



An introduction to Open Science

Doing transparent, credible and efficient research

SCIENCE

The **smartest heads** in the world immerse themselves into a research topic for years.

In that process, they become the experts – nobody knows more about that topic. The boundaries of knowledge have been **pushed forward**.

When the scientist are confident in their findings, they publish them in the best scientific journals, with the highest standards of **quality, rigor, and integrity**.

**HOW MUCH OF THAT LITERATURE
DO YOU THINK IS TRUE?**

How to be successful in academia

aka.

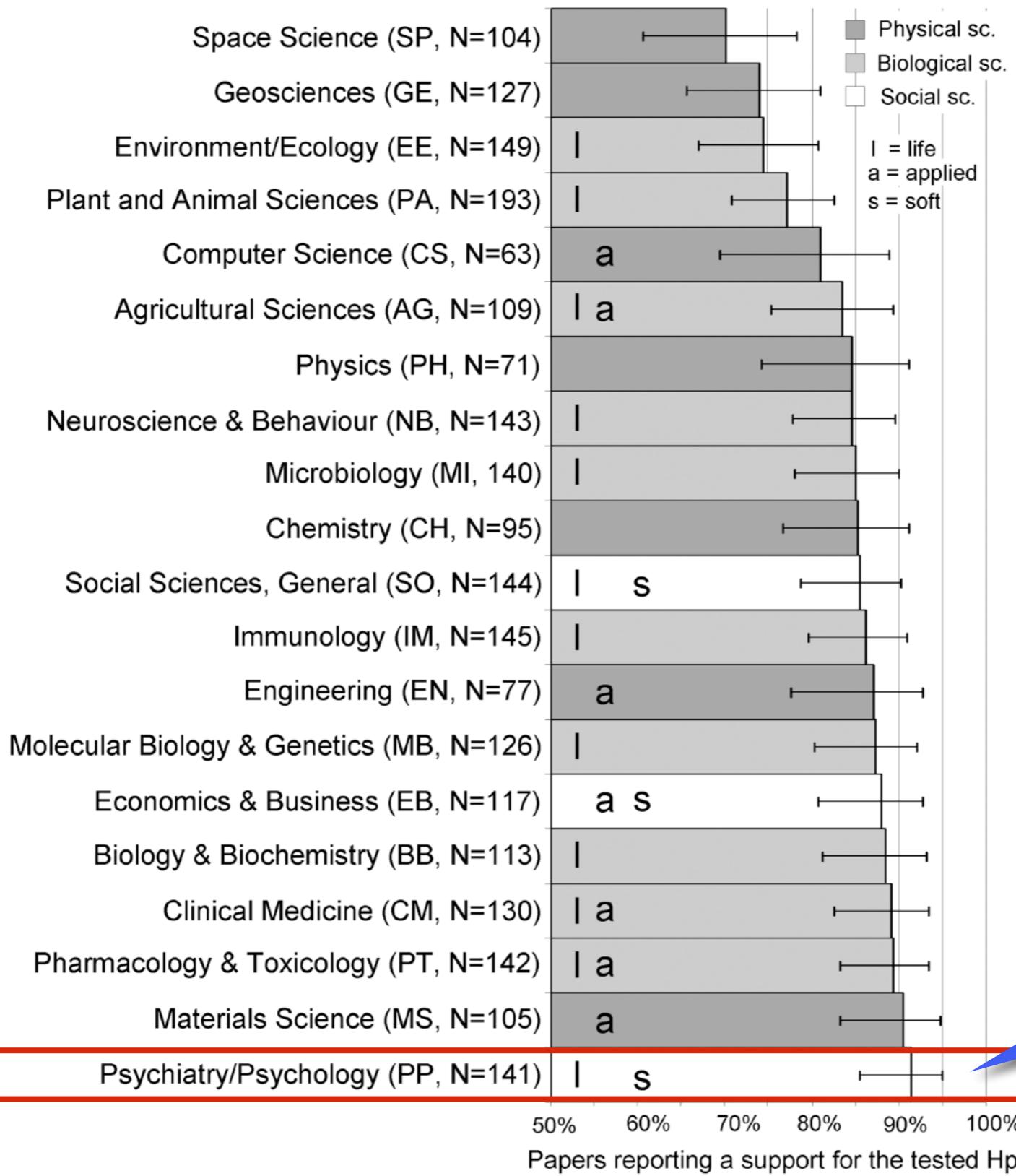
Hack your way to scientific glory

How to become a Professor?

