Advancing Computational Reproducibility in the Social Sciences

Creating and using digital reproduction records as pedagogical tools

December 14, 2020

Katie Hoeberling

Berkeley Initiative for Transparency in the Social Sciences (BITSS)

khoeberling@berkeley.edu

The **Berkeley Initiative for Transparency in the Social Sciences** (BITSS) promotes ethical, transparent and reproducible research practices to improve the integrity of science and inspire better public policy.

Generate **evidence**
Increase access to open science **education**
Strengthen the scientific **ecosystem**

Failures to reproduce or replicate

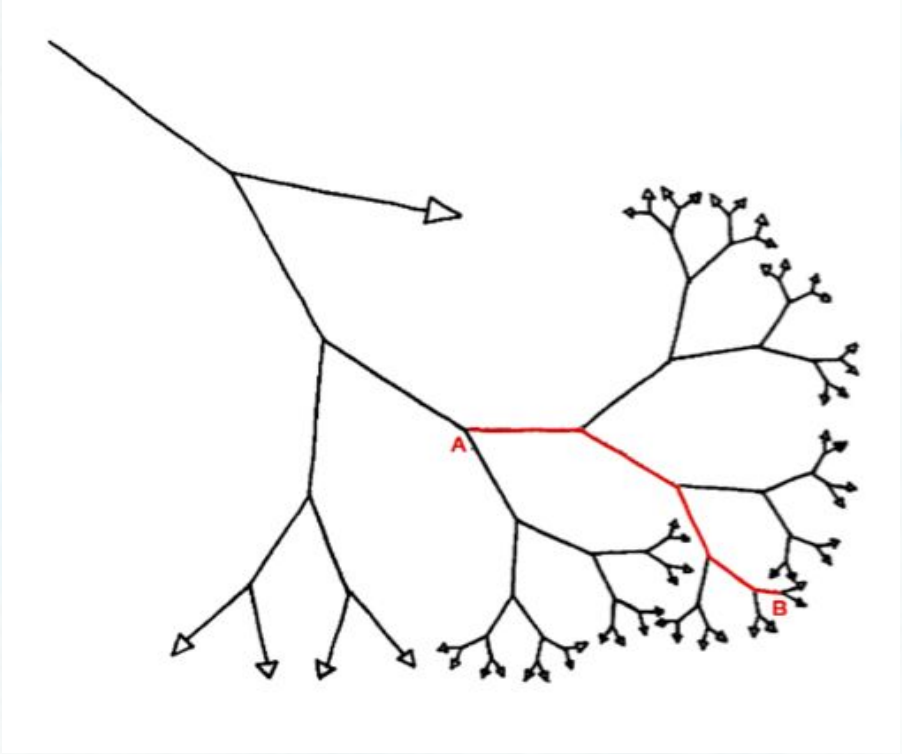
Replication in Social Sciences (same method, different sample)	Reproduction in Economics (same data and methods)
OSC (2015): 30%-60%	Chang & Li (2015): 43%
Camerer et. al. (2016): ~60%	Gertler et. al. (2017): 14%
Nosek & Camerer et. al. (2018): ~60%	Kingi et. al. (2018): 43%
Klein et. al. (2018): 50%	Wood et. al. (2018): 25%

What is “Scholarship?”

“An article about computational results is advertising, not scholarship. The actual scholarship is the full software environment, code and data, that produced the result.”

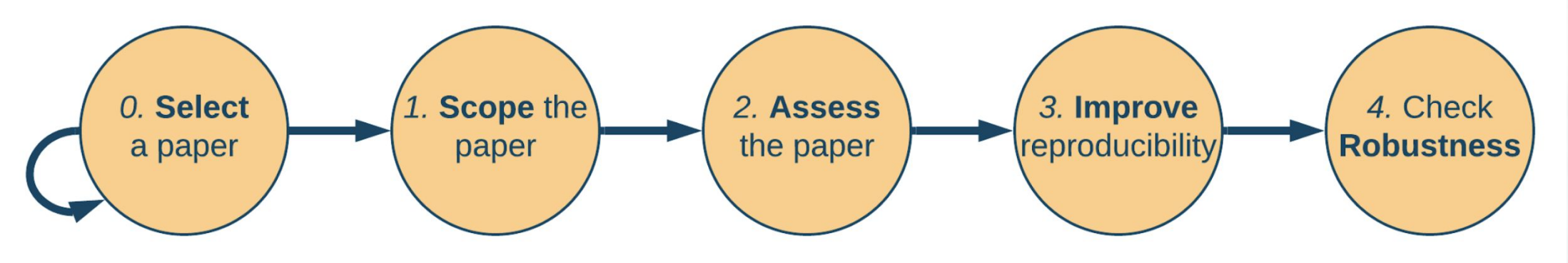
– Buckheit and Donoho (1995) paraphrasing Claerbout and Karrenbach (1992)

- Accelerating understanding
- Supporting learning
- Improving inclusion and participation



^^ From Jorge Luis Borges’s “El jardín de senderos que se bifurcan”

