCareerEditor@aaas.org.

Science

The hard road to reproducibility

Lorena A. Barba