



Reproducible Science Origins & Definitions

Cyril Pernet
Melanie Ganz
Robert Oostenveld

Rules of engagement

- you can stop me anytime to ask questions
- whenever a gray rectangle appears, an activity/answer is needed, often in small groups (say 3 people)
- group yourselves now

Question box

Intended Learning Outcomes

- List factors limiting reproducibility
- Being able to explain differences between 'commonly' used terms like reproducibility, replicability, generalizability, etc ..
- Being able to argue why sharing matters for reproducibility

Why Most Published Research Findings Are False

John P. A. Ioannidis

2005. *PLoS Medicine*, 2(8), e124. doi:
10.1371/journal.pmed.0020124

“There is increasing concern about the reliability of biomedical research, with recent articles suggesting that up to 85% of research funding is wasted.”

Bustin, S. A. (2015). The reproducibility of biomedical research: Sleepers awake!
Biomolecular Detection and Quantification

THE LANCET

Online First Current Issue All Issues Special Issues Multimedia Information for Authors

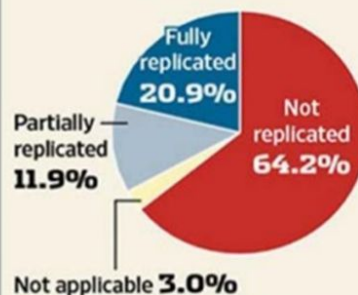
All Content Search Advanced Search

Research: increasing value, reducing waste

Published: January 8, 2014

No Cure

When Bayer tried to replicate results of 67 studies published in academic journals, nearly two-thirds failed.



Source: Nature Reviews Drug Discovery

nature

International weekly journal of science

Home News & Comment Research Careers & Jobs Current Issue Arch

News & Comment News 2015 May Article

NATURE | NEWS

First results from psychology's largest reproducibility test

The scientific truth

A scientific theory is based on testing hypotheses, i.e. predictions.

Testing of hypotheses is done via experiments following the scientific method.



Where can lack of reproducibility happen in this 'cycle' ? and why it matters if one wants to get to the (current) truth

Implications can be huge (practically and financially)

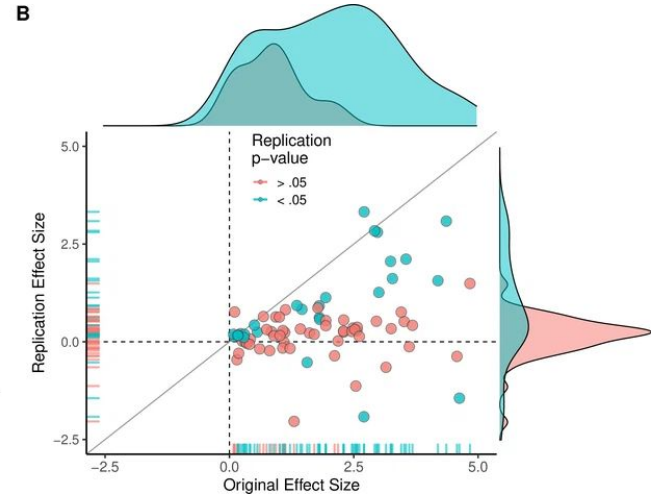
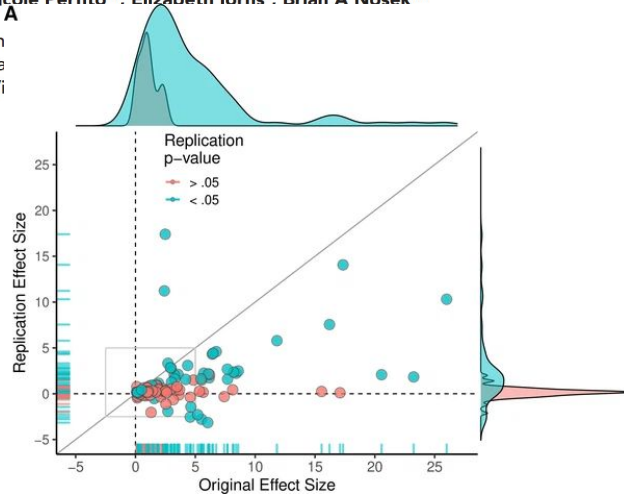