Accelerating Computational Reproducibility in the Social Sciences (The ACRe Project)



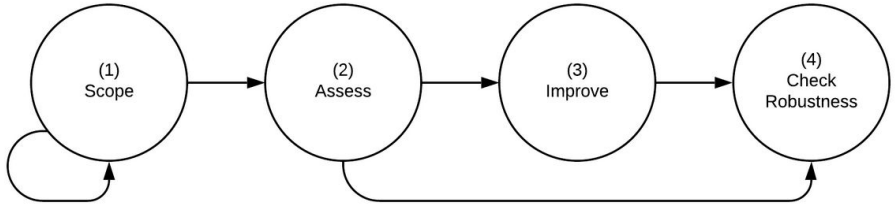
- Curriculum development
- Training for students and instructors
- The Social Science Reproduction Platform: Crowdsourced reproductions, improvements, and community discussion
- Assessments of reproducibility across journals, fields

The ACRE Guide

Step by step instructions for conducting and recording a reproduction, including chapters on:

- Choosing a paper
- Assessing reproducibility
- Making improvements
- Robustness checks
- Constructive conversations w/ authors

https://bitss.github.io/ACRE/



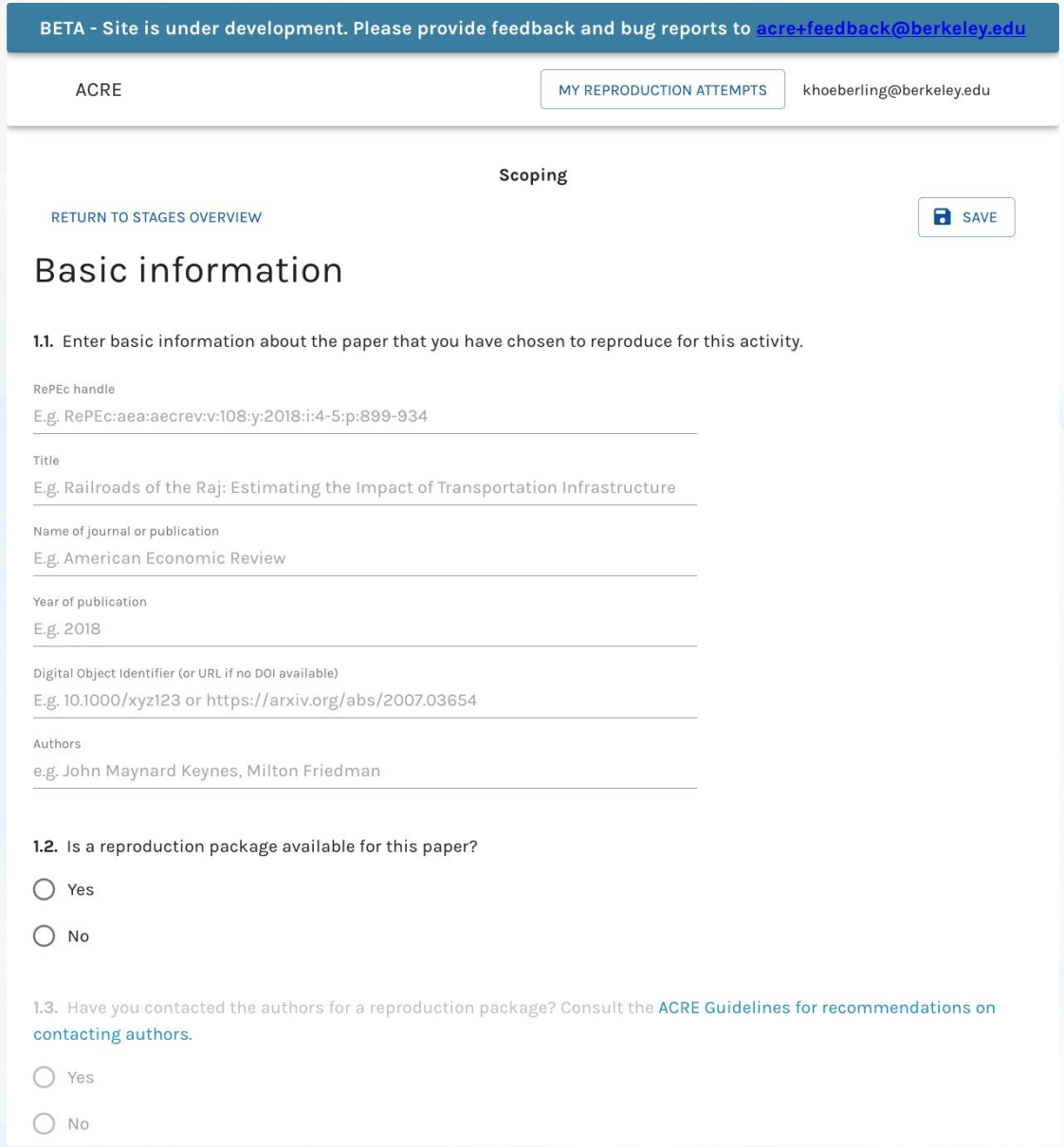
(1) Scoping	(2) Assessment	(3) Improvement		(4) Robustness
		Display-item-level	Paper-level	
☐ Select paper	☐ Describe inputs	☐ + Raw data	☐ + Version control	☐ Analytical choices
☐ Check ACRE	☐ Reproduction diagrams	☐ + Analysis data	☐ + Documentation	☐ Type of choice
☐ Declare paper	☐ Reproduction score	☐ + Analysis code	☐ + Dynamic document	☐ Choice value
☐ Read paper		☐ + Cleaning code	☐ + File structure	☐ Justify and test alternatives
☐ Declare estimates		☐ Debug analysis code		
		☐ Debug cleaning code		
Record results in Survey 1	Record results in Survey 2			Record results in Survey 3

