UNIVERSITY OF LEEDS

Opening the Social Sciences

Why You Should Care and How You Can Do It

Eike Mark Rinke (POLIS)

SRMC Seminar Series, 28th June 2023

BITSS

TIER

UK RN

ReproducibiliTea

SOCARXIV
open archive of the social sciences

mzes open social science conference 2019

Join our Leeds UKRN network on MS Teams!

Join our Leeds UKRN network on MS Teams!

Why Open Science / Research / Scholarship?

- Science as system of public knowledge production

- “Open practice” is about two things:
  - ***sharing*** (OA, OER, OS, OD)
  - ***showing*** (OM, OD)

PROFESSION SYMPOSIUM

Opening Political Science

Open Minds, Open Methods: Transparency and Inclusion in Pursuit of Better Scholarship

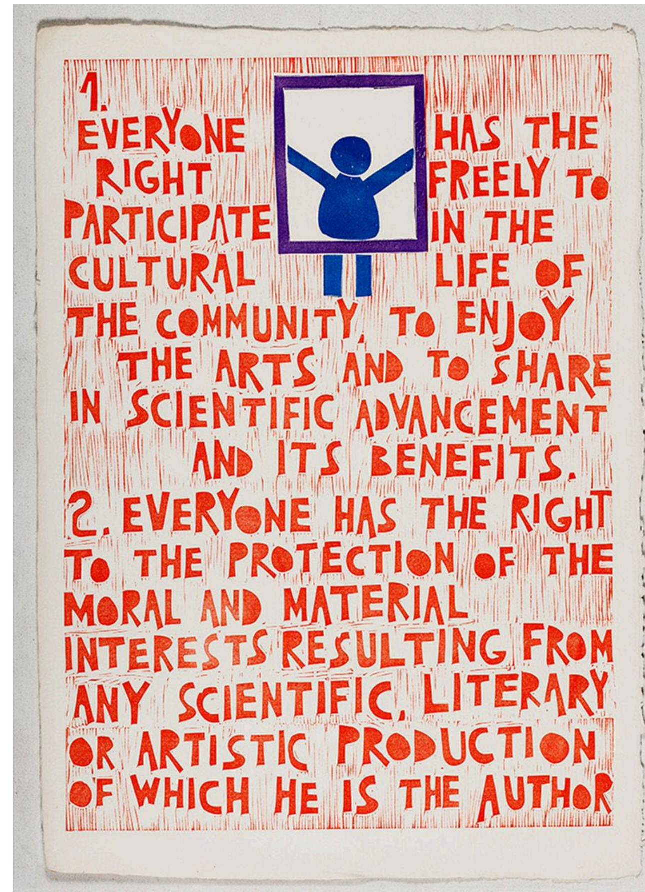
Eike Mark Rinke, *University of Leeds*
Alexander Wuttke, *University of Mannheim*

its purposes, it helps t
the nature of research
ledge claims to be cr
knowledge claims fro
An important tradi
going back to at least
Oreskes (2019), argues
social character. In thi
is distinctive in that sc
evaluate one another's
to critique and engag
incremental—knowled
only "stand on the sho
like a scrupulous ortho

[Rinke & Wuttke (2021)]

Why Open Science / Research / Scholarship?

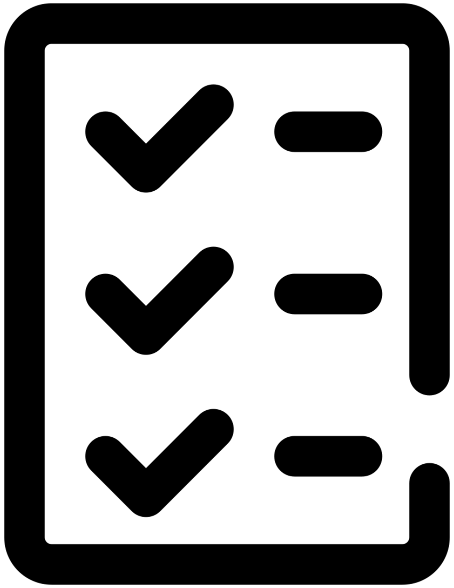
- Two interlocked ends of open science:
  1. Social-moral: Sharing → accessibility → equality & equity (Porsdam Mann et al., 2018; Radder, 2017)



Porsdam Mann et al. (2018)

Why Open Science / Research / Scholarship?

- Two interlocked ends of open science:
  2. Epistemic: Showing → reliability (e.g., reproducibility & replicability) → credibility (Ridder, 2013)



An emerging view of Open Social Sciences as

- **a *broad* tent** (covering all types and aspects of social research)

- **a *context-sensitive* framework** (avoiding imposition of meanings across methodological and epistemological boundaries)

nature human behaviour

PERSPECTIVE

PUBLISHED: 10 JANUARY 2017 | VOLUME: 1 | ARTICLE NUMBER: 0021

OPEN

UNIVERSITY OF LEEDS

A manifesto for reproducible science

Marcus R. Munafò[1,2*], Brian A. Nosek[3,4], Dorothy V. M. Bishop[5], Katherine S. Button[6], Christopher D. Chambers[7], Nathalie Percie du Sert[8], Uri Simonsohn[9], Eric-Jan Wagenmakers[10], Jennifer J. Ware[11] 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.

What proportion of published research is likely to be false? Low sample size, small effect sizes, data dredging (also known as *P*-hacking), conflicts of interest, large numbers of scientists working competitively in silos without combining their efforts, and so on, may conspire to dramatically increase the probability that a published finding is incorrect[1]. The field of metascience — the scientific study of science itself — is flourishing and has generated substantial empirical evidence for the existence and prevalence of threats to efficiency in knowledge accumulation (refs 2–7; Fig. 1).

Data from many fields suggests reproducibility is lower than is desirable[8–14]; one analysis estimates that 85% of biomedical research efforts are wasted[14], while 90% of respondents to a recent survey in *Nature* agreed that there is a 'reproducibility crisis'[15]. Whether 'crisis' is the appropriate term to describe the current state or trajectory of science is debatable, but accumulated evidence indicates that there is substantial room for improvement with regard to research practices to maximize the efficiency of the research community's use of the public's financial investment in research.

Here we propose a series of measures that we believe will improve research efficiency and robustness of scientific findings by directly targeting specific threats to reproducible science. We argue for the adoption, evaluation and ongoing improvement of these measures to optimize the pace and efficiency of knowledge accumulation. The measures are organized into the following categories[16]: methods, reporting and dissemination, reproducibility, evaluation and incentives. They are not intended to be exhaustive, but provide a broad, practical and evidence-based set of actions that can be implemented by researchers, institutions, journals and funders. The measures and their current implementation are summarized in Table 1.

The problem

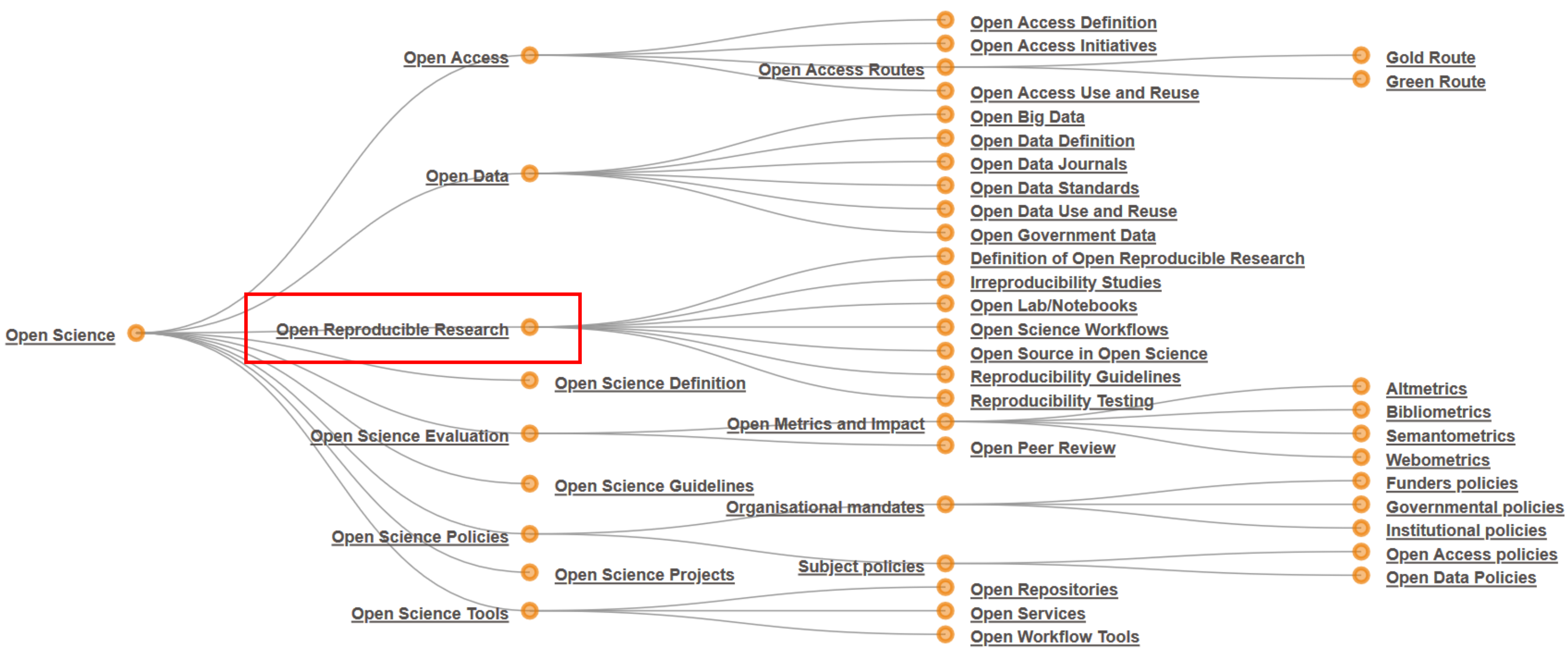
A hallmark of scientific creativity is the ability to see novel and unexpected patterns in data. John Snow's identification of links between cholera and water supply[17], Paul Broca's work on language lateralization[18] and Jocelyn Bell Burnell's discovery of pulsars[19] are examples of breakthroughs achieved by interpreting observations in a new way. However, a major challenge for scientists is to be open to new and important insights while simultaneously avoiding being misled by our tendency to see structure in randomness. The combination of apophenia (the tendency to see patterns in random data), confirmation bias (the tendency to focus on evidence that is in line with our expectations or favoured explanation) and hindsight bias (the tendency to see an event as having been predictable only after it has occurred) can easily lead us to false conclusions[20]. Thomas Levenson documents the example of astronomers who became convinced they had seen the fictitious planet Vulcan because their contemporary theories predicted its existence[21]. Experimenter effects are an example of this kind of bias[22].

