UNIVERSITY OF CAMBRIDGE
Social Sciences Research Methods Centre

Transparency in Social Science Research & Teaching

Dr Nicole Janz

Science and Engineering South Consortium: *Time to Share* Conference
Cambridge, July 2016

The Economist

OCTOBER 19TH–25TH 2013

Economist.com

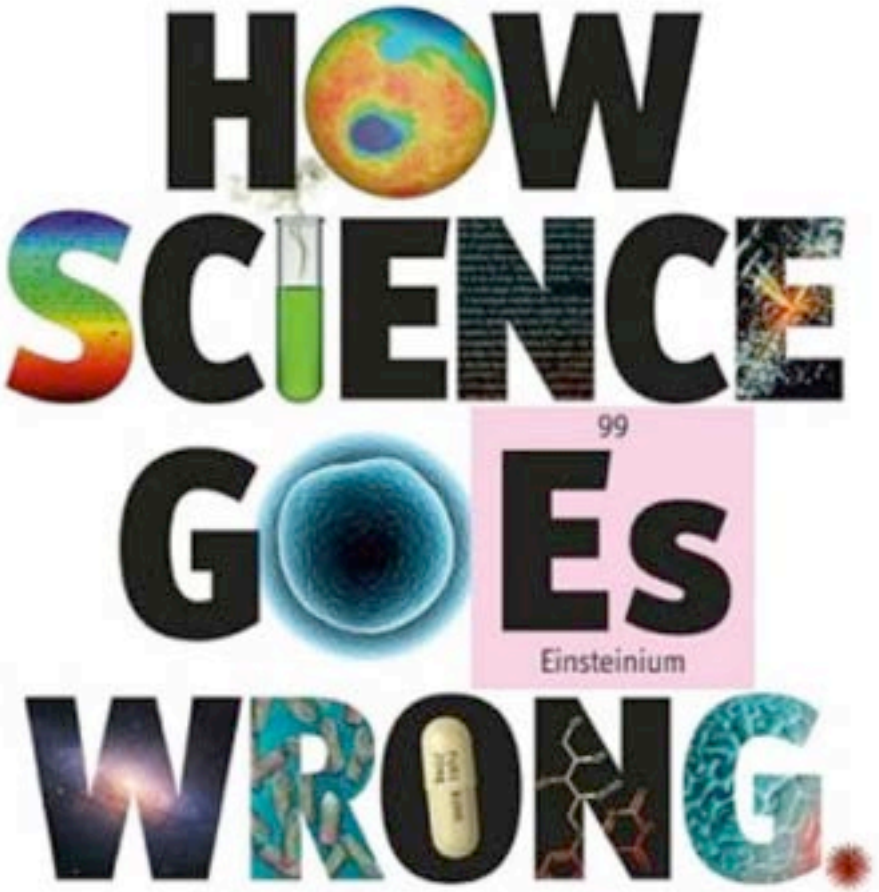
Britain's angry white men

How to do a nuclear deal with Iran

Investment tips from Nobel economists

Junk bonds are back

The meaning of Sachin Tendulkar



SundayReview

The New York Times

Why Do So Many Studies Fail to Replicate?

Gray Matter

By JAY VAN BAVEL MAY 27, 2016

Unreliable research

Trouble at the lab

Scientists like to think of science as self-correcting. To an alarming degree, it is not