Investigating the replicability of preclinical cancer biology

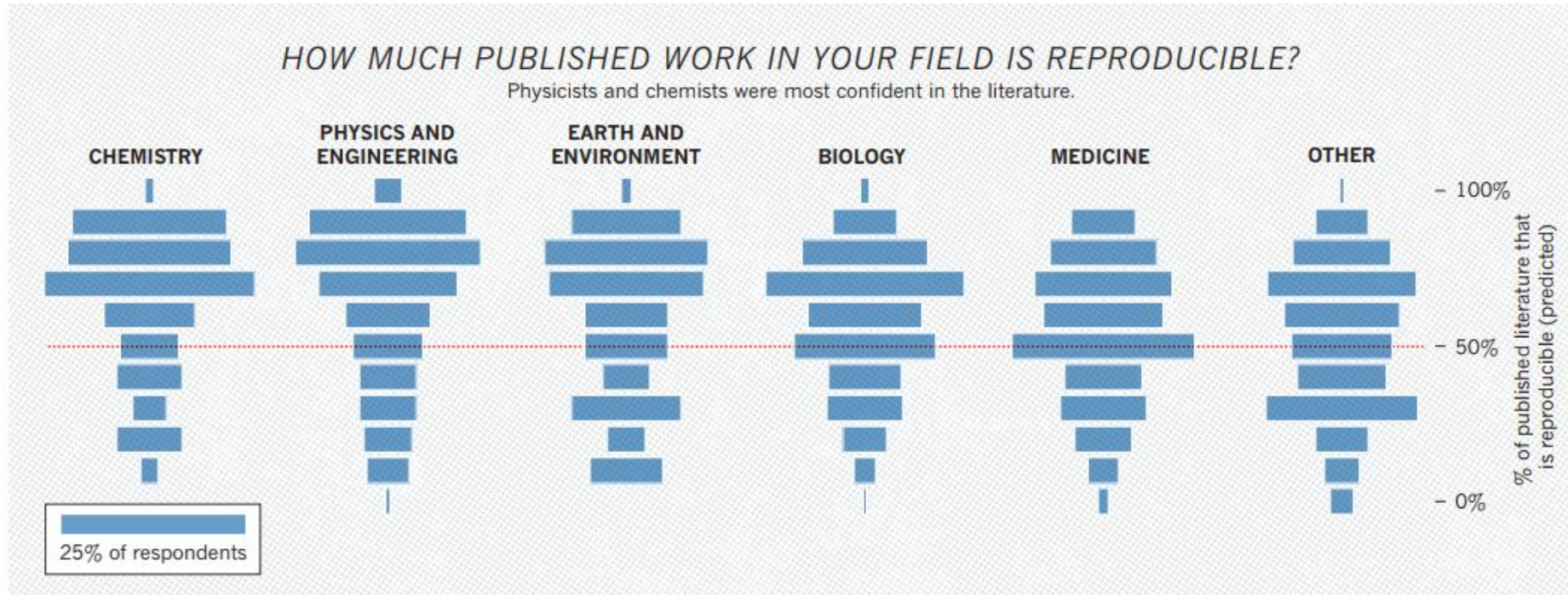
Timothy M Errington^{1*}, Maya Mathur², Courtney K Soderberg¹,
Alexandria Denis^{1†}, Nicole Perfito^{1‡}, Elizabeth Iorns³, Brian A Nosek^{1,4}

¹Center for Open Science
Stanford University, Sta
States; ⁴University of Vi

50 experiments from 23 papers, 158 effects
the median effect size in the replications was 85% smaller
For positive effects, 40% of replications (39/97)
For null effects 80% of replications (12/15)