Over-interpretation of noise is facilitated by the extent to which data analysis is rapid, flexible and automated[23]. In a high-dimensional dataset, there may be hundreds or thousands of reasonable alternative approaches to analysing the same data[24,25]. For example, in a systematic review of functional magnetic resonance imaging (fMRI) studies, Carp showed that there were almost as many unique analytical pipelines as there were studies[26]. If several thousand potential analytical pipelines can be applied to high-dimensional data, the generation of false-positive findings is highly likely. For example, applying almost 7,000 analytical pipelines to a single fMRI dataset resulted in over 90% of brain voxels showing significant activation in at least one analysis[27].

[1]MRC Integrative Epidemiology Unit, University of Bristol, Bristol BS8 2BN, UK. [2]UK Centre for Tobacco and Alcohol Studies, School of Experimental Psychology, University of Bristol, 12a Priory Road, Bristol BS8 1TU, UK. [3]Department of Psychology, University of Virginia, Charlottesville, Virginia 22904, USA. [4]Center for Open Science, Charlottesville, Virginia 22903, USA. [5]Department of Experimental Psychology, University of Oxford, 9 South Parks Road, Oxford OX1 3UD, UK. [6]Department of Psychology, University of Bath, Bath BS2 7AY, UK. [7]Cardiff University Brain Research Imaging Centre, School of Psychology, Cardiff University, Cardiff CF24 4HQ, UK. [8]National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs), London NW1 2BE, UK. [9]The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [10]Department of Psychology, University of Amsterdam, Amsterdam 1018 WT, Netherlands. [11]CHDI Management/CHDI Foundation, New York, New York 10001, USA. [12]Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford 94304, California, USA. [13]Stanford Prevention Research Center, Department of Medicine and Department of Health Research and Policy, Stanford University School of Medicine, Stanford 94305, California, USA. [14]Department of Statistics, Stanford University School of Humanities and Sciences, Stanford 94305, California, USA.
*e-mail: marcus.munafo@bristol.ac.uk

Munafò et al. (2017)

What's Open Science/Research/Scholarship *as practice*?



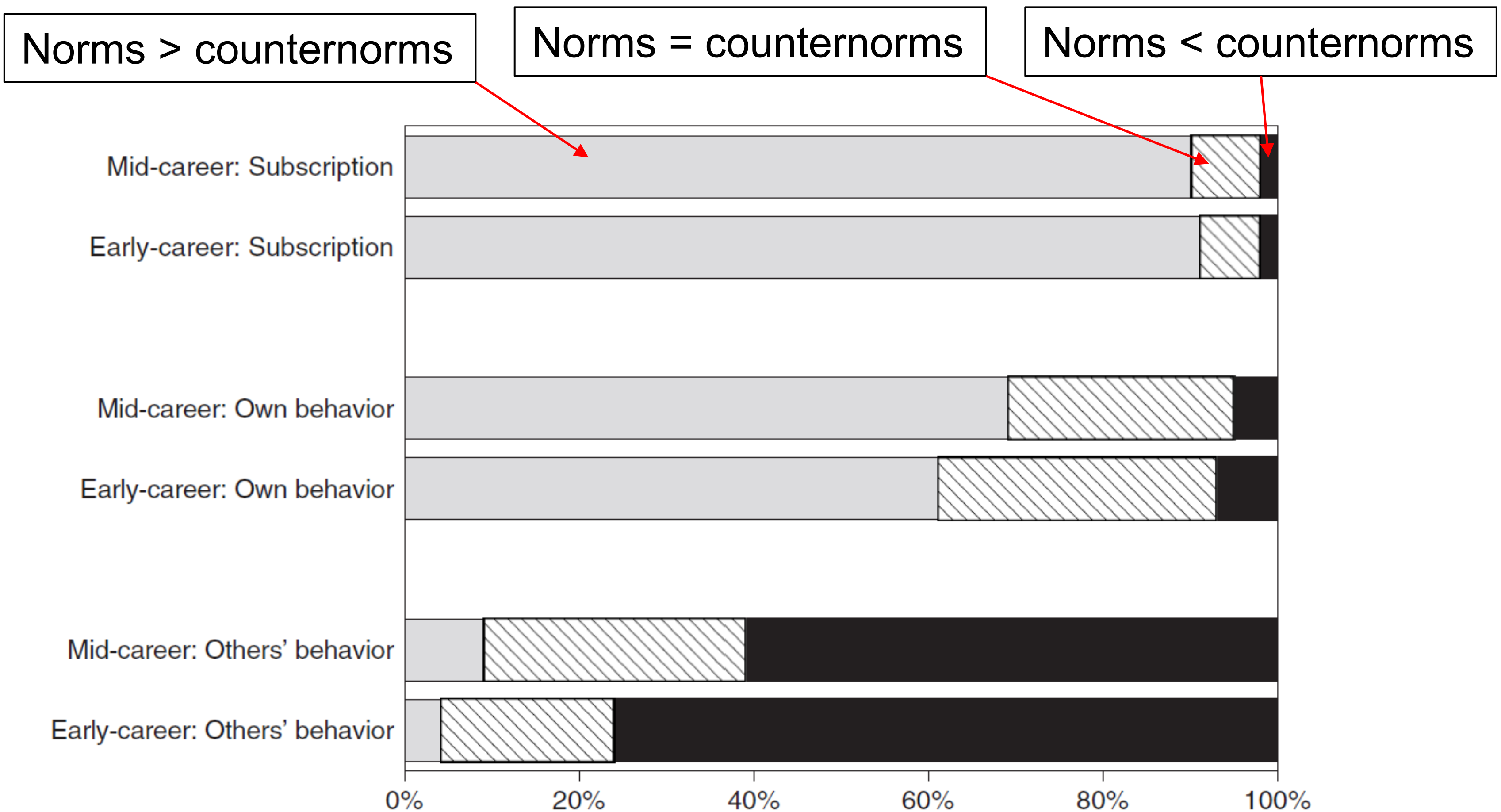
FOSTER Plus

Academic norms and counternorms

Communality: Scientists openly share findings with colleagues.	*Secrecy*: Scientists protect their newest findings to ensure priority in publishing, patenting, or applications.
Universalism: Scientists evaluate research only on its merit, i.e., according to accepted standards of the field.	*Particularism*: Scientists assess new knowledge and its applications based on the reputation and past productivity of the individual or research group.
Disinterestedness: Scientists are motivated by the desire for knowledge and discovery, and not by the possibility of personal gain.	*Self-Interestedness*: Scientists compete with others in the same field for funding and recognition of their achievements.
Organized Skepticism: Scientists consider all new evidence, hypotheses, theories, and innovations, even those that challenge or contradict their own work.	*Organized Dogmatism*: Scientists invest their careers in promoting their own most important findings, theories, or innovations.
Governance: Scientists are responsible for the direction and control of science through governance, self-regulation and peer review.	*Administration*: Scientists rely on administrators to direct the scientific enterprise through management decisions.
Quality: Scientists judge each others' contributions to science primarily on the basis of quality.	*Quantity*: Scientists assess each others' work primarily on the basis of numbers of publications and grants.