Oct 19th 2013 | From the print edition

Timekeeper Like 22K Tweet

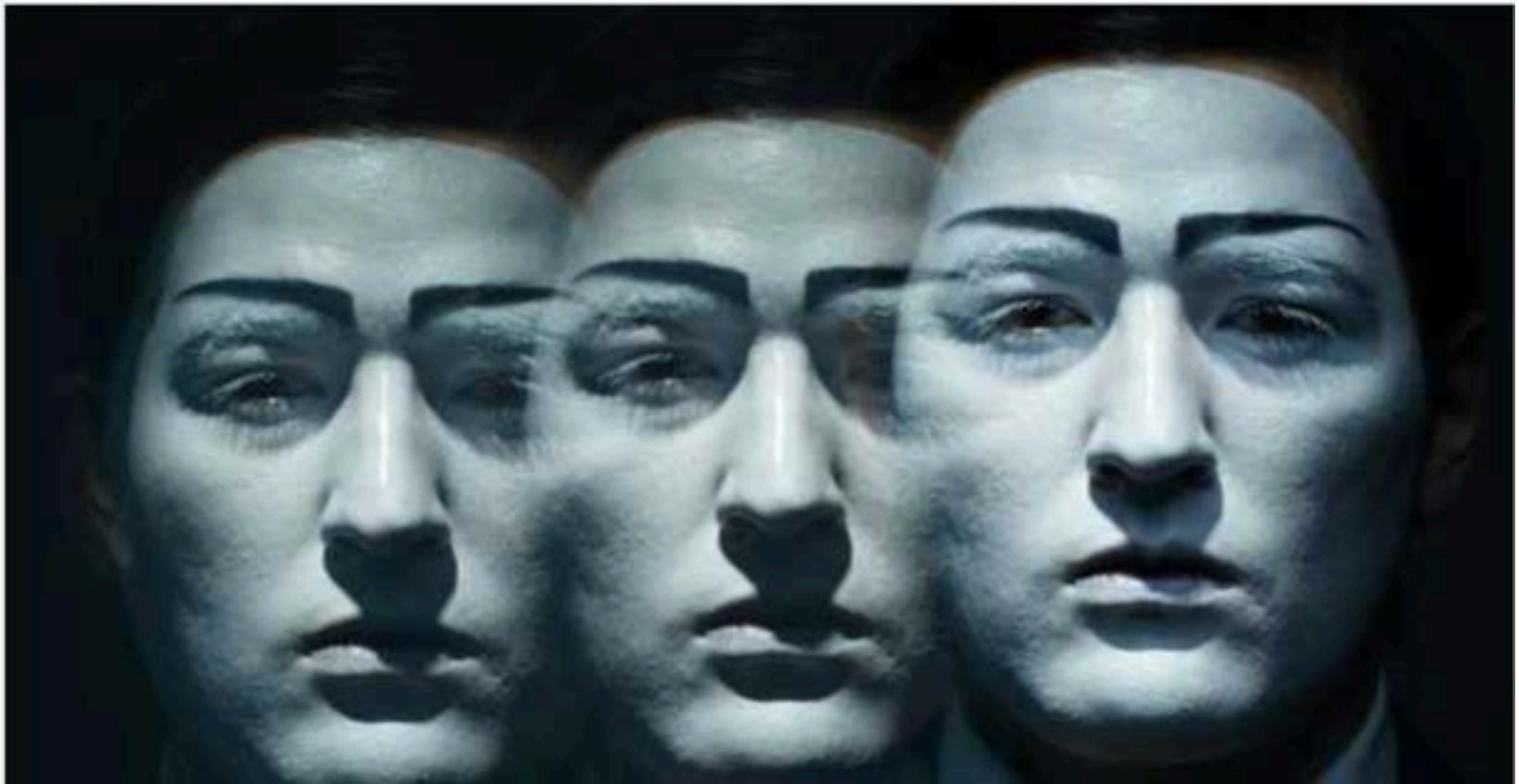
"I SEE a train wreck looming," warned Daniel Kahneman, an eminent psychologist, in an

More than half of psychology papers are not reproducible

Initiative to replicate findings of 100 prominent studies raises further questions about health of discipline

August 27, 2015

By Paul Jump Twitter: @PaulJump



How to replicate

What is a replication?

Duplication	Replication
Verify research results	Test the robustness of the original research results
exact same data set	***new data***
exact same methods	***new models***

Which study should I pick?

Relevant research with **impact**

The perfect replication project

Abstract

I'm the perfect replication project because I combine all these, or at least most of these, features: interesting & relevant questions, results that are accepted but have never been checked, fail to control for important variables, use out-dated measurements, make you wonder if the results apply in different contexts, I'm pointed at in "limitations" and "future research" sections of articles, I'm in an area 'ripe for replication'.

Keywords: *replication, relevant, improvement*

Outdated measures

Results widely **accepted** but never checked

Missing **control** variables

Examples of a ‘good pick’

Reinhart & Rogoff. 2010. “Growth in a Time of Debt.”

Argument: high debt is associated with lower growth

Impact:

- high journal (The American Economic Review)
- research was used by governments to justify austerity measures

Growth in a Time of Debt

By Carmen M. Reinhart and Kenneth S. Rogoff*

Practical steps in a replication study

1 Select paper

2 Access data & code

3 Identify each variable

2-3 weeks

4 Reproduce tables, figures

5 Compare

3-4 weeks

If you got to this point, you completed a **duplication.**

Practical steps in a replication study (II)

6 Add value
- new data
- new variables
- new model specifications
- theoretical contributions

4-6 weeks

7 Compare

8 Get feedback from peers

9 Journal submission

months

You now completed a full **replication!**

Communicating failed replications

What replicators write

"We … find that coding errors, selective exclusion of available data, and unconventional weighting of summary statistics lead to **serious errors**" (Herndon et al. 2013)

"If we cannot even reproduce the original results using the same publicly available data, there is **no need for further commentary**." (Miller et al, 2001)

How original authors often **respond**

"less realistic", "inconsistent with the substantive literature," and "**of limited utility**" (Mansfield, Milner, and Rosendorff 2002)

"fundamentally **flawed**"
(Peffley, Knigge, and Hurwitz 2001)

"statistical, computational, and reporting errors that **invalidate its conclusions**" (Gerber and Green 2005:301).