Actual (not desired) relevance in professorship hiring committees	Rank
Number of peer-reviewed publications	1
Fit of research profile to the hiring department	2
Quality of research talk	3
Number of publications	4
Volume of acquired third-party funding	5
Number of first authorships	6
...	...

N = 1453 psychology researchers, 66% were actually members of a professorship hiring committee.

How to get lots of publications?



92% of published
papers have
significant,
positive results

Learn to *p*-hack like the pros!



***p*-hacking** (*n.*). Tune your data analysis in a way that you achieve a significant *p*-value in situations where it would have been non-significant.

Questionable research practices (QRPs) (*n.*). Practices of data collection and data analysis that are not outright fraud, but also not really kosher.

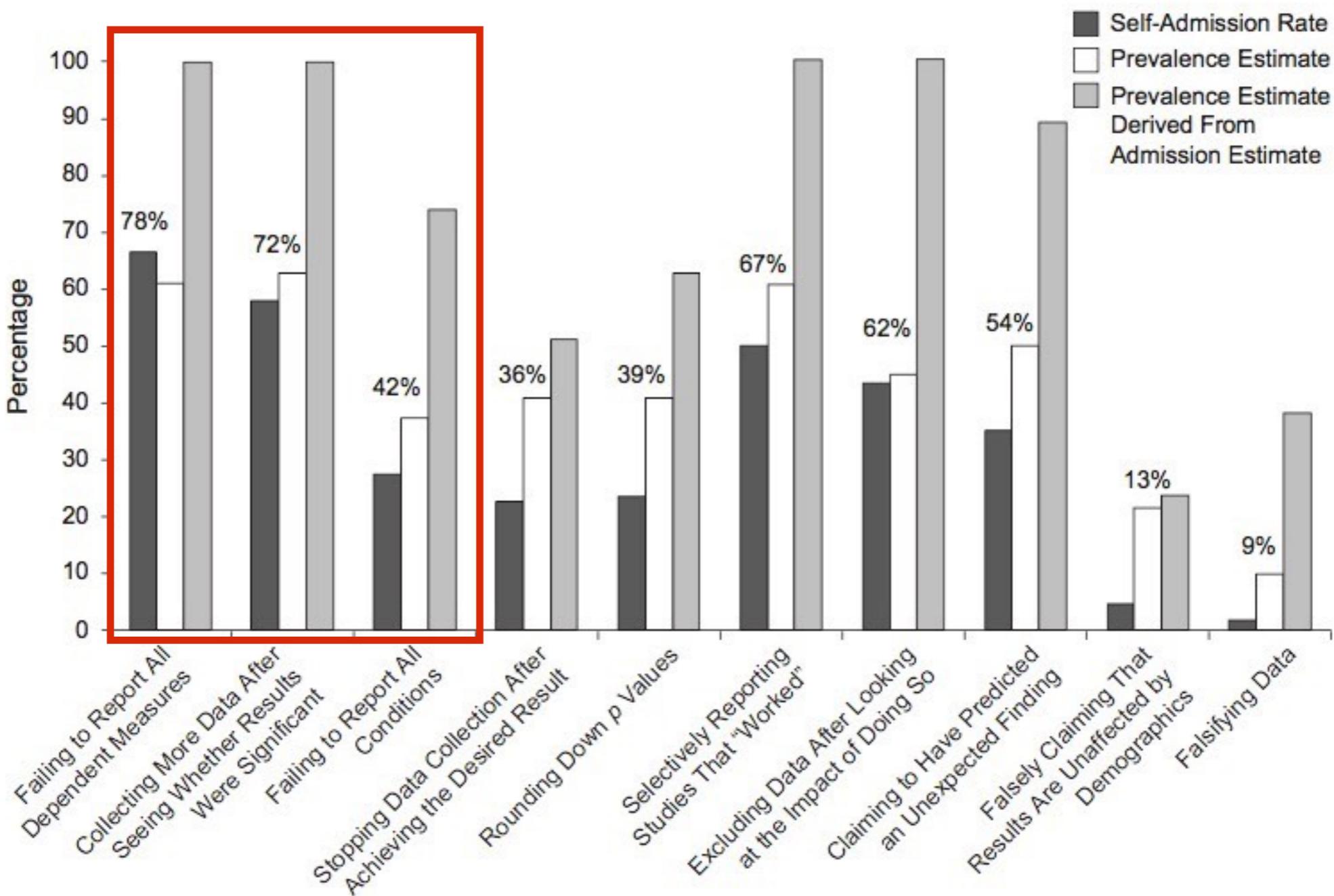
→ Learn how to *p*-hack in the next lecture!

Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling *IN PSYCHOLOGY*

Psychological Science
23(5) 524–532
© The Author(s) 2012
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797611430953
<http://pss.sagepub.com>


Leslie K. John¹, George Loewenstein², and Drazen Prelec³

¹Marketing Unit, Harvard Business School; ²Department of Social & Decision Sciences, Carnegie Mellon University; and ³Sloan School of Management and Departments of Economics and Brain & Cognitive Sciences, Massachusetts Institute of Technology



Survey among 6,813 academic researchers in The Netherlands: Self-reported prevalence of fabrication and falsification in the last 3 years

		Disciplinary field			
QRP	Description (In the last three years.)	Life and medical sciences	Social and behavioural sciences	Natural and engineering sciences	Arts and humanities
Fabrication	Making up of data or results				
Falsification	Manipulating research materials, data or results				
Any FF	Fabrication and/or Falsification				

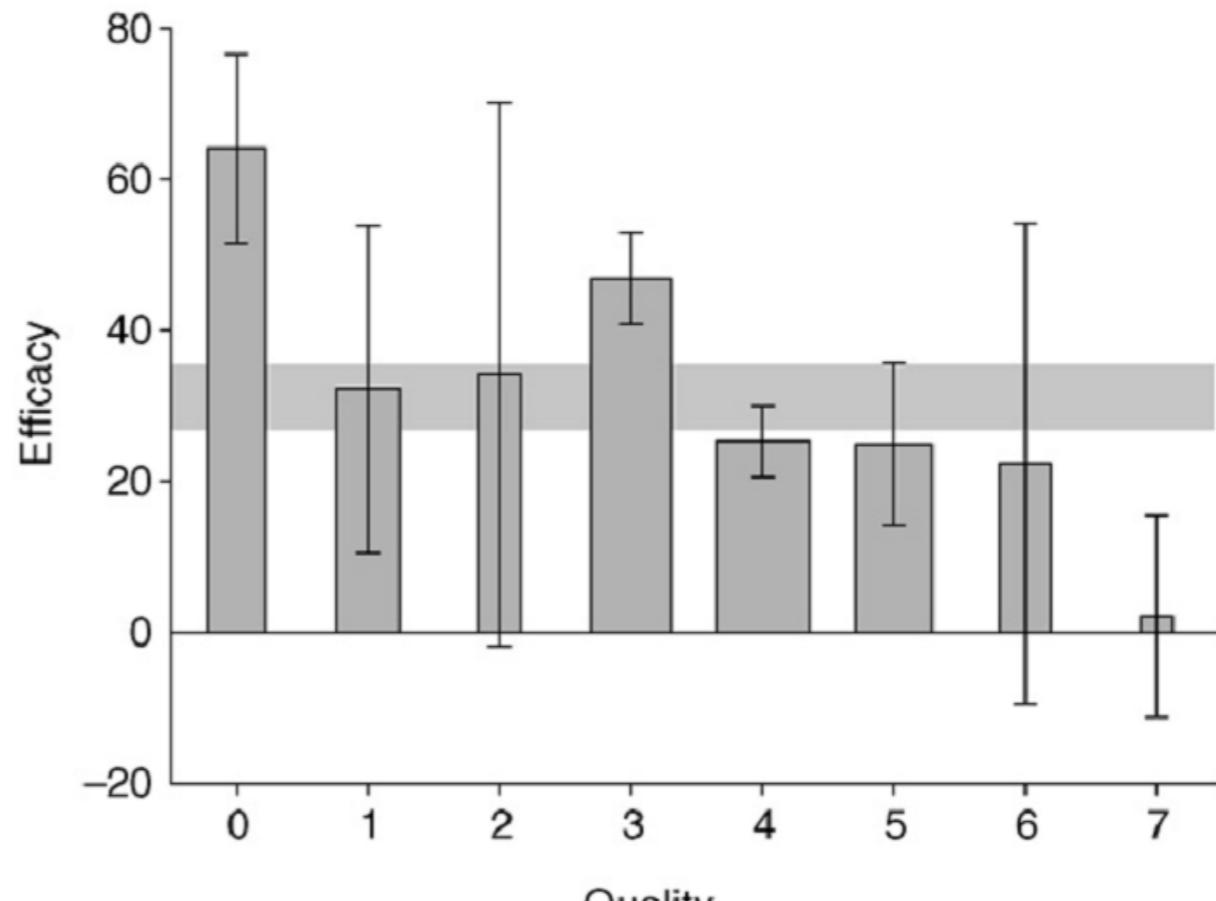
(Un)Intentional?