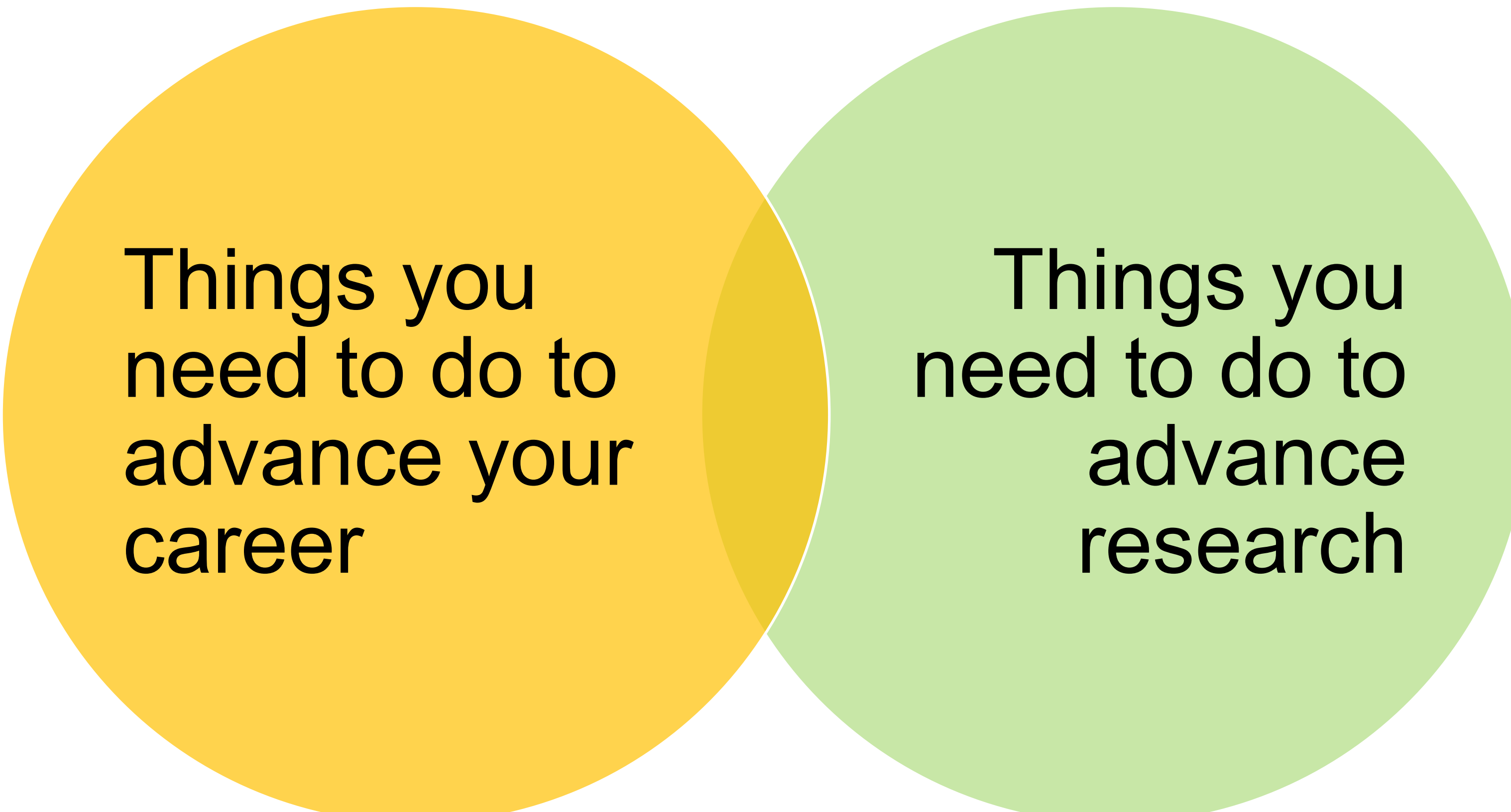
Anderson et al. (2007)

Ideals and reality



Anderson et al. (2007)

Theme 1: Open research is about normative dissonance

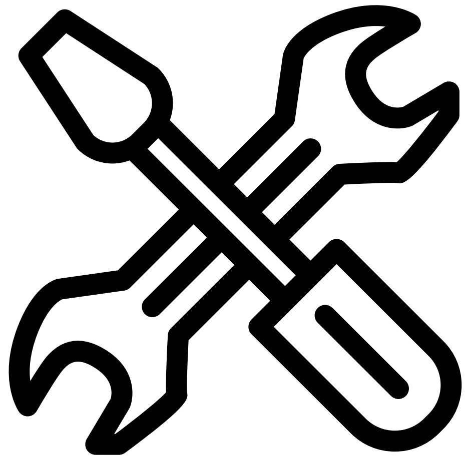


Being open

- Towards yourself
- Towards the research community
- Towards society

Theme 2: Meta-research helps us assess the dissonance

Theme 3: Open research offers means to respond to it



UNIVERSITY OF LEEDS

Meta-research: Assessing the credibility of social research

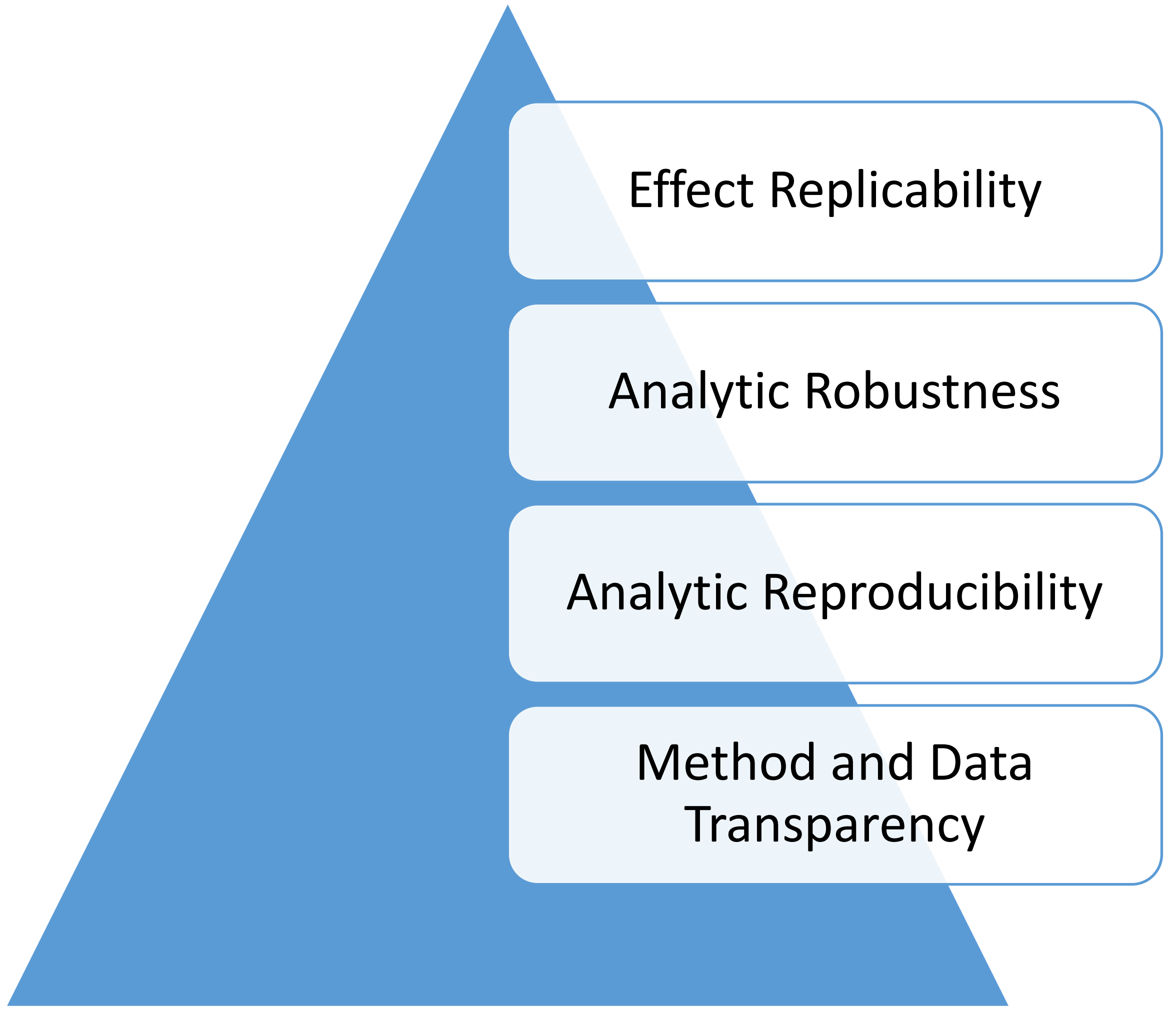
UNIVERSITY OF LEEDS

A credible finding or hypothesis is one that has repeatedly survived **high-quality, risky attempts at proving it wrong**.