Publishing a replication study

- Good replication studies get published
- Write a **solid** paper (puzzle, relevance, hypothesis, research design, findings, discussion) – as if it was an **original** piece.
- In some fields (politics): **Don't sell it as a replication** paper

Quarterly Journal of Political Science, 2010, 5: 339–356

Voting Costs and Voter Turnout in Competitive Elections

Bernard Fraga[1] and Eitan Hersh[2,*]

Our estimation approach builds off of the methodology and data used by Gomez et al. (2007) …, adding measures of electoral closeness in order to focus on how the randomly assigned cost (rain) has a different impact depending on the electoral environment.

same way even to rain, then serious doubt should meet claims that voters will react

Political Regimes and International Trade: The Democratic Difference Revisited

XINYUAN DAI *University of Illinois at Urbana–Champaign*

How do domestic political institutions affect the outcomes of international trade negotiations? Specifically, are the aggregate trade barriers agreed upon by a democratic pair lower than those by a pair composed of a democracy and an autocracy? I revisit these important questions

I revisit these important questions by highlighting some problematic aspect of the analysis by Mansfield, Milner, and Rosendorff (2000).

Contrary to their central conclusion, I find that whether the aggregate trade barriers are lower for a democratic pair than those for a mixed pair depends on the preferences of the decision makers involved.

Journals Open to Replication (selection)

Political Science

Research & Politics ^

*

*

Psychology

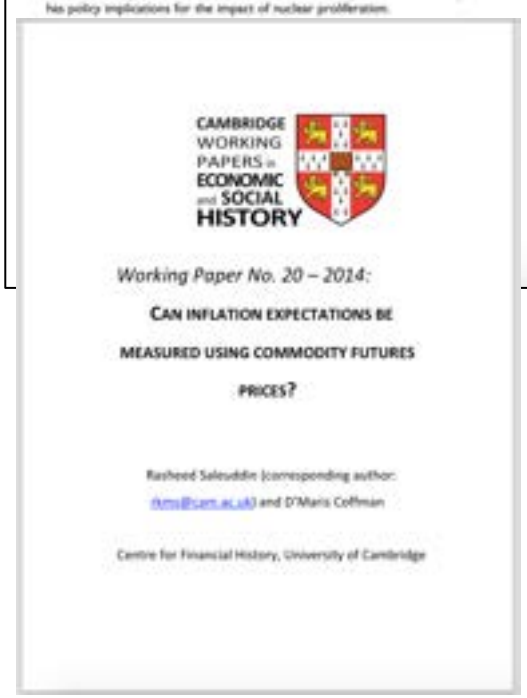
+

Economics

#

*original study was published in the same journal
+ home of the original 'Many Labs' project
# special issue dedicated to replications (March 2015)
^this journal invites replication studies

Replications by Early Career Researchers



International
Studies
Perspectives

International Studies Perspectives (2015), 1–16.

Bringing the Gold Standard into the Classroom: Replication in University Teaching[1]

NICOLE JANZ
University of Cambridge

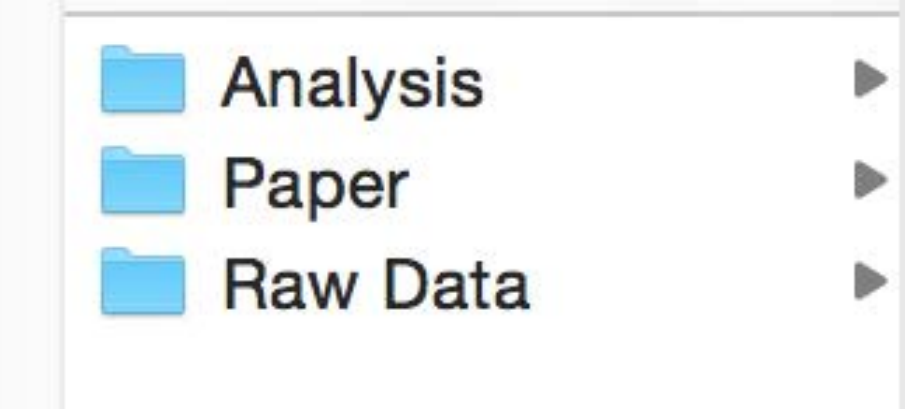
Reproducibility is held to be the gold standard for scientific research. The credibility of published work depends on being able to replicate the results. However, there are few incentives to conduct replication studies in political science. Replications are difficult to conduct, time-consuming, and hard to publish because of a presumed lack of originality. This article sees a solution in a profound change in graduate teaching. Universities should introduce replications as class assignments in methods training or invest in new stand-alone replication workshops to establish a culture of replication and reproducibility. This article will

How to work transparently

Working reproducibly

Starting out…

- Plan file structure
- Never touch raw data!



