

What is Reproducibility and why it matters for you

Coding in Public Workshop
2024



Table of Contents

- What is reproducibility?
- Why you should care
- Special considerations for social and data sciences
- Q&A
- Break out groups #1

What is reproducibility

What is reproducibility?

Getting “the same” results



[This Photo](#) by Unknown Author is licensed under [CC BY](#)

Reproducibility in brief

Essential part of the scientific method

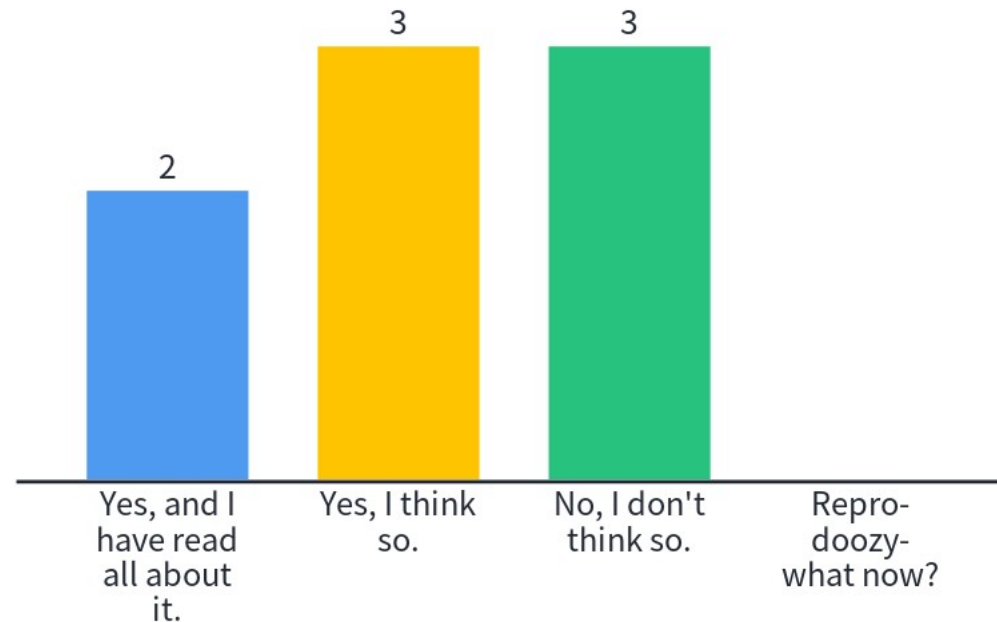


Maybe repeat an *entire* scientific work

Maybe narrowly recreate an analysis



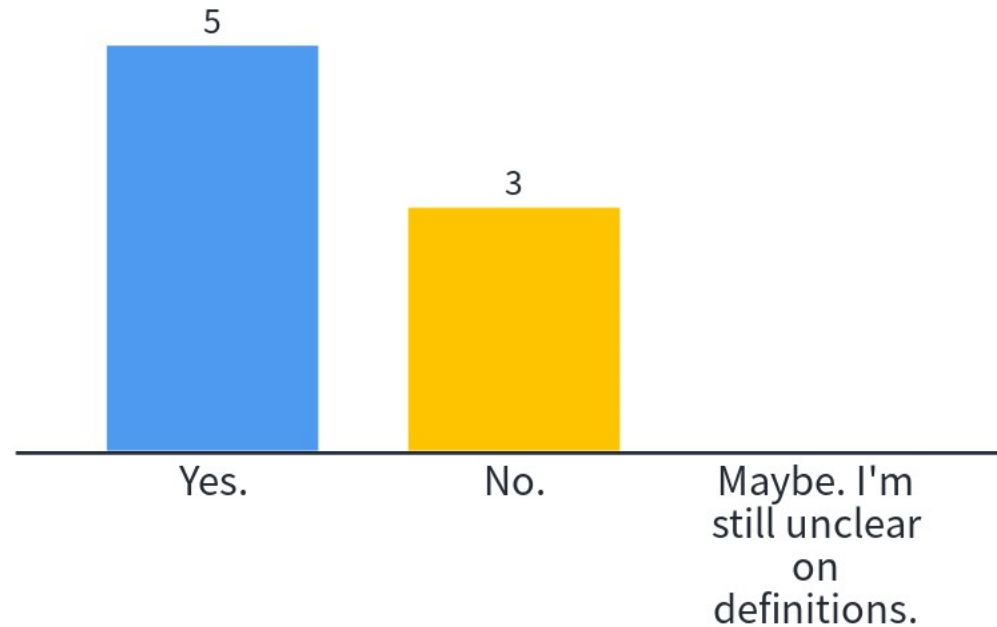
Have you heard of the “reproducibility crisis”?