LeBel et al. 2018

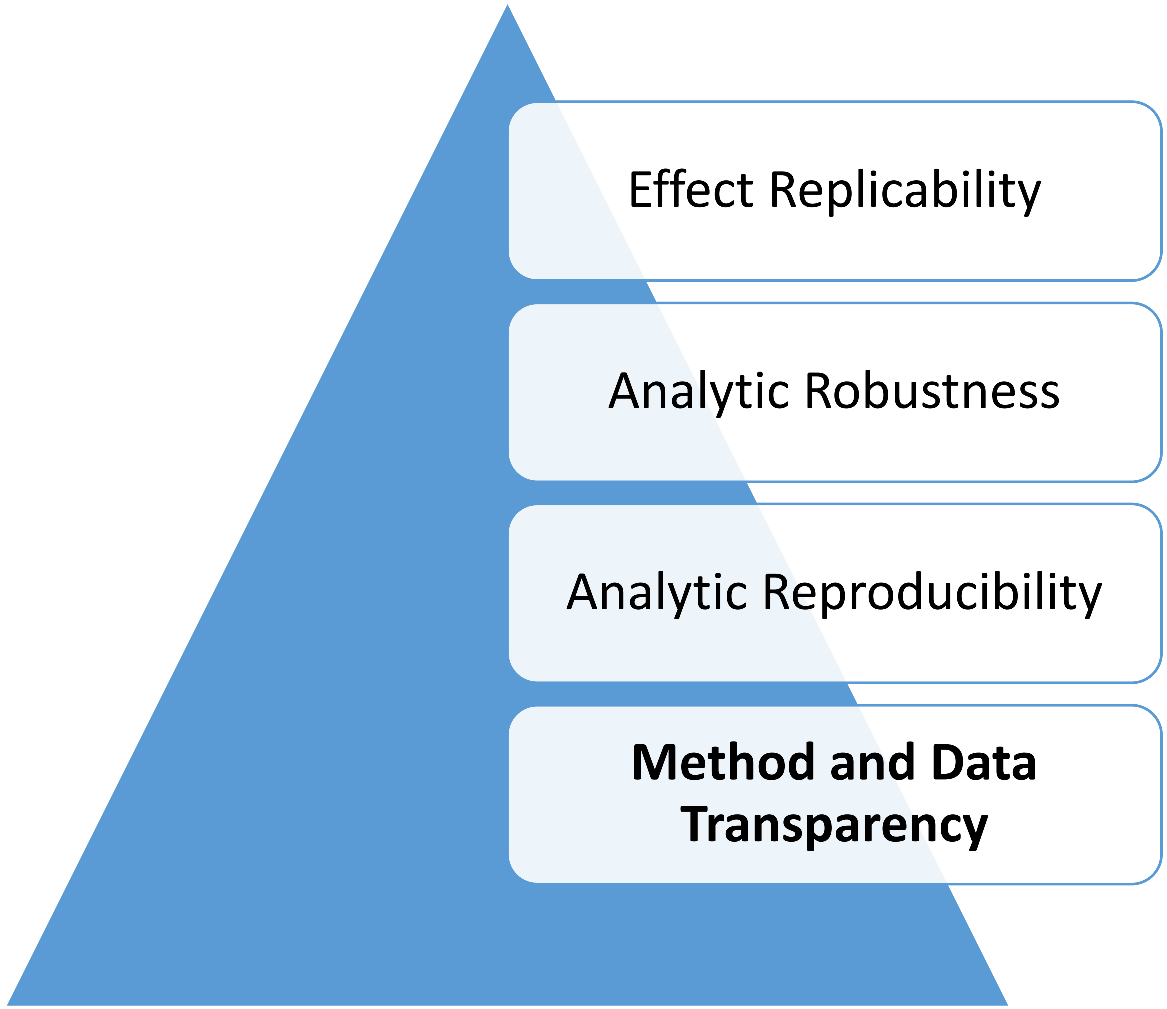
UNIVERSITY OF LEEDS

Criteria for research credibility



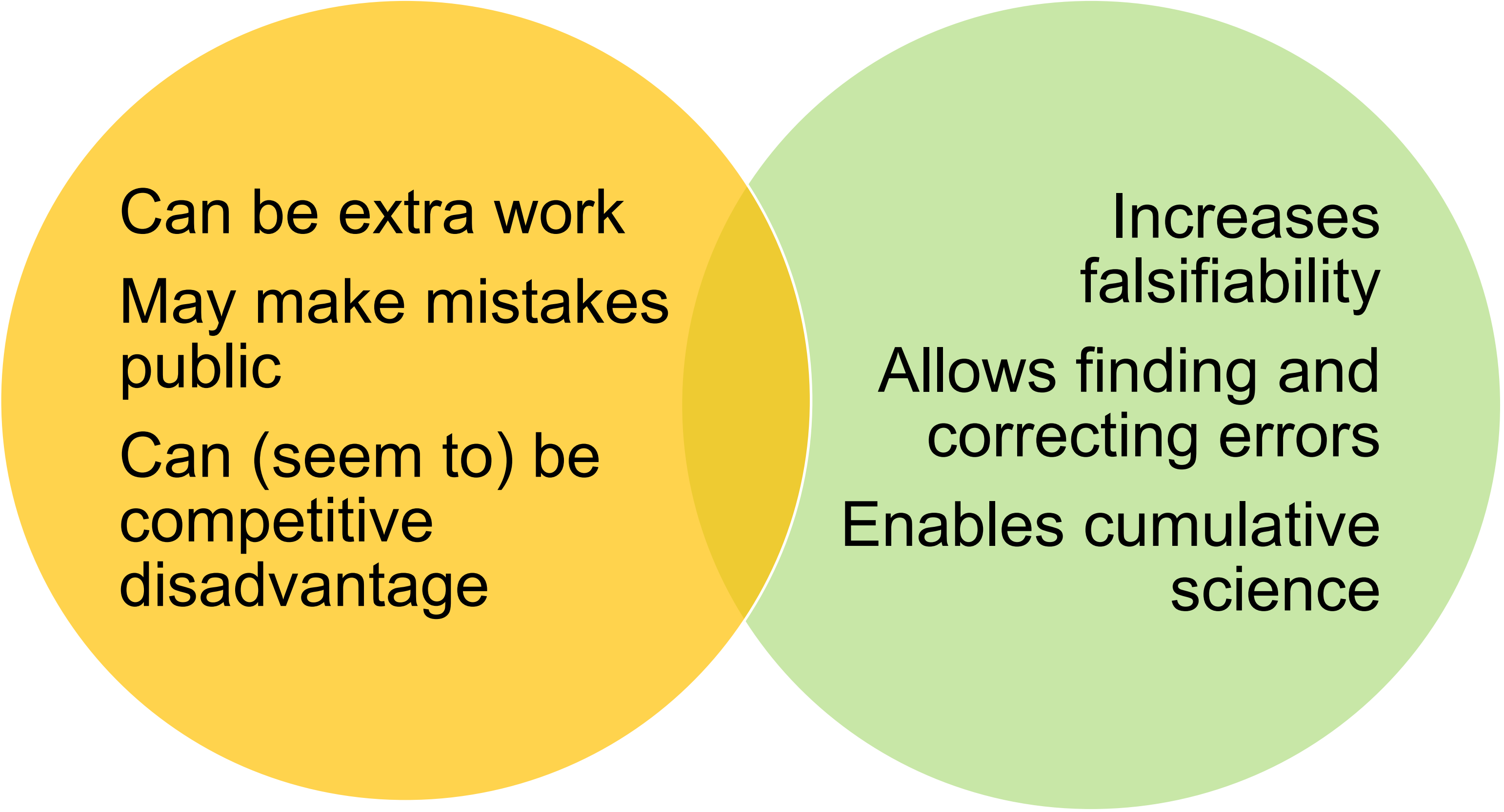
LeBel et al. 2018

Criteria for research credibility



[LeBel et al. 2018]

Method and Data Transparency

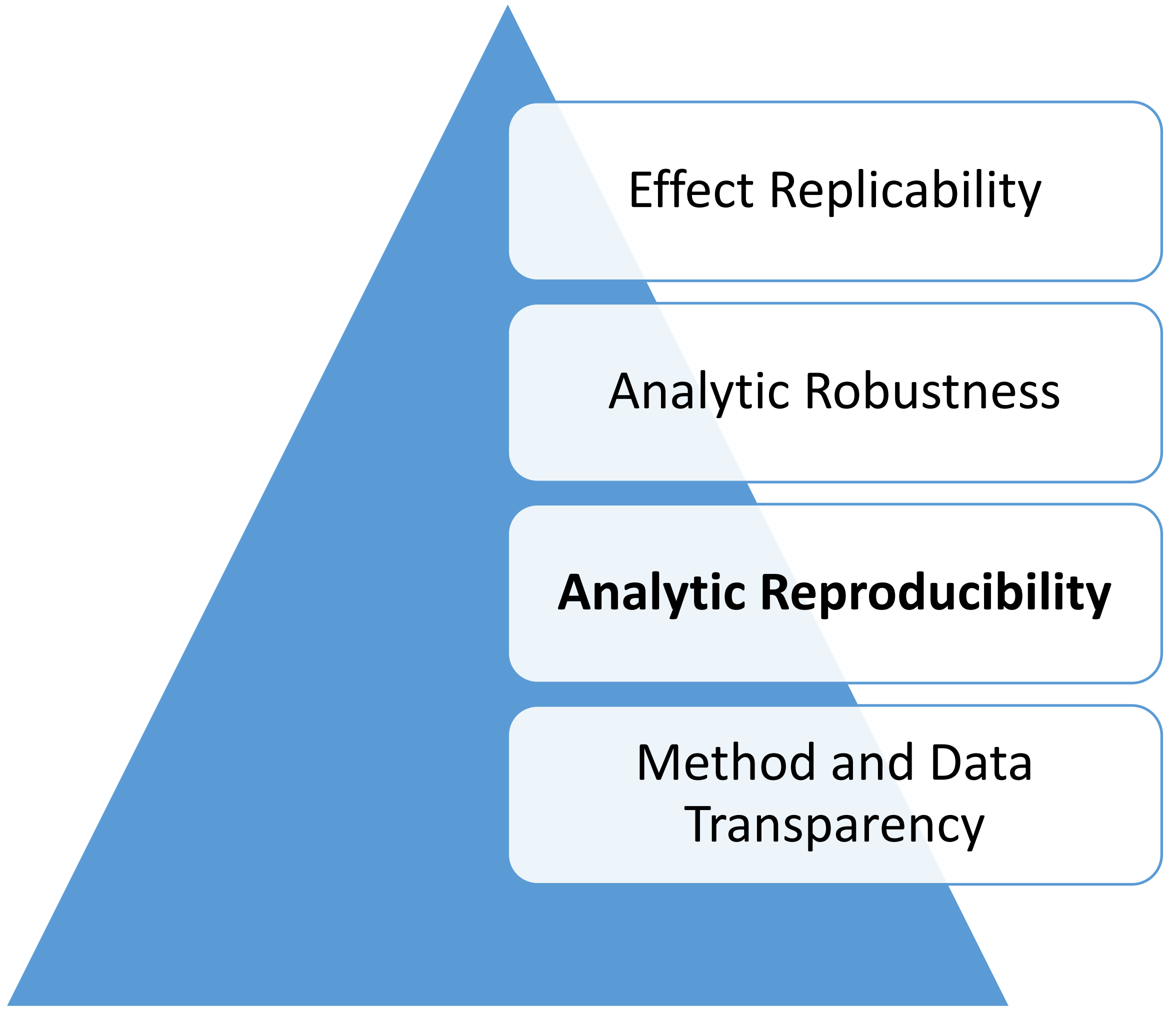


Method and data transparency: What does meta-research say?