Analysis

- Comment your code
- Keep a log of decisions

Writing up

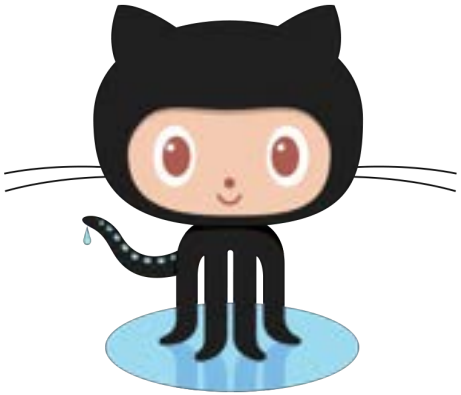
- clarity in methodology section; appendix

Before you submit: Replicate your results !!!

What to share - quantitative

1. Readme file
2. Dataset
3. Software commands
4. Information to reconstruct data

OSF



Github

The Dataverse Project

discourse analysis
content analysis
ethno-graphy
process tracing
participant observ.
focus groups
interviews
Qualitative Analysis

AJPS AMERICAN JOURNAL of POLITICAL SCIENCE

AJPS, South Kedzie Hall, 368 Farm Lane, East Lansing, MI 48824
ajps@msu.edu, (517) 884-7836

GUIDELINES FOR PREPARING REPLICATION FILES

Version 2.1, May 19, 2016

William G. Jacoby
Robert N. Lupton

Michigan State University

The *American Journal of Political Science* requires the authors of all accepted manuscripts to provide replication files before the article enters the production stage of the publication process. The replication files for each article must be made available as a Dataset (i.e., a collection of files) located in the *AJPS* Dataverse on the Harvard Dataverse Network. Instructions for getting started on the *AJPS* Dataverse can be found in the "Quick Reference for Uploading Replication Files,"

Support your claims - qualitative

satellite images, interview transcripts, personal diaries, video clips, newspaper articles, speeches…

Transparency Appendix:

- how you **evaluated** persuasiveness & consistency of evidence
- logic and **steps** in process tracing
- Upload **files** and **fragments** e.g. partial transcripts (100-150 w.)

QDR BETA
QUALITATIVE DATA REPOSITORY

When to protect the data

- Confidential / proprietary data
- Protect individuals
- Informed consent obtained?

- **Anonymization**
- **Justify** why you withhold data

What's in it **for me?**

avoid
disaster

easier to
write up

persuade
reviewers

enables
continuity

build your
reputation

Markowetz, F. (2015) Five selfish reasons to work reproducibly. Genome Biology

Get in Touch

Twitter: **@PolSciReplicate**

http://**PoliticalScienceReplication**.wordpress.com/

Vising scholar 2016 and Catalyst

Ambassador

Useful resources

Materials - Replication

- King, Gary. (2006). **How to Write a Publishable Paper as a Class Project**,copy at: http://gking.harvard.edu/papers

- Janz, N. (2015) **Bringing the Gold Standard Into the Class Room: Replication in University Teaching**, International Studies Perspectives, Article first published online: 9 March 2015. Copy at: http://tinyurl.com/q2qnrvn

- Brandt et al. (2014) **The Replication Recipe: What makes for a convincing replication?** Journal of Experimental Social Psychology, Vol 50, pp. 217-224. Copy at: http://tinyurl.com/poe474k

Materials – Transparent Workflow

- Christensen, Garret (2016). **Manual of Best Practices in Transparent Social Science Research** https://github.com/garretchristensen/BestPracticesManual

- Open Science Framework. **Transparency and Openness Promotion (TOP) Guidelines**. https://cos.io/top/

- **TIER Documentation Protocol** https://www.haverford.edu/project-tier/protocol-v2

Adding value to a replication

1. Theoretical contribution: questioning the arguments
2. Statistical contribution

Sample size: Power calculations (how big should the sample be?) More years, more countries (units) New samples (experiments) Different subsets of your data set (e.g. only OECD countries) Missing data handling (multiple imputation)	**Model specification:** Standard errors treatment, LDV, lags Interactions Dummy variables Omitted variables Reversed causality Adjusted / improved / advanced models
Changing measurements: Change of variables: %GDP, log transformation, different ways of dealing with negative values for logging, different measurement for the same variable	**Robustness/Sensitivity checks:** How much do betas and standard errors change when we change model specifications? Are they very 'sensitive' even to small changes/outliers?