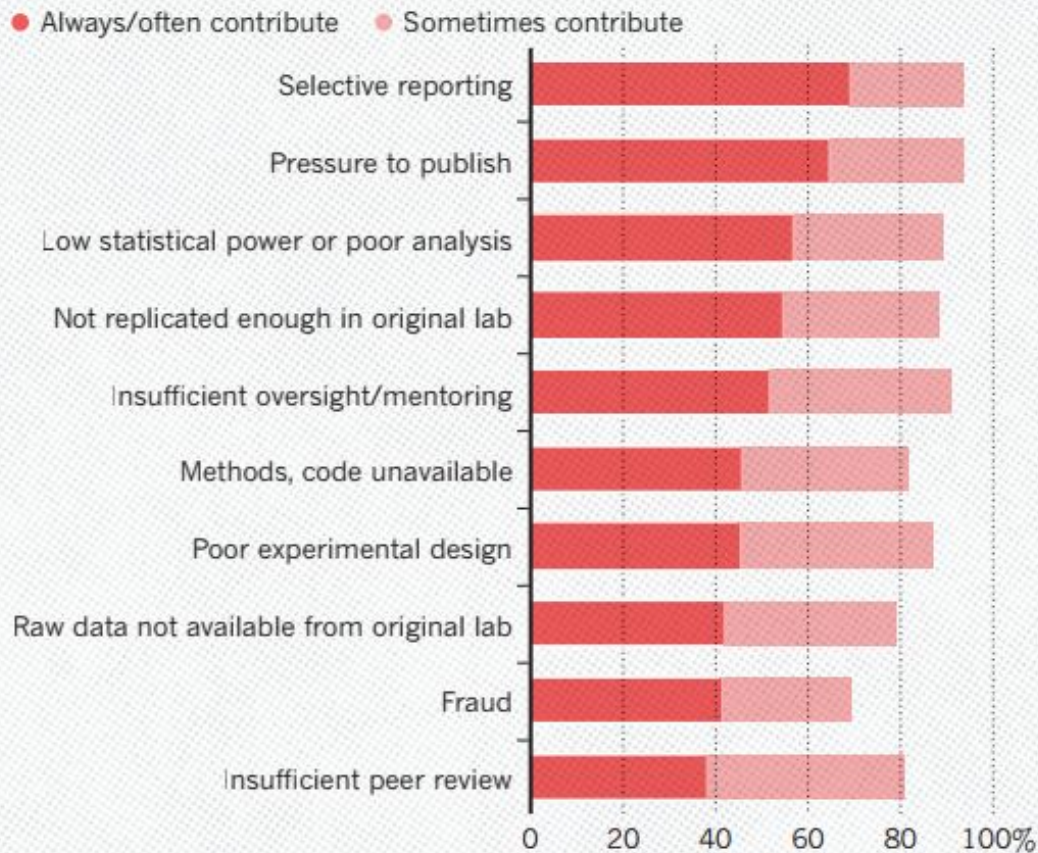
Nature survey 2016. N = 1500



<https://www.nature.com/articles/533452a>

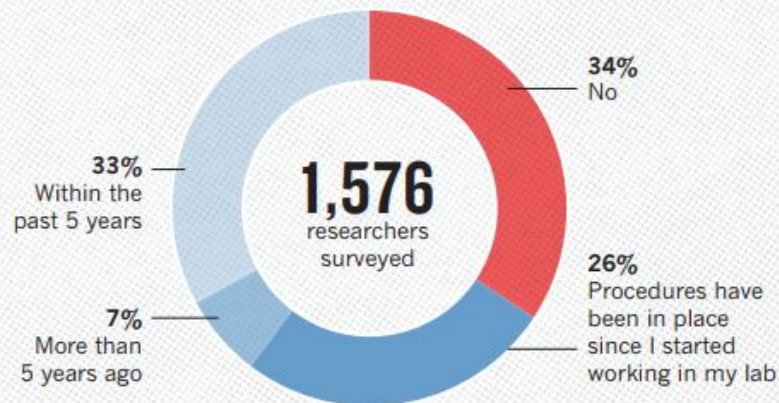
WHAT FACTORS CONTRIBUTE TO IRREPRODUCIBLE RESEARCH?