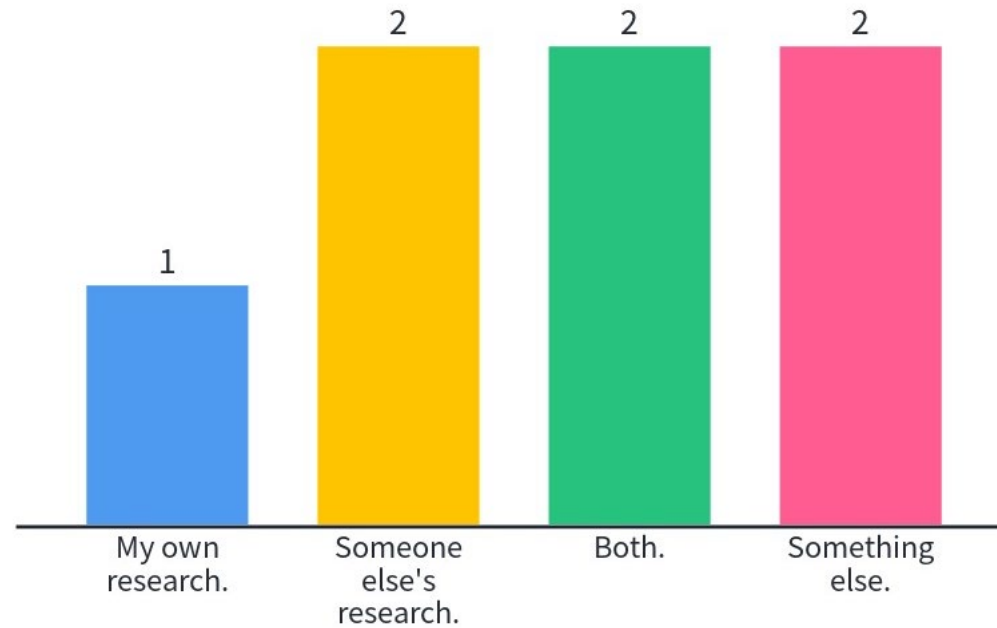
The reproducibility crisis!

- *Social Psychology* and “repligate” (Chambers 2014)
- More than 70% of researchers have tried and failed to reproduce another scientist's experiments (Baker 2016, Chambers 2014)
- More than 50% have failed to reproduce their own experiments (Baker 2016, Chambers 2014)
- Few published their efforts to reproduce (especially “failed” ones) (Baker 2016, Chambers 2014)
- Few contacted the original researchers with questions about reproduction fails (Baker 2016, Chambers 2014)
- New activism/movements about open science, ‘metascience’, etc.

Have you tried to reproduce a research project?



What did you try/want to reproduce?



Reproducible research is like a good recipe

NOT like the technical challenge on Great British Bake Off



What do you think is driving this crisis?

pressure to publish

methods

data

code

ignorance



Does the crisis of reproducibility affect some researchers more than others?

The field of the research

Lack of training

Less technological fields: literature for example

Less emphasis in certain fields about documentation/code/metadata

Fear of research being taken

Type of work; article, poster, paper, workshop etc

Why you should care

Consequences for society

- Inappropriate take up e.g. pro-austerity “proof” (Alesina 2010, Krugman 2013, Devries 2011)
- Wasted time, money, research efficiency, etc. (Freedman 2015, Mack 2014, Roesch 2020)
- Loss of trust (Chambers 2014, Freedman 2015, Peterson 2021)



[This Photo](#) by Unknown Author is licensed under [CC BY-NC-ND](#)

Consequences for science and scientists

- Wasted time, money, research efficiency, etc. (Freedman 2015)
- Distorted research culture (Baker 2016, Freedman 2015, Roesch 2020)
- Loss of trust (Chambers 2014, Freedman 2015)
- Reputational damage (Chambers 2014)



<https://kellystanford.co.uk/science-pusheen>

What impacts matter to you most?