- Intentional?
 - Evil researcher who only cares about his/her career and not at all about truth-seeking?
- Unintentional?
 - Wrong education?
 - Wrong/uncritical standards of the field?
 - Pushed by supervisors, reviewers, or editors?
 - <http://bulliedintobadscience.org/>
 - Distorting effects on the published record are probably comparable, but the ethical evaluations differs strongly.

The higher the study quality, the smaller the effect

Publication	Year	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	Score
Sharkey	1994	/	/	/	/	/	/	/	/	/	/	3
Kuroda	1996	/	/	/	/	/	/	/	/	/	/	3
Sharkey	1996	/	/	/	/	/	/	/	/	/	/	5
Aoyama	1997	/	/	/	/	/	/	/	/	/	/	3
Butcher	1997	/	/	/	/	/	/	/	/	/	/	3
Sharkey	1997	/	/	/	/	/	/	/	/	/	/	1
Toung	1997	/	/	/	/	/	/	/	/	/	/	2
Takamatsu	1998	/	/	/	/	/	/	/	/	/	/	4
Bochelen	1999	/	/	/	/	/	/	/	/	/	/	5
Janelidze	1999	/	/	/	/	/	/	/	/	/	/	0
Kuroda	1999	/	/	/	/	/	/	/	/	/	/	4
Toung	1999	/	/	/	/	/	/	/	/	/	/	5
Yoshimoto	1999	/	/	/	/	/	/	/	/	/	/	3
Aronowski	2000	/	/	/	/	/	/	/	/	/	/	4
Miyazawa	2000	/	/	/	/	/	/	/	/	/	/	2
Arii	2001	/	/	/	/	/	/	/	/	/	/	4
Ebisu	2001	/	/	/	/	/	/	/	/	/	/	5
Fredduzzi	2001	/	/	/	/	/	/	/	/	/	/	7
McCarter	2001	/	/	/	/	/	/	/	/	/	/	3
McGregor	2001	/	/	/	/	/	/	/	/	/	/	0
Takamatsu	2001	/	/	/	/	/	/	/	/	/	/	2
Maeda	2002	/	/	/	/	/	/	/	/	/	/	4
Brecht	2003	/	/	/	/	/	/	/	/	/	/	4
Furuchi	2003a	/	/	/	/	/	/	/	/	/	/	4
Furuichi	2003b	/	/	/	/	/	/	/	/	/	/	4
Aronowski	2004	/	/	/	/	/	/	/	/	/	/	0
Chung	2004	/	/	/	/	/	/	/	/	/	/	3
Nito	2004	/	/	/	/	/	/	/	/	/	/	6
Shichinohe	2004	/	/	/	/	/	/	/	/	/	/	4