Many top-rated factors relate to intense competition and time pressure.



HAVE YOU ESTABLISHED PROCEDURES FOR REPRODUCIBILITY?

Among the most popular strategies was having different lab members redo experiments.



Learning about organizing your data, sharing data, good code practices, code sharing, computational reproducibility, and few stats tricks will greatly enhance the reproducibility of your work.

What can we not redo?

- There is a problem in science, research is not **repeatable**, **replicable**, **reproducible**, **generalizable**.

**Are those terms interchangeable?
What are your definitions?
- 5 minutes brain storming -**

Reproducibility, repeatability and replicability

- Repeatability: repeat evaluations under identical conditions (can you re-run your own analysis?)
- Reproducibility: re-compute the same things (is your research transparent so that someone can re-do your analysis?)
- Replicability: achieve consistent results (using a new sample can we obtain the same results?)
- Statistical analyses aim at statistical generalizability: inferring from a sample to a target population.
- Scientific generalizability, on the other hand, refers to applying a model based on a particular target population to other populations.

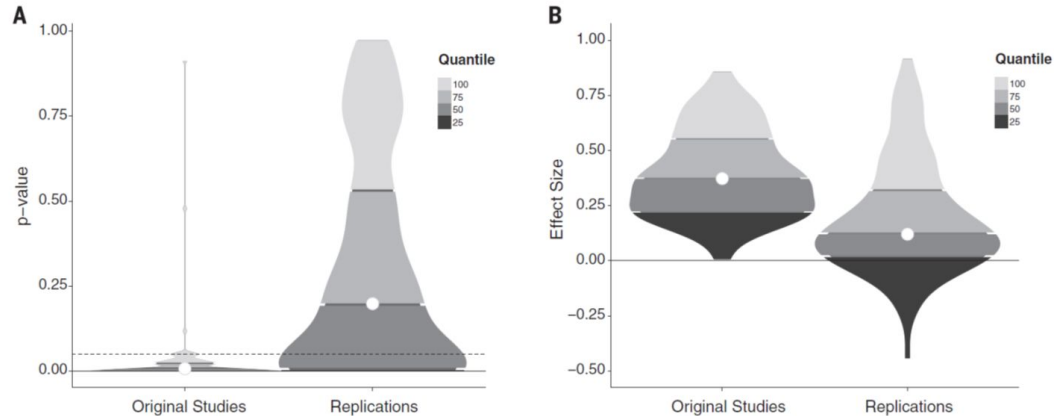
Reproducibility (again)

- Reproducibility gives credibility to a result while replicability increases evidential weight.
- With Method reproducibility, we need to ensure that someone else can redo the same (=sharing data and code), while with result reproducibility (i.e. replicability), we want the same results from different data. In this case, only equivalence of effect sizes matter and not binary results (significant or not) really matter. Inferential reproducibility refers to conclusions obtained vis-à-vis a hypothesis of interest and is best obtained by using cumulative evidence.

Why focus on effect sizes?

PSYCHOLOGY

Estimating the reproducibility of psychological science



Most of research is evaluated using null hypothesis significance testing, i.e. $P(\text{data}|\text{null})$. When there is no effect, p-values are uniformly distributed so we expect to publish false positives!