- In politics and IR, 21% of all statistical inference papers published in 2020/21 have open data, 5% of all experiments are preregistered. — Scoggins & Robertson (2023)
- 27% of all data links in *APSR* in 2013 were dead in 2014. — Gertler/Bullock (2017)
- Data is not available upon request: Data *actually* shared for 17% of articles with "upon request" data-availability statement — Hussey (2023)

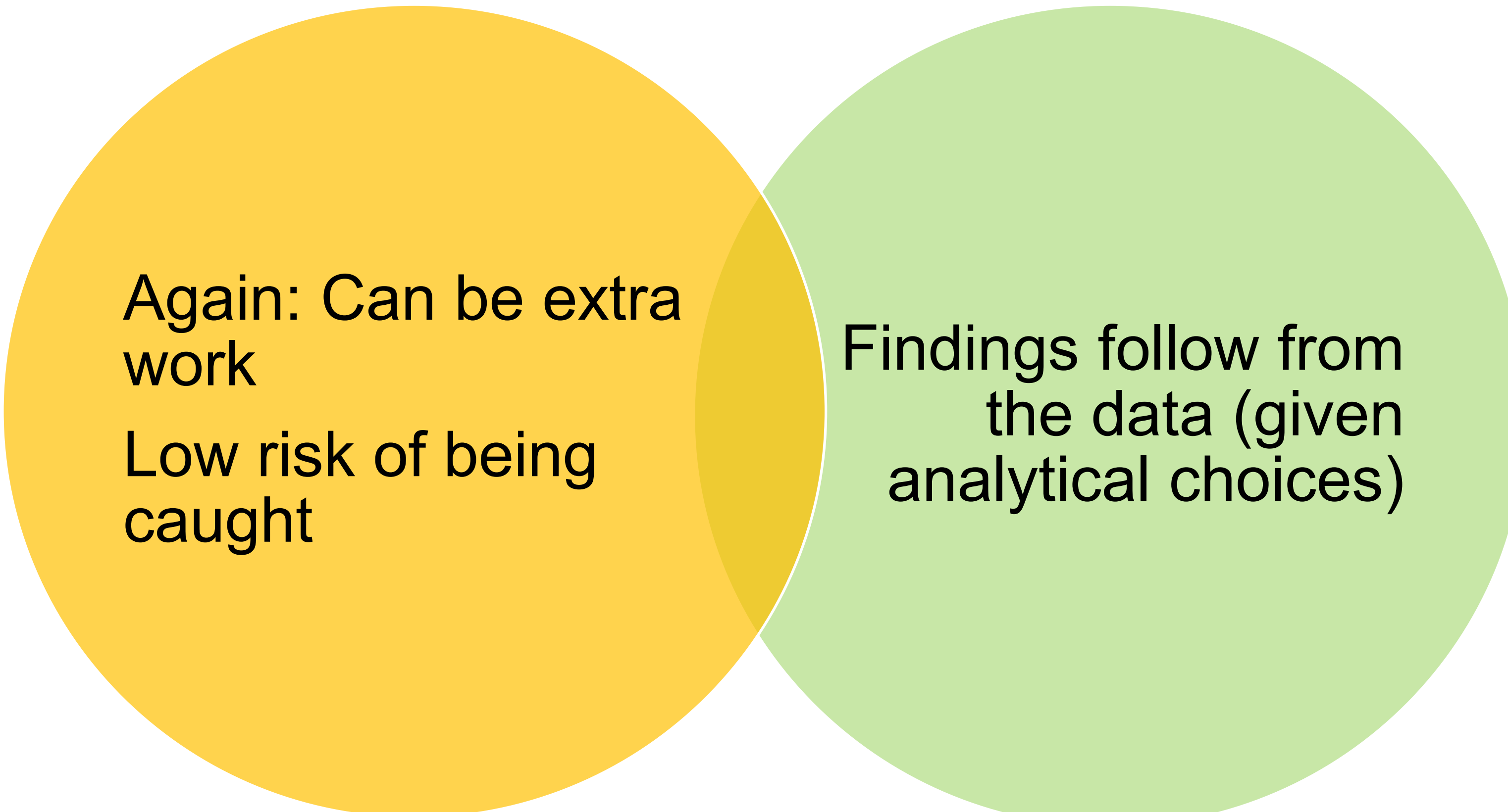
But: Mandatory data and method transparency slowly becoming standard practice

Criteria for research credibility



LeBel et al. 2018

Analytic Reproducibility



Analytic Reproducibility

- 74% of scripts posted on Dataverse (2010-20) failed to complete without error, 56% failed when code cleaning was applied — *Trisovic et al. (2022)*
- 58% of papers in QJPS reported results that differed from those generated by author's own code — *Eubank (2016)*

But: More tools (e.g., SocSci Reproduction Platform) and stricter journal procedures around reproduction

Criteria for research credibility

Effect Replicability

Analytic Robustness

Analytic Reproducibility

Method and Data Transparency

[LeBel et al. 2018]

Analytic Robustness & Replicability

Academic currency: publication

Known publication biases

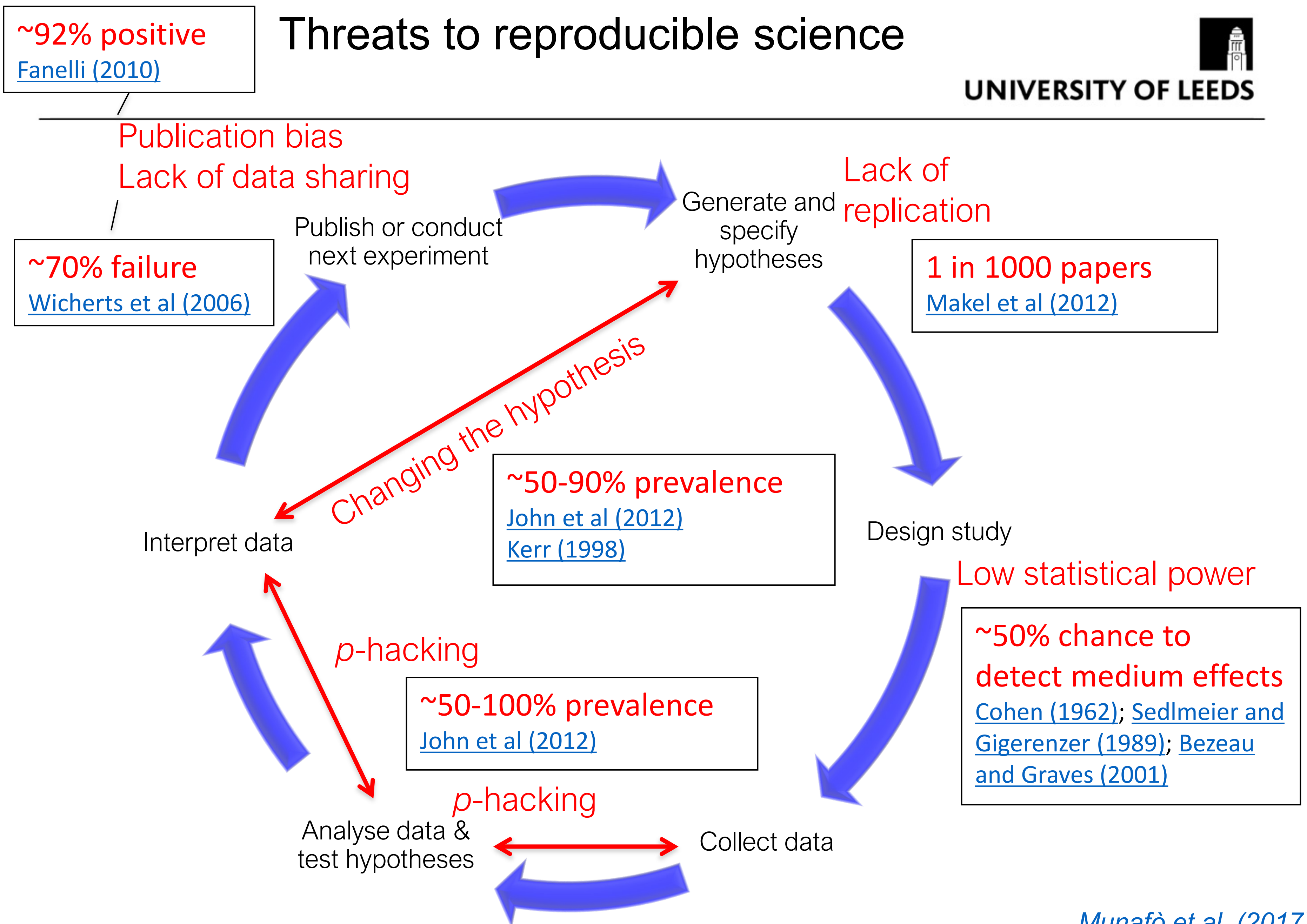
Researcher degrees of freedom

More robust findings:

- Less sensitive to seemingly minor analytical decisions
- more replicable
- greater predictive power

Threats to reproducible science

UNIVERSITY OF LEEDS



Munafò et al. (2017)

Publication bias

Findings across studies:
Representative of entire evidence base generated by published and unpublished studies?