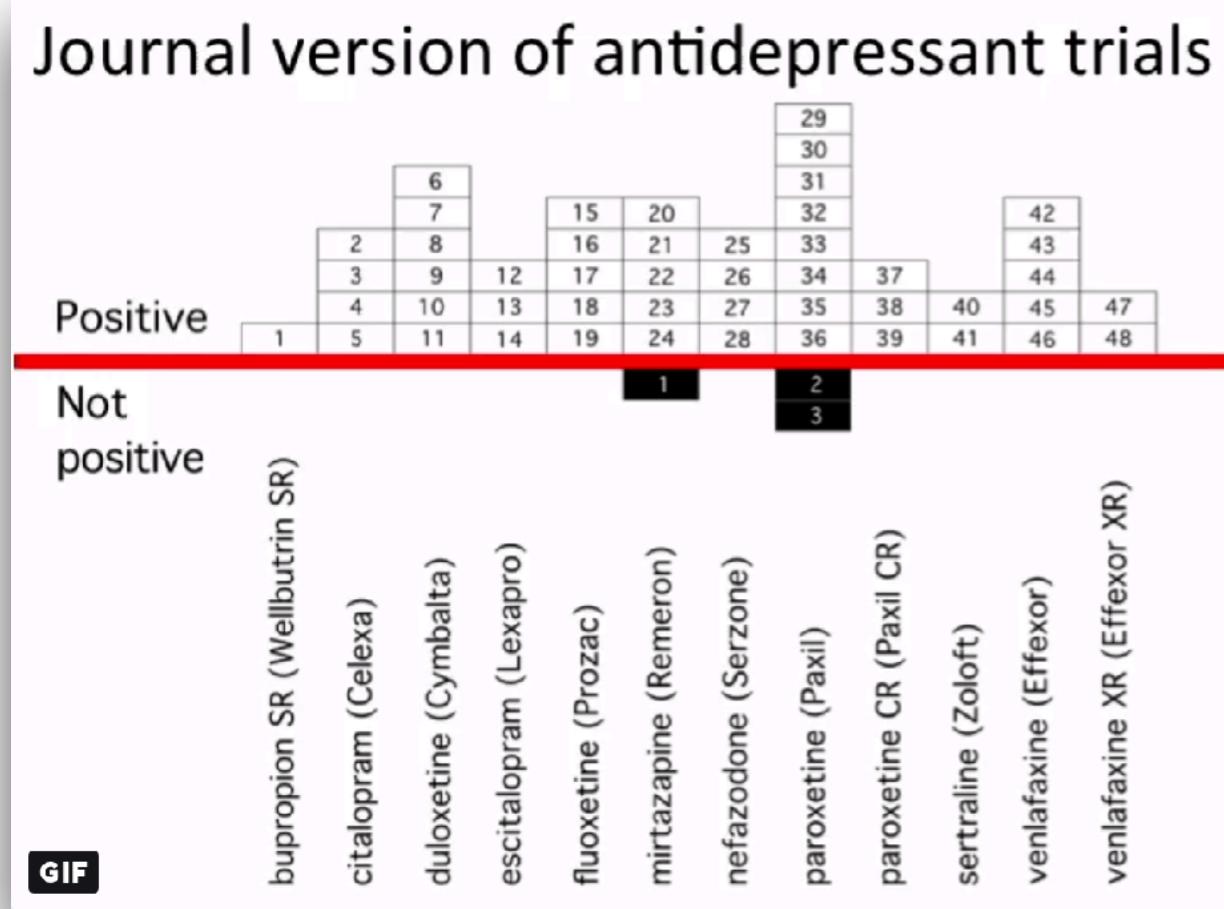
Studies fulfilling the criteria of (1) peer reviewed publication; (2) control of temperature; (3) random allocation to treatment or control; (4) blinded induction of ischaemia; (5) blinded assessment of outcome; (6) use of anaesthetic without significant intrinsic neuroprotective activity; (7) animal model (aged, diabetic or hypertensive); (8) sample size calculation; (9) compliance with animal welfare regulations; and (10) statement of potential conflict of interests.