Focusing on effect size, i.e. quantifying effect allows focusing prediction of quantities allowing building models. → [check the p_value folder in the github repo](#)

More definitions: the Turing Way

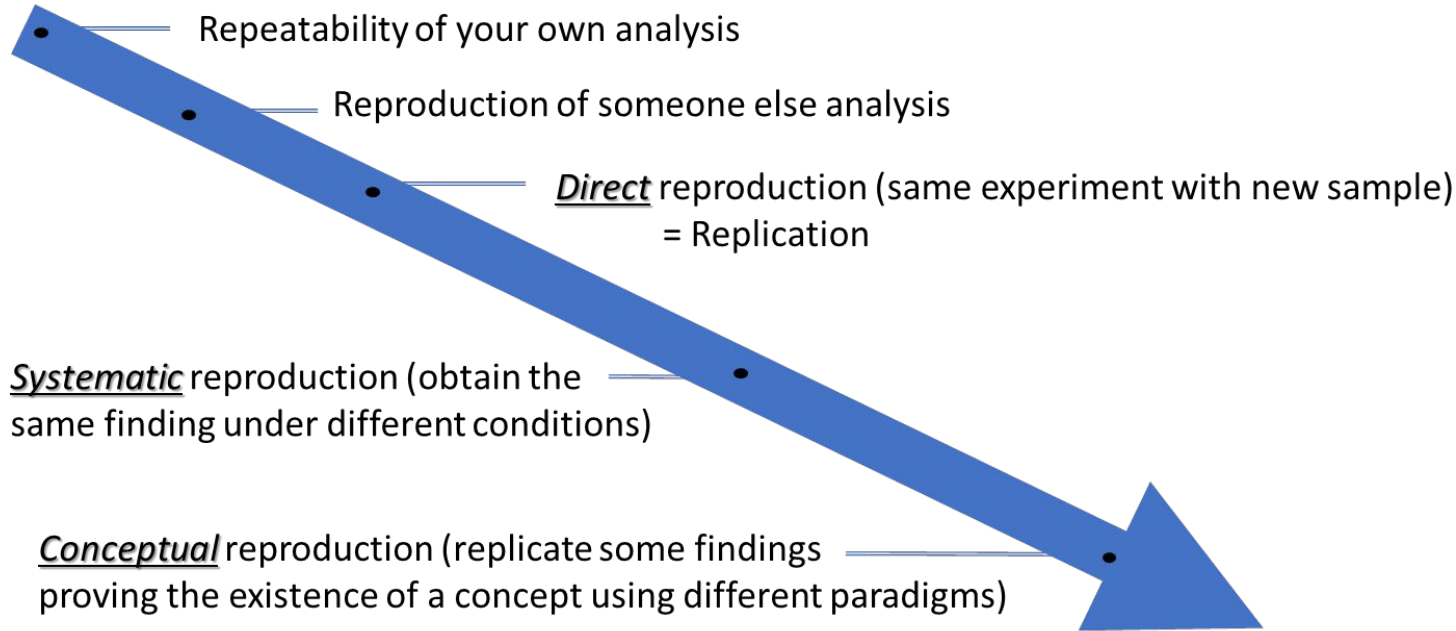
		Data	
		Same	Different
Analysis	Same	<i>Repeatable and Reproducible</i>	Replicable
	Different	Robust	Generalisable

The Turing Way: A Handbook for Reproducible Data Science (Version v0.0.4). Zenodo.

<http://doi.org/10.5281/zenodo.3233986>

<https://the-turing-way.netlify.app/reproducible-research/overview/overview-definitions.html>