Finding

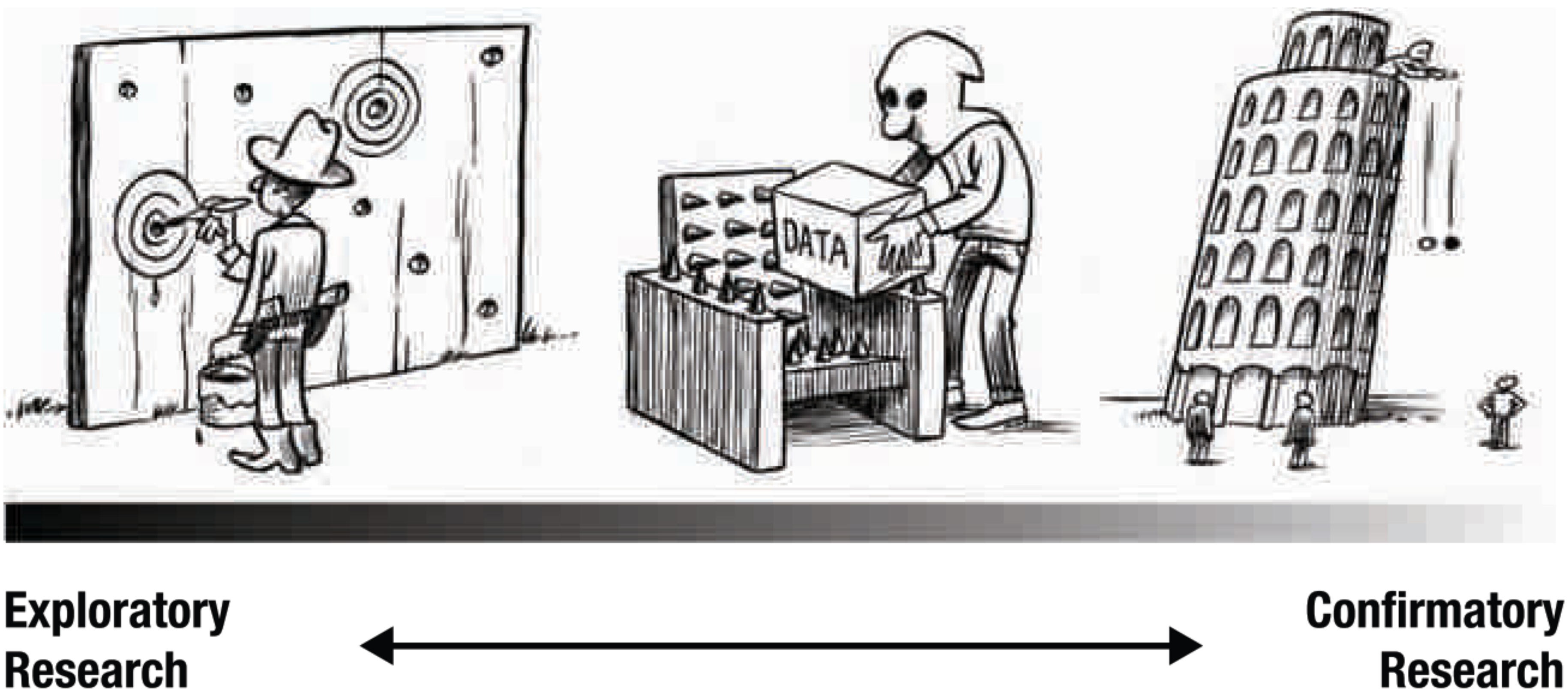
QRPs

Findings within studies:
Representative of entire evidence base generated by this study?

Low credibility of research

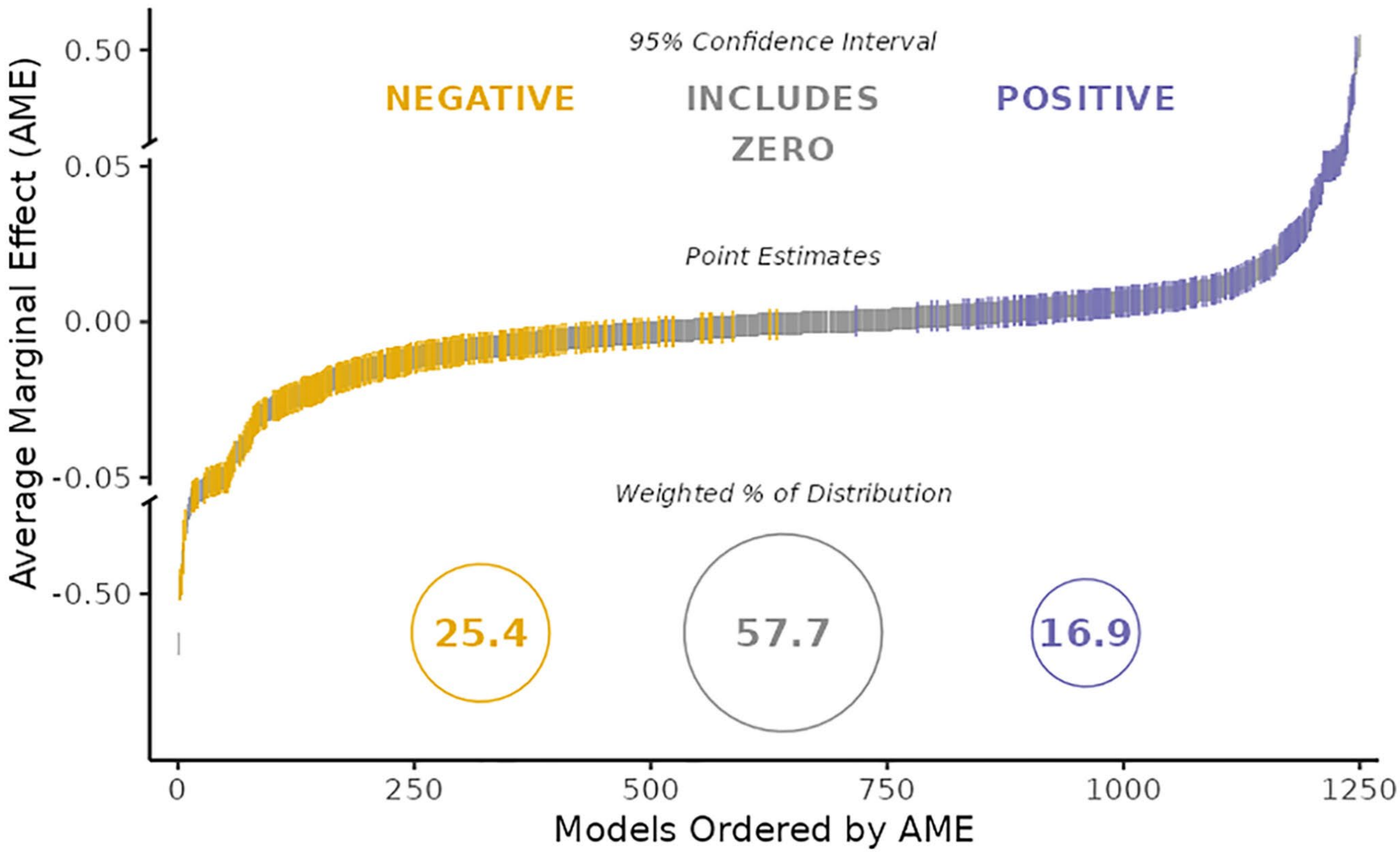
UNIVERSITY OF LEEDS

HARKing, *p*-Hacking and Hypothesis-Testing Research



[Hoek (in Wagenmakers et al. 2012)]

Analytical Robustness



Breznau, Rinke & Wuttke et al. (2022)

Analytical Robustness

UNIVERSITY OF LEEDS

PNAS COMMENTARY

Check for updates

The apparent prevalence of outcome variation from hidden "dark methods" is a challenge for social science

Colin F. Camerer[a,1]

Every working scientist knows that in the details are both devils and angels. Lots of small design decisions have to be made in collecting and analyzing data, and those decisions affect conclusions. But beginning scientists, from rookies in school science fairs to students in early years of a rigorous Ph.D. program, are often surprised how much small decisions matter. Despite this recognition that details matter, when science is communicated, many small decisions made privately by a science team are hidden from view. It is difficult to disclose every detail (and usually little disclosure is required). Such hidden decisions can be thought of as "dark methods," like dark matter which cannot be directly seen because it does not reflect light, but which is evident from its other effects. The Herculean effort resulting in the new many-analyst study (1) which is the subject of my Commentary should force a painful reckoning about the extent of these dark method choices and their influence on conclusions. Design decisions of each team that were coded (107 of them) explained at most 10 to 20% of the outcome variance. Assuming that the coding itself is not too noisy, it seems that hidden decisions account for the lion's share of what different teams conclude.

