From Open Access to Open Science:
Enhancing Transparency in Social Science Research

Richard Ball
Professor of Economics

Norm Medeiros
Associate Librarian

Directors, Project TIER

Open Access Week
Magill Library
26 October 2017

Project TIER is supported by the Alfred P. Sloan Foundation
10th Annual International Open Access Week, sponsored by SPARC (Scholarly Publishing and Academic Resource Coalition)

Goal: raise awareness of the importance of free, immediate access to published research

Open Access implies access to some finished product, frequently a publication such as a journal article

Project TIER is concerned with the processes that create that finished product—i.e, workflows that are transparent and lead to reproducible results. Well organized replication documentation (data, code, and documentation) promote transparency and reproducibility, and lead to improved student learning.
Project TIER promotes transparency and reproducibility in social science research by

- Developing curriculum and tools for teaching students to conduct empirical research transparently and reproducibly
  - TIER Protocol
  - syllabi and documents from TIER-influenced courses
  - OSF (Open Science Framework) template and demo project
  - exercises (Stata, R Markdown)

- Disseminating these curricula and tools to faculty interested in adopting them—in all the social sciences, at the undergraduate and graduate levels via workshops, conference talks, and our web site
Project TIER grew organically out of our experience teaching an introductory statistics course for students majoring in economics and other social sciences.

2003-2008: The primordial soup

- Observed need for greater clarity in research papers written by students
- Developed guidelines for documenting research intended to ensure this clarity
Over many semesters, we gradually extended and refined these documentation guidelines. In the end, we broke them up into two pieces:

- **Specifications** that describe the content and organization of the final documentation that should be prepared to accompany a completed research paper

- Recommendations on steps to take throughout the *process* of conducting a research paper to integrate the construction of the required final documentation into the entire life-cycle of a research project

All this can be seen at our website...
Example of a real student project:


This thesis is stored on [OSF (the Open Science Framework)](https://osf.io), and is available publicly in [Haverford’s Institutional Research repository](https://repository.haverford.edu).
Following these guidelines transformed the way students conducted their entire research projects

- The final replication documentation results from processes that occur throughout the research lifecycle. Reproducibility is not tacked on at the end.
- These methods distinguish Open Science (a process) from Open Access (a product).

What’s more, these processes that lead to comprehensive and valid replication documentation also (1) increase the efficiency with which students work, (2) help avoid disasters, and (3) really help students understand what they are doing.

It also means that communication with instructors while working on a project is more effective—we can see what students are doing and provide better guidance.
2009-2012: Start of outreach activities

- Presentations at conferences related to teaching, including the first American Economic Association CTREE (Conference on Teaching and Research in Economic Education) in 2011

- Published a paper on this work

Landmark Studies


- 154 requests for data and code
- Received 90 responses with some information
- Reviewed the first 54 cases
- Only 8 replicated (15%)
Landmark Studies (cont’d)


- 39% found to replicate

- Prompted subsequent studies in Cancer Biology and the Social Sciences
Landmark Studies (cont’d)


• Reviewed 59 articles from 13 leading journals

• Used author-deposited data and code

• Initially only able to reproduce 33% of the papers

• After contacting authors, rate of reproducibility went to 49%
Cancer Research Is Broken
There's a replication crisis in biomedicine—and no one even knows how deep it runs.

How the Reproducibility Crisis in Academia is Affecting Scientific Research

Tackling the 'credibility crisis' in science
January 4, 2016

Psychology's Ongoing Credibility Crisis
New studies have intensified the debate over psychology's "reproducibility" problems.

About 40% of economics experiments fail replication survey
By John Bohannon | Mar. 3, 2015, 2:00 PM

The Credibility Crisis in Computational Science: An Information Issue

Science, Now Under Scrutiny Itself

How do we solve science's 'credibility problem'?
April 7, 2015 8:43am EDT
High profile cases


- Analyzed data from 44 countries over 200 years to demonstrate a sharp decrease in economic growth once a country’s debt reaches 90% of its GDP. Led to austerity policies in parts of the world.


- Found coding errors, exclusion of data, and unconventional weighting led to Reinhart and Rogoff’s findings.
High profile cases (cont’d)


- Claimed a relationship between the establishment of legalized abortion and a subsequent decrease in crime 18 years following.


- Found computer programming errors that failed to control for influencers of crime.
High profile cases (cont’d)


- Demonstrate poor compliance with journal archive policies
- Describe failed replications for several high profile studies

*If the sums do not add up, the science is wrong. If there are no sums to be added up, no one can tell whether the science is right or wrong.*

– Donald Laming
Organizations

BITSS (Berkeley Institute for Transparency in the Social Sciences)

Center for Open Science

Replication Wiki

AIRLEAP (Association for Integrity and Responsible Leadership in Economics and Associated Professions)

....and many others
The nudge that got us to jump into the fray and launch Project TIER:

- ICPSR/Sloan Challenge grant (2013); the official start of Project TIER

**Main activities**

- Faculty development workshops twice per year at HC, 8-10 participants on average (upcoming Nov 17-18, 2017)

- Purpose: to help interested instructors teach transparent, reproducible methods in their own classes

- Have found these two-day workshops help faculty adopt the practices

- Held first workshop on the road at Occidental College in January 2017. Looking to hold a similar workshop in Spring 2018 at UCLA.
Workshop attendees

as well as from Brazil, Ecuador, Egypt, and the Netherlands
Other activities

- Grad student workshops—on site at particular grad programs—
  (Duke, Clark, Penn, Cornell, University of Colorado)

- Faculty Fellowships
  - 5-6 individuals per year
  - 3 cohorts so far (2015/16, 2016/17, 2017/18)
  - service project and outreach
TIER in the classroom
One local example

• Bryn Mawr, Sociology 265, Prof. Nate Wright
  
  o Individual students with multiple projects
  o Could easily support collaborative projects with differential access, such as ECON203
What is the real purpose of all this? Why teach students to conduct research transparently?

- Transparency and reproducibility are necessary for research to be scientifically valid and credible.

- Students who have these skills when they graduate from college are well prepared for jobs as research assistants, consultants, financial analysts, etc.

- Documenting their work throughout the research process helps students keep track of what they are doing, and consequently they have a much better understanding of what they are doing.

- The instructor’s ability to provide guidance while students are working on a project is radically enhanced—it is a whole new world.
Most fundamentally, we are teaching students that when they say they have found evidence that supports some kind of claim, they need to be able to show the evidence to others, and reconstruct the argument that leads from the “raw” evidence to the claim.¹

In our view, helping students understand and internalize this principle is one of the principal goals of education.

¹ Frequent and emphatic tweets do not do the trick.