Reproduction as a continuum



COMMENT

Open Access

Five selfish reasons to work reproducibly



Florian Markowetz

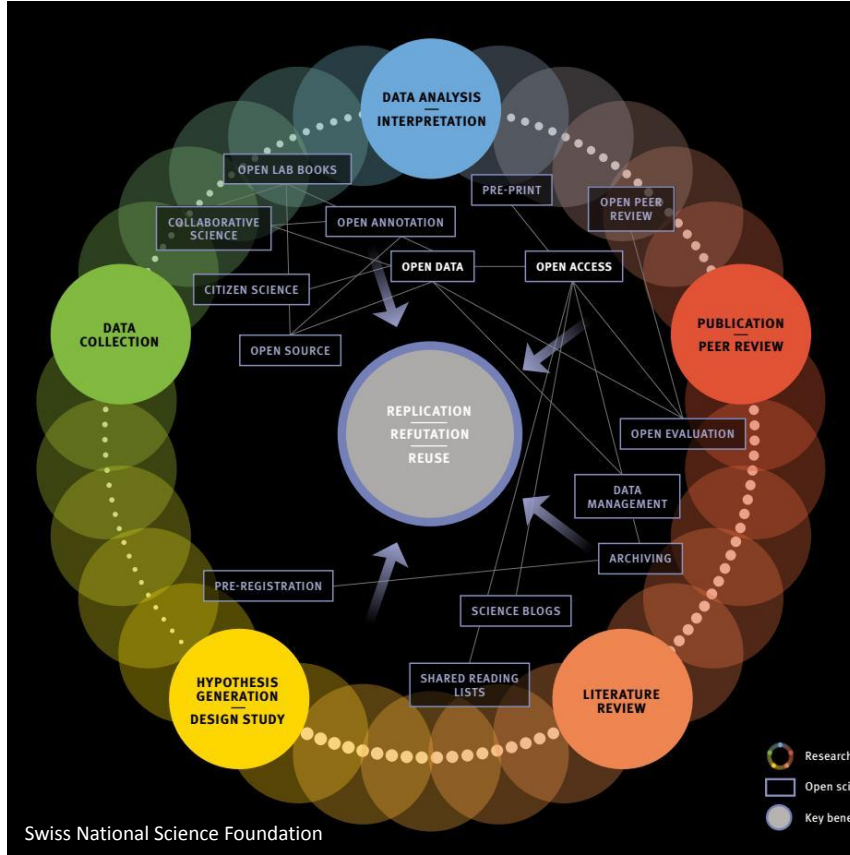
1. helps to avoid disaster (publish spurious result, cannot repeat your experiment)
 2. makes it easier to write papers (redo an analysis, get someone else to check and do things)
 3. helps reviewers see it your way (understand your fancy new analysis)
 4. enables continuity of your work (pick-up work and redo - improve)
 5. helps to build your reputation (citation advantage)
- <https://doi.org/10.1371/journal.pone.0230416>

A word on replicating studies and analyses

- 'No isolated experiment, however significant in itself, can suffice for the experimental demonstration of any natural phenomenon' (Fisher, 1960)
- If a study hasn't been replicated, nothing warrants results are true – replication allows saving time and money in the long run.
- 'A means for acquainting yourself with the methods used in a study, the original author's line of thinking, the complications he or she must have faced, and the solutions they devised to those problems' (Gary King, Harvard univ.)

Open Science

What is open science ?

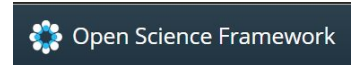
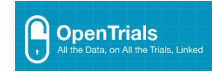


Open science is the movement to make **scientific research, data and dissemination** accessible to all levels of an inquiring society, amateur or professional. https://en.wikipedia.org/wiki/Open_science

Open publication
Open evaluation



Open hypothesis-design
= pre-registration



Journals !

Open data

- Data sharing
- Code sharing



Data Sharing

“Experience has shown the advantage of occasionally rediscussing statistical conclusions, by starting from the same documents as their author. I have begun to think that no one ought to publish biometric results, without lodging a well arranged and well bound manuscript copy of all his data, in some place where it should be accessible, under reasonable restrictions, to those who desire to verify his work.”

Galton Biometry. Biometrika. 1901;1(1):7–10.

What is open data useful for?

Code Sharing

Behavioral Ecology and Sociobiology (2021) 75: 103
<https://doi.org/10.1007/s00265-021-03036-x>

EDITORIAL



Why sharing data and code during peer review can enhance behavioral ecology research

Esteban Fernández-Juricic¹

What is code sharing useful for?