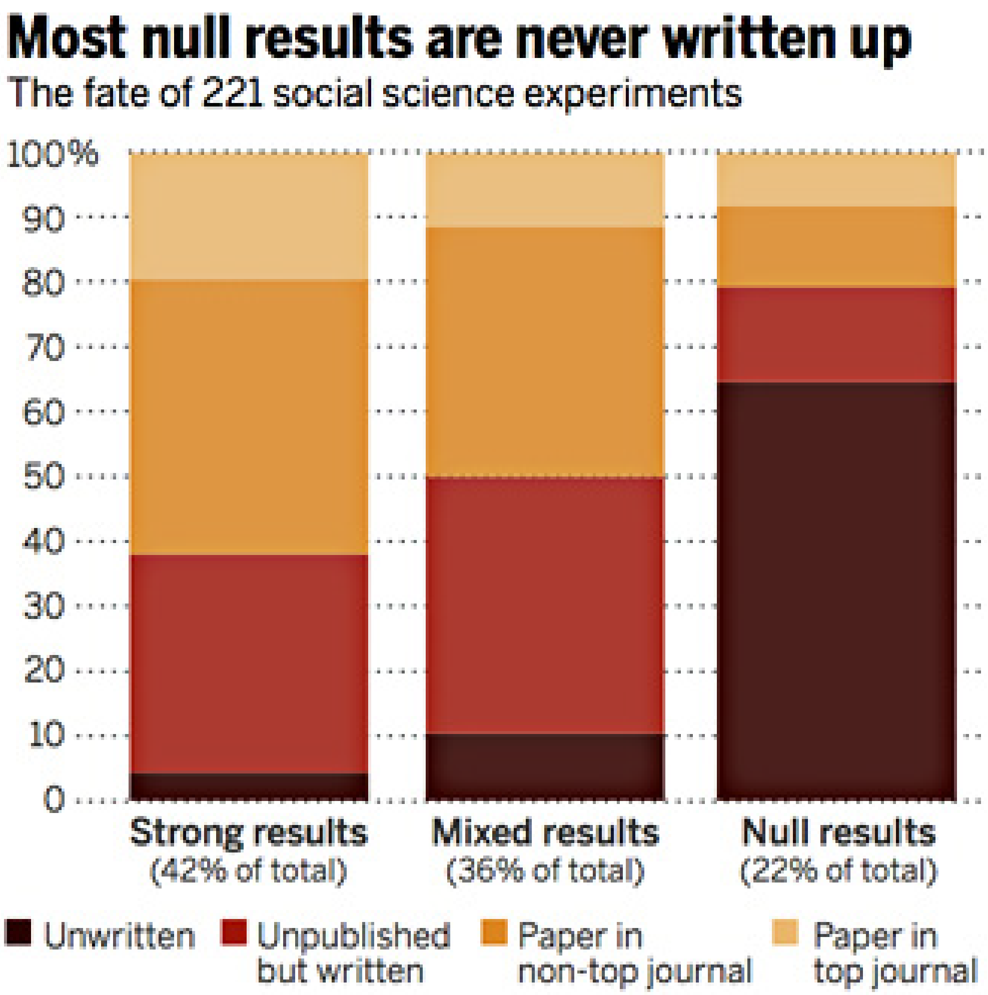
In ref. 1, they recruited 73 teams to test the hypothesis that "immigration reduces public support for government provision of social policies." Whether this hypothesis is true is obviously an important question, especially now and very likely in the world's future as well. The hypothesis is pendent variables, subsets of data, etc. The contribution of these coded decisions explained only a little more than 10% of variance in results between teams.

The authors conclude that even when trying to carefully code these design decisions (in order specifically to shed light on typically dark methods), the coded variables do not explain much. Eighty Percent of the variance in team-reported results is due to some other variables that are not coded. Fig. 1*A* illustrates both variability in team outcomes and the weak relation between high-level design features and those outcomes.

The challenges posed by the surprising influence of dark methods come after almost two decades of other questions about how well current practices cumulate scientific regularity (2). Social scientists—as well as those in other fields, especially medicine—are now well aware of the feared and actual impacts of p-hacking, selective inference, and both scientist-driven and editorial publication bias. A small wave of direct replications in psychology, economics, and in general science journals, intended to reproduce previous experimental protocols as closely as possible, typically found that many or most results do not replicate strongly (3–5). (My rule of thumb is that the long-run effect size of a genuine discovery will be at 2/3 as large as the original effect). But most social sciences have also turned toward self-correction, albeit at the slow pace of turning a large oil tanker rather than a sports car. Preregistration, journal

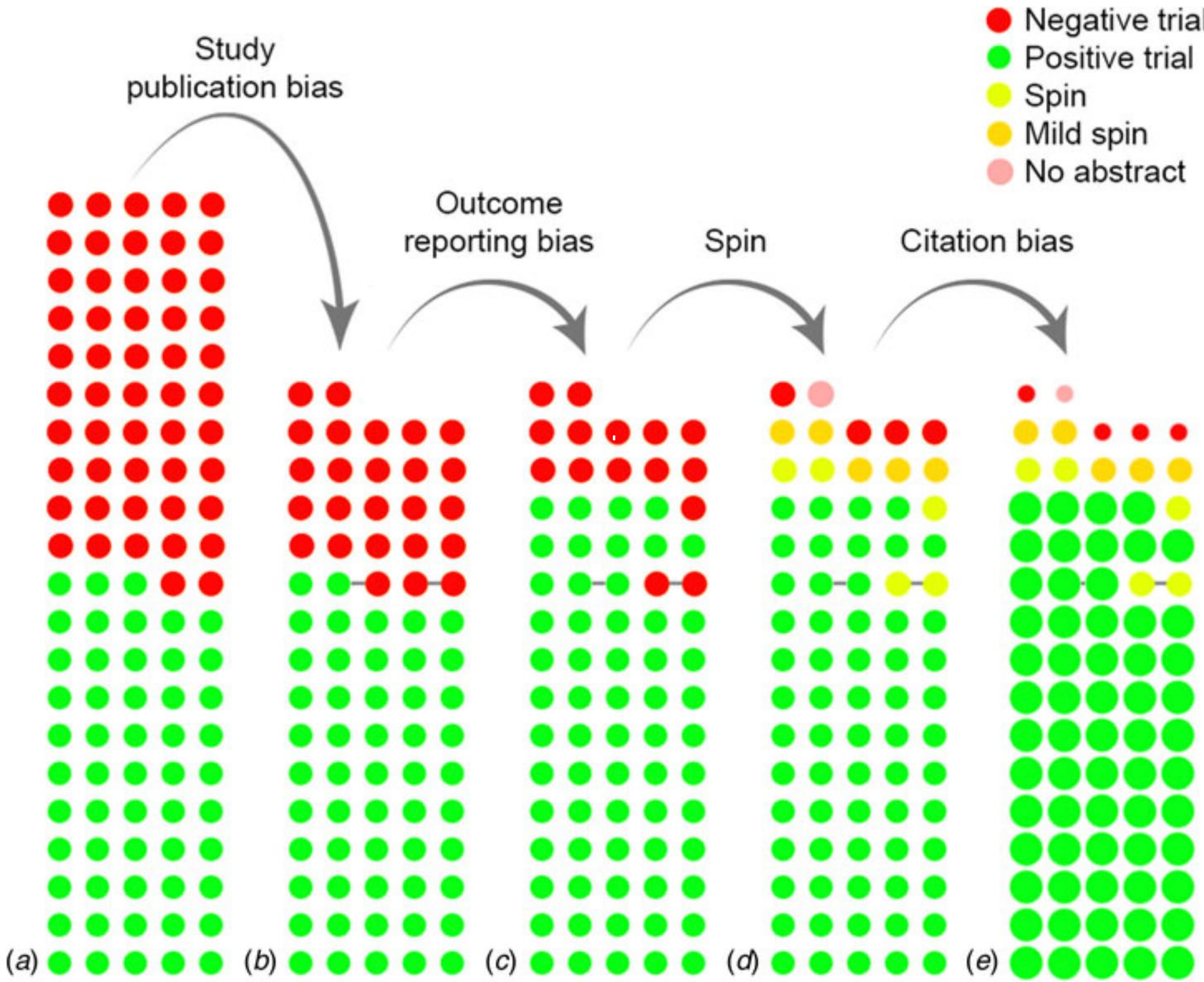
Camerer (2022)

Publication Bias… … is (was?) rampant.