Trusting research

Societal damage and
impact on how people view
social science subjects

Scientific progress
slows down

bad research and
'corrupt' research
culture

Reputation damage
Miss understanding or miss
use

Do these impacts affect some researchers more than others?

norms and expectations
in different fields

removing inconvenient
data points and p-
hacking

Longstanding problems of research culture

Share or protect knowledge?
(Kasmire 2021)

Who gets the credit?



[This Photo](#) by Unknown Author is licensed under [CC BY](#)



[This Photo](#) by Unknown Author is licensed under [CC BY-SA-NC](#)

Collaboration? Or Competition?

- Fear of being scooped is a long-standing problem
- Credit is always hard to apportion, especially after things have got out of control
- Both need proper documentation!

More share/protect problems in research culture

- Education
- Hiring/promoting
- Publishing
- Funding
- Post-publication engagement
- Reputation
- Other?

Which can contribute to **Questionable Research Practices!**

What questionable research practices or “soft fraud” have you seen/done/heard about?

MMR immunisation

Not reporting all results,
removing extremes,
tweaking p-values

P-hacking

Removing inconvenient
data points and p-
hacking

Manipulation of plots
(x/y variables)

Avoid the questionable, be reproducible

Good “research recipes” should include:

- Data (raw to research-ready, details of acquisition, representativeness, etc.)
- Tools (materials, software, packages, versions, etc.)
- Decisions made (in cleaning, processing, recoding, etc.)
- Analysis (order of operations, code where possible, graphs and numbers, etc.)
- Results (clear and as objective as possible)
- Access (to data, code, notebooks, recordings, transcripts, etc.) IF POSSIBLE

Reproducible research recipes are not easy

They take time, effort, and knowledge which means...

Reproducible research recipes can feel very vulnerable.

Vulnerability

Showing your work leaves you open to:

- Criticism about your effort or knowledge (not all valid)
- Competition (not always a bad thing)
- Requests that take time (not always worthwhile)
- SPAM (the modern age)

What makes researchers feel (especially) vulnerable?

reporting negative results, admitting the research didn't go as planned.

Imposter syndrome and fear of people seeing my inelegant code

Calculation failure

social vulnerability

Criticism/Imposter syndrome

Concern that you've done something wrong
Concern that competitors will try to catch you out
Concern that not sharing results or data legally and ethically

Fear of critiquing or playing down the research

What might reduce researcher vulnerability?

Good examples

Mentoring, good colleagues, practice

Mentorship from someone with experience in reproducibility/documentation/tools to use

Introduce these concepts at the undergraduate level

collaboration, having support of colleagues

Further training
Start internal with sharing code at early stages

Special considerations for social and data sciences

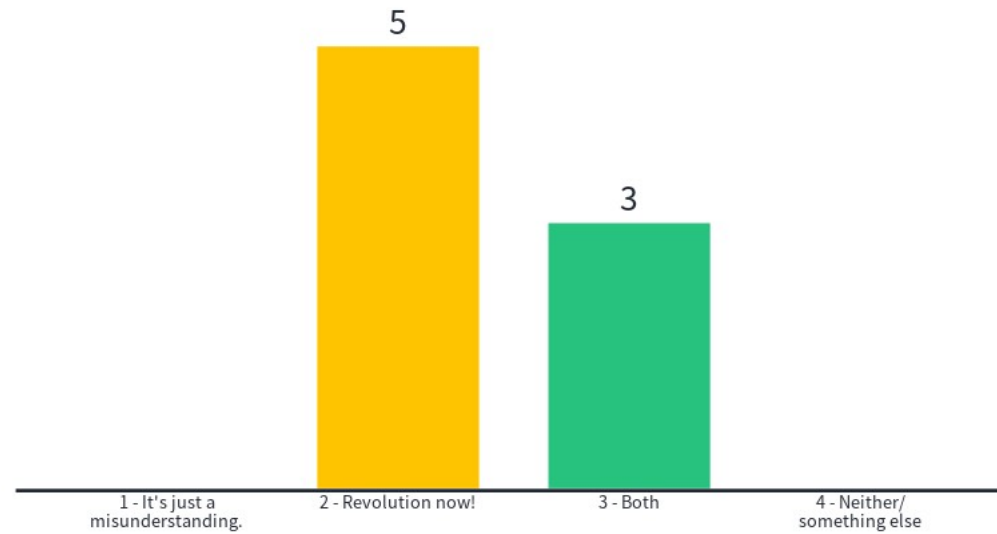
Reaction #1 – It's just a misunderstanding

- “Crisis” is overblown, reactions may be ham-fisted
(Peterson 2021)
- Reproduction efforts should require permission/collaboration
of original author
(Chambers 2014)

Reaction #2 – Revolution now!