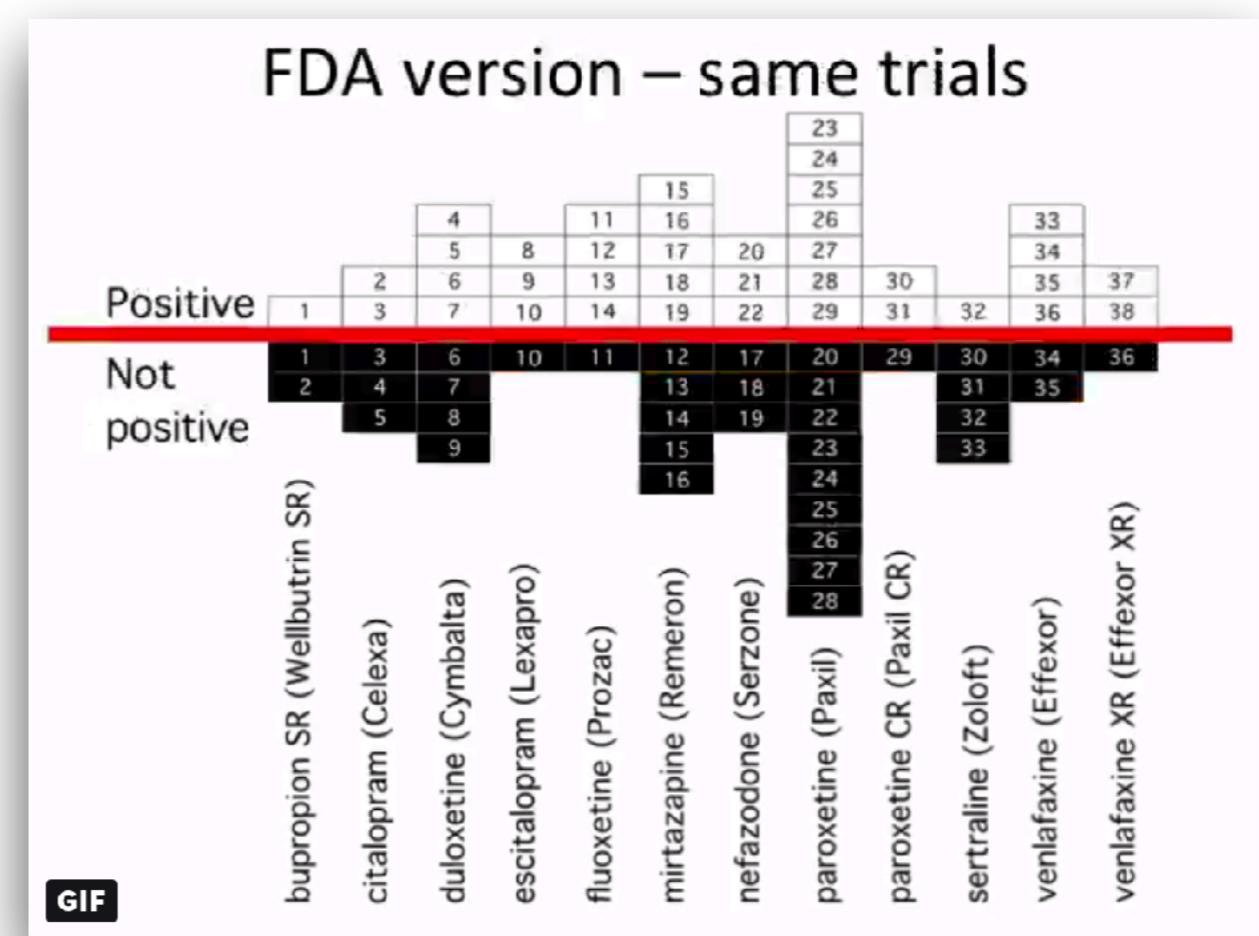
Publication bias (and other biases)

Add some publication bias to the mix: Efficacy of anti-depressants (Turner et al. 2008)

Trials published in journals
48 positive, 3 negative