Key Concept

SCIENTIFIC DATA

Amended: Addendum

OPEN
SUBJECT CATEGORIES
» Research data
» Publication
characteristics

**Comment: The FAIR Guiding
Principles for scientific data
management and stewardship**

Mark D. Wilkinson *et al.*[#]

**F
A
I
R**

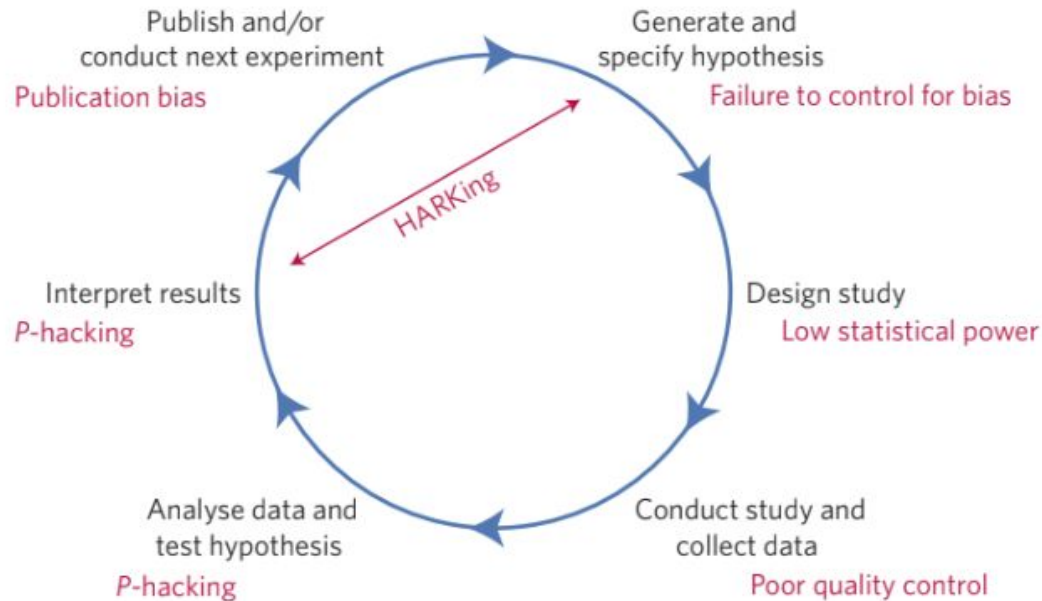
<https://www.go-fair.org/fair-principles/>

For the rest of the course ...

A manifesto for reproducible science

Marcus R. Munafò^{1,2*}, Brian A. Nosek^{3,4}, Dorothy V. M. Bishop⁵, Katherine S. Button⁶,
Christopher D. Chambers⁷, Nathalie Percie du Sert⁸, Uri Simonsohn⁹, Eric-Jan Wagenmakers¹⁰,
Jennifer J. Ware¹¹ and John P. A. Ioannidis^{12,13,14}

Improving the reliability and efficiency of scientific research will increase the credibility of the published scientific literature and accelerate discovery. Here we argue for the adoption of measures to optimize key elements of the scientific process: methods, reporting and dissemination, reproducibility, evaluation and incentives. There is some evidence from both simulations and empirical studies supporting the likely effectiveness of these measures, but their broad adoption by researchers, institutions, funders and journals will require iterative evaluation and improvement. We discuss the goals of these measures, and how they can be implemented, in the hope that this will facilitate action toward improving the transparency, reproducibility and efficiency of scientific research.



The various lectures and tools and exercises aim at providing you with the conceptual and practical tools to make more reproducible science and therefore advance science in general.

Figure1 | Threats to reproducible science. An idealized version of the hypothetico-deductive model of the scientific method is shown. Various potential threats to this model exist (indicated in red), including lack of replication⁵, hypothesizing after the results are known (HARKing)⁷, poor study design, low statistical power², analytical flexibility⁵¹, P-hacking⁴, publication bias³ and lack of data sharing⁶. Together these will serve to undermine the robustness of published research, and may also impact on the ability of science to self-correct.

Intended Learning Outcomes

- List factors limiting reproducibility
- Being able to explain differences between 'commonly' used terms like reproducibility, replicability, generalizability, etc ..
- Being able to argument why sharing matters for reproducibility