- Make “soft” sciences more like “hard” sciences
(Chambers 2014)
- Market based solutions e.g. incentives, apps, journals, etc.
(Roesch 2020)

Which of these reactions makes more sense?



All the interesting stuff in “soft” sciences is...

- Not usually easy to measure, at least directly
- Can be VERY difficult to formalise
- Don't often have agreed on definitions
- Rely on subjective judgements/scores
- The results cannot always be meaningfully separated from the context of the research

Still, even in “soft sciences”

But! More evidence is more persuasive!

“Soft” science research is unlikely to have a big effect if it isn’t well supported and open to scrutiny.

Problems with reactions #1 and #2

- Hard sciences are not exempt
(Baker 2016, Freedman 2015, Roesch 2020)
- Efforts to improve show *some* improvement...
But only in the most high-impact journals (Ball 2018)
- Nobody like to be forced to change, especially through seemingly irrelevant “solutions” imposed from above (Peterson 2021)

Multiple, mutually reinforcing solutions for big change

Better training,
mentorship,
experience

Self/buddy
replication,
replication
journals

Consolidate or
standardise
methods/material
s/protocols

More time,
rewards, support,
incentives

Change
publishing
practice

Share data, code,
etc.

Attitude change

Field-generated
or field-specific
solutions

Mack 2014,
Freedman 2015,
Baker 2016, Ball
2018, Roesch
2020, Peterson

Do these solutions make the vulnerability issues better or worse?

Relief

Monetary incentives are quite powerful, and I can see that leading to more reproducible work. Specific reproducibility journals are also a good idea, could have a ripple effect on other journals.

Yes, but it will take time for it to materialise. Culture changes doesn't happen fast.

Redefine the problem and goal

- 100% reproducibility is not the goal
- Go cautiously
- Embrace complexity
- Embrace collaboration and/or competition
- Strive for open science
- Automate the boring stuff (Sweigart 2019)
- Recognise and deal with vulnerability

Any additional thoughts on reproducibility or vulnerability in “soft” sciences?

Quali data partic
difficult as you said .

Different journals ask for
different things, datasets,
code etc - should be
standardised

Coding templates for
publication, good practice
guides for the field of
research.

References

- Alesina, A. a. A., Silvia (2010). "Large changes in fiscal policy: taxes versus spending." Tax policy and the economy **24**(1): 35--68
- Baker, M. (2016). "1,500 scientists lift the lid on reproducibility." Nature **533**: 452-454.
- Ball, P. (2018). "High-profile journals put to reproducibility test." Nature.
- Chambers, C. (2014). Physics envy: Do 'hard' sciences hold the solution to the replication crisis in psychology? The Guardian.
- Devries, P. a. G., Jaime and Leigh, Daniel and Pescatori, Andrea}, (2011). "A new action-based dataset of fiscal consolidation." IMF working paper.
- Freedman, L. P. a. C., Iain M. and Simcoe, Timothy S. (2015). "The economics of reproducibility in preclinical research." PLoS biology **13**(6).
- Kasmire, J. (2021). "Managing balance: Pursuit of equilibrium permeates the history of science and influences contemporary investigations." Humanistic Management Journal **6**(1): 133--146.
- Krugman, P. (2013). "How the case for austerity has crumbled." The New York Review of Books **6**.
- Mack, C. A. (2014). "In Praise of the Null Result." Journal of Micro/Nanolithography, MEMS, and MOEMS **13**(3).
- Peterson, D. (2021). "The replication crisis won't be solved with broad brushstrokes." Nature **594**(7862): 151--151
- Roesch, E. a. R., Nicolas P (2020). " New journal for reproduction and replication results." Nature **581**(7806).
- Sweigart, A. (2019). automate the boring stuff with Python: practical programming for total beginners, No Starch Press.

Q&A

Break out #1