[Franco et al. 2014]

Publication Bias and Selective Reporting



[de Vries et al. (2018)]

Replicability of Experimental Studies

UNIVERSITY OF LEEDS

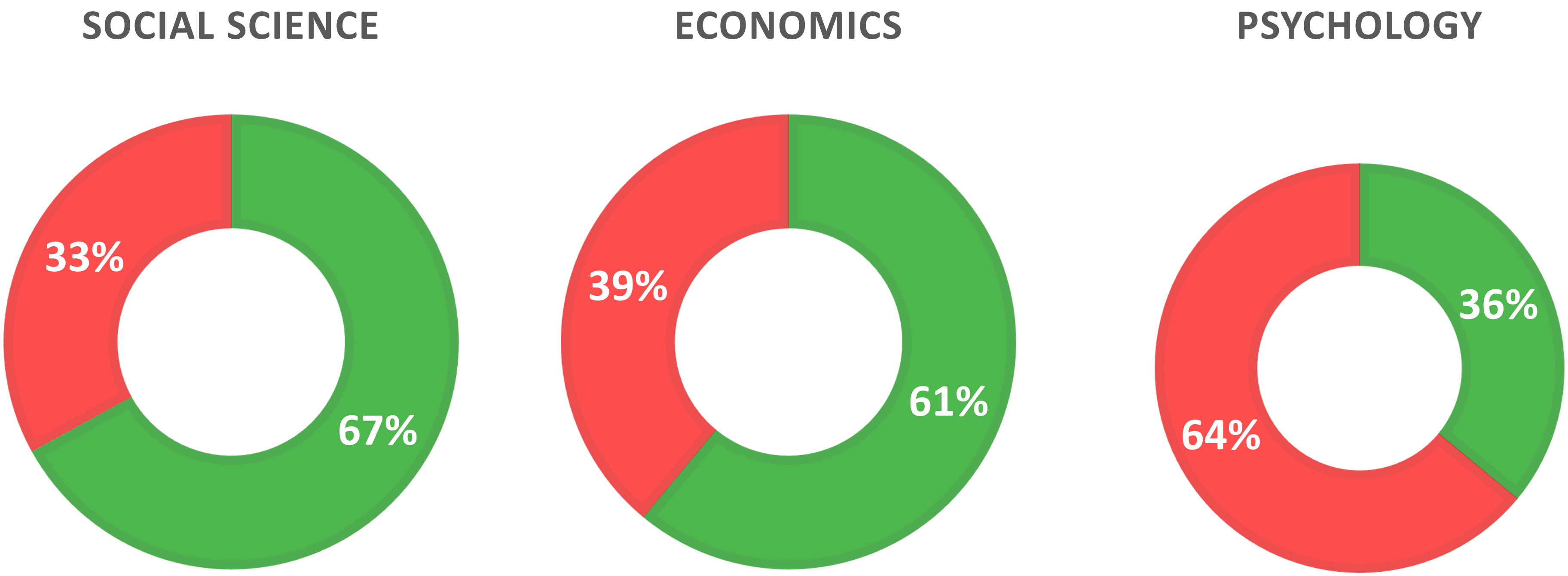


Figure based on:
[Camerer et al. (2018)]

A Solution? Pre-registration & Results-blind Peer Review

Center for Open Science

UNIVERSITY OF LEEDS

THE UNIVERSITY OF CHICAGO PRESS JOURNALS

This journal | Search | Advanced Search | Log In | Register

Subscriber/member | Browse issues | Forthcoming | Contributors | About | SPSA

The Journal of Politics

Editor in Chief: Vera Troeger
Sponsored by the Southern Political Science Association

SPSA SOUTHERN POLITICAL SCIENCE ASSOCIATION

Registered Report Guidelines

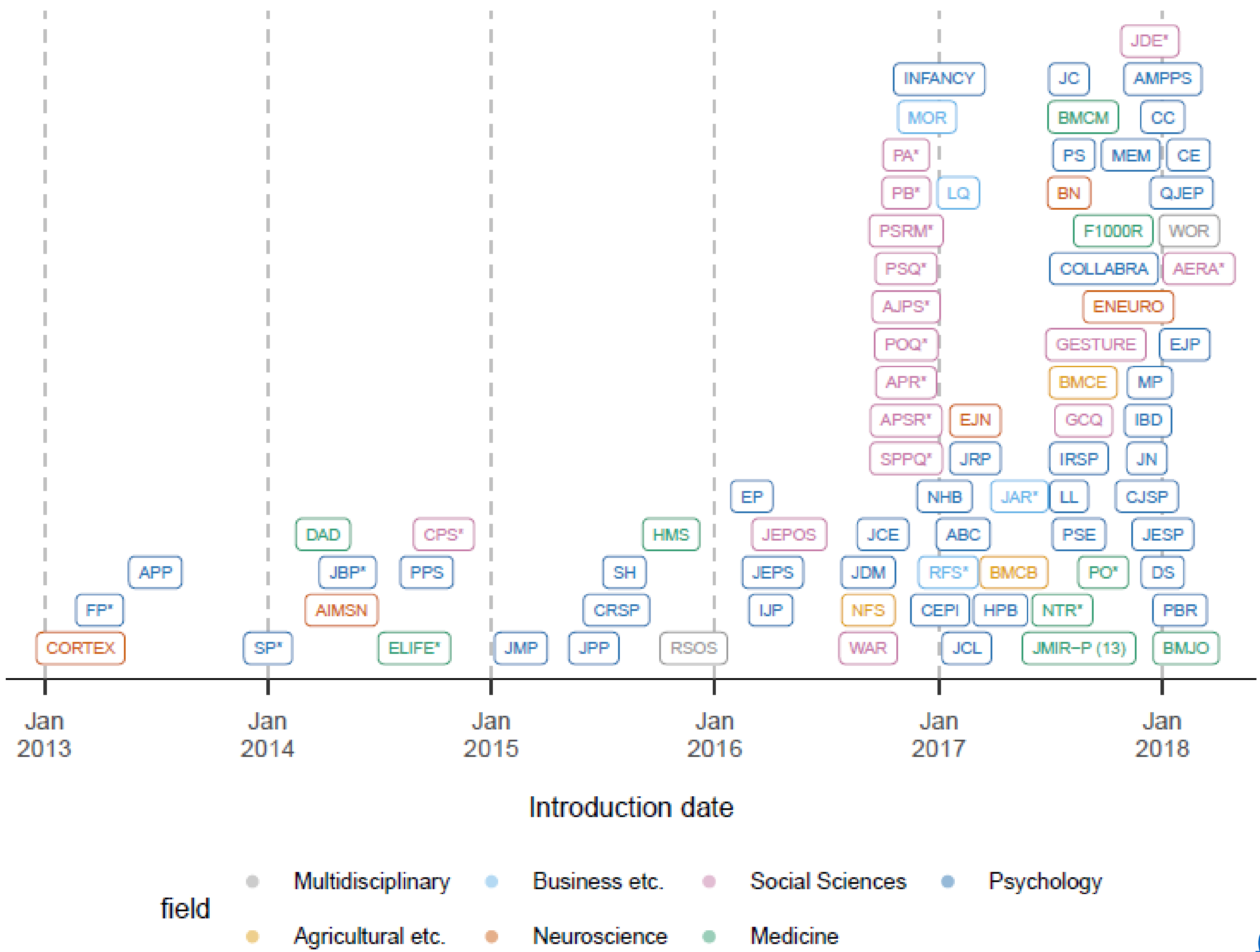
This page contains information on Registered Reports. Please use the following list of contents to find the information you require.

Source

Registered Reports

UNIVERSITY OF LEEDS

Journals offering Registered Reports

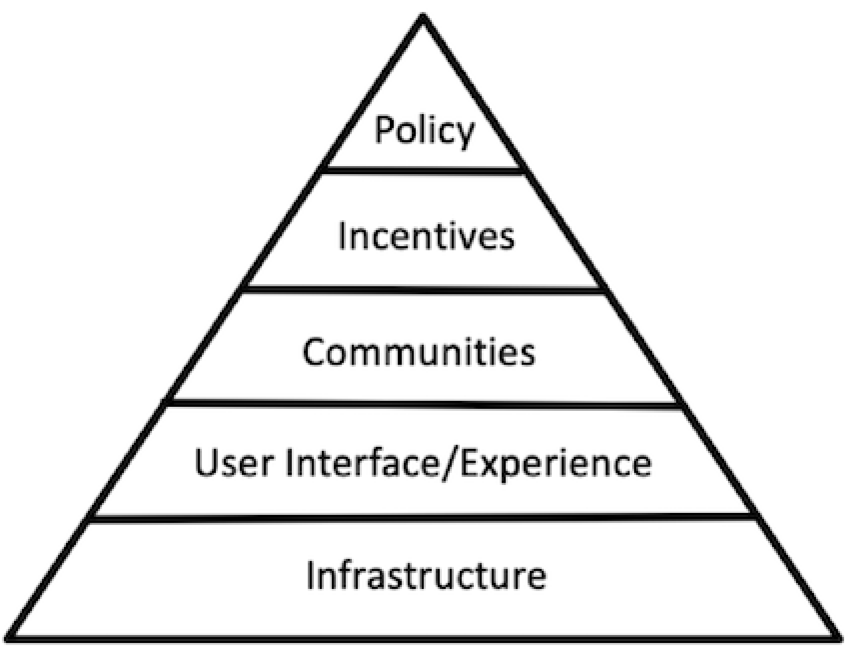


[Hardwicke & Ioannidis (2018, Appendix B)]

Competing considerations regarding research transparency

- personal intellectual considerations (confidence in own research, helping future self etc.)
- public intellectual considerations (showcasing rigor and power of research)
- resource considerations and opportunity costs (time and money)
- logistical considerations (practical possibilities)

Strategy for Culture Change



Make it required

Make it rewarding

Make it normative

Make it easy

Make it possible

[Center for Open Science]

Some *general* open research practices

1. Collaborate
2. Foster open science skills
3. Incentivize open research practices
4. Preregister studies and submit registered reports
5. Publish materials, data, and code
6. Methodological appendices
7. Adopt reproducible workflows
8. Adopt open reporting standards

Some open *qualitative* research practices

1. Pre-registration
2. Methodological Appendices
3. Annotation (Software Assisted)
4. QDA Software Output
5. …

Some open *quantitative* research practices

1. Conduct replication studies

2. Implement Transparency and Openness Promotion (TOP) Guidelines

3. …

Some learning resources

1. Foster Open Science E-Learning Courses: https://www.fosteropenscience.eu/courses
2. Open Science MOOC: https://opensciencemooc.eu/
3. Transparent and Open Social Science Research MOOC: https://www.bitss.org/education/mooc-parent-page/

OPEN SCIENCE: JUST SCIENCE DONE RIGHT

Imming & Tennant (2018)