Social Science Reproduction Platform

An open source online platform for crowdsourcing reproductions

- Record citable reproductions and improvements (assigned DOIs w/ Crossref)
- Review and provide feedback on others' reproductions (Discourse forum)
- Aggregated results on a Metrics Dashboard

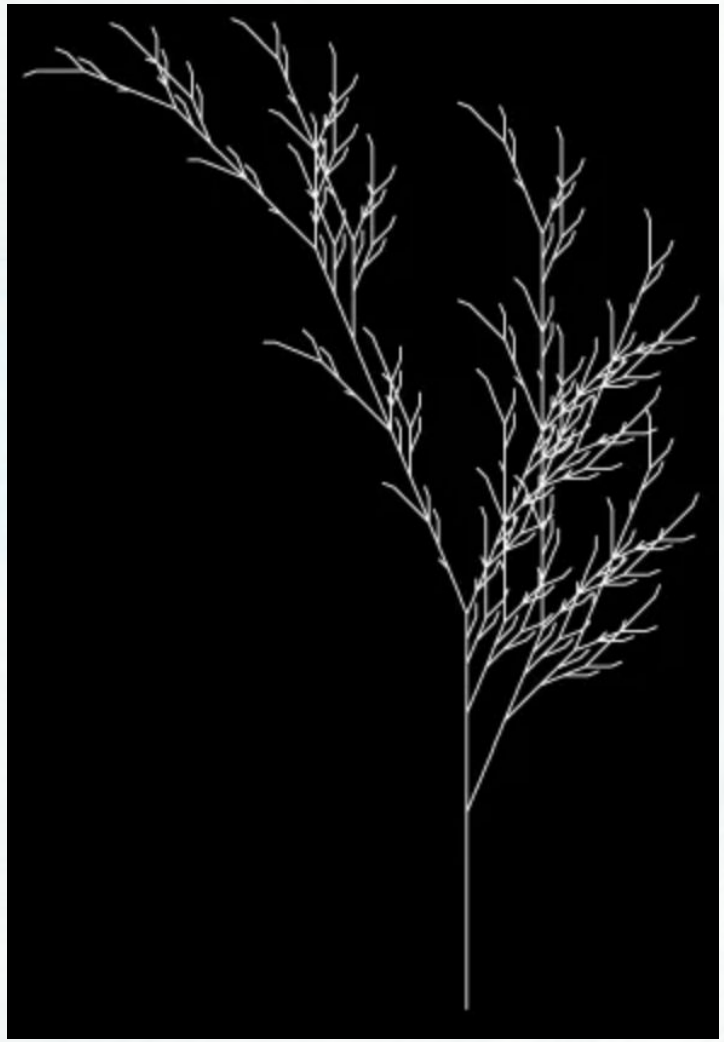
https://www.socialsciencereproduction.org/



Open questions, next steps

- Improving SSRP's pedagogical features
- How to allow **anonymous reproductions**
  - Balancing risks to reputations w/ risks of creating space for trolls
  - De-identifying reproductions is not foolproof
  - Temporary embargo periods
- Assessing **quality**
- Getting **buy-in** from social scientists

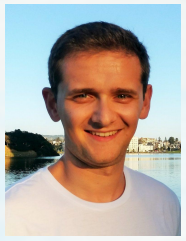
We welcome feedback and community contribution to the Guide and platform!

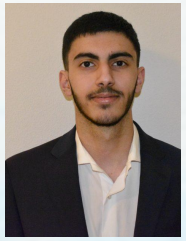


^^ From Antonio Sánchez Chinchón via Fronkinstin

Acknowledgements





- Conception, supervision: Ted Miguel, Fernando Hoces de la Guardia, Lars Vilhuber
- Project management: Aleks Bogdanoski, Katie Hoeberling
- Web development: Albert Chae
- Form development, testing: Joel Ferguson (GSR); Faris Alrubaish, Samira Gholami, Charles Love, Aleksandra Ma, Emma Ng, Lauren Peterson, Jiayu (Ella) Shi (URAPs)
- Funding: Arnold Ventures
- Curriculum piloting: Students from Ted's Development Economics courses, Marc Richter and Dina Pomeranz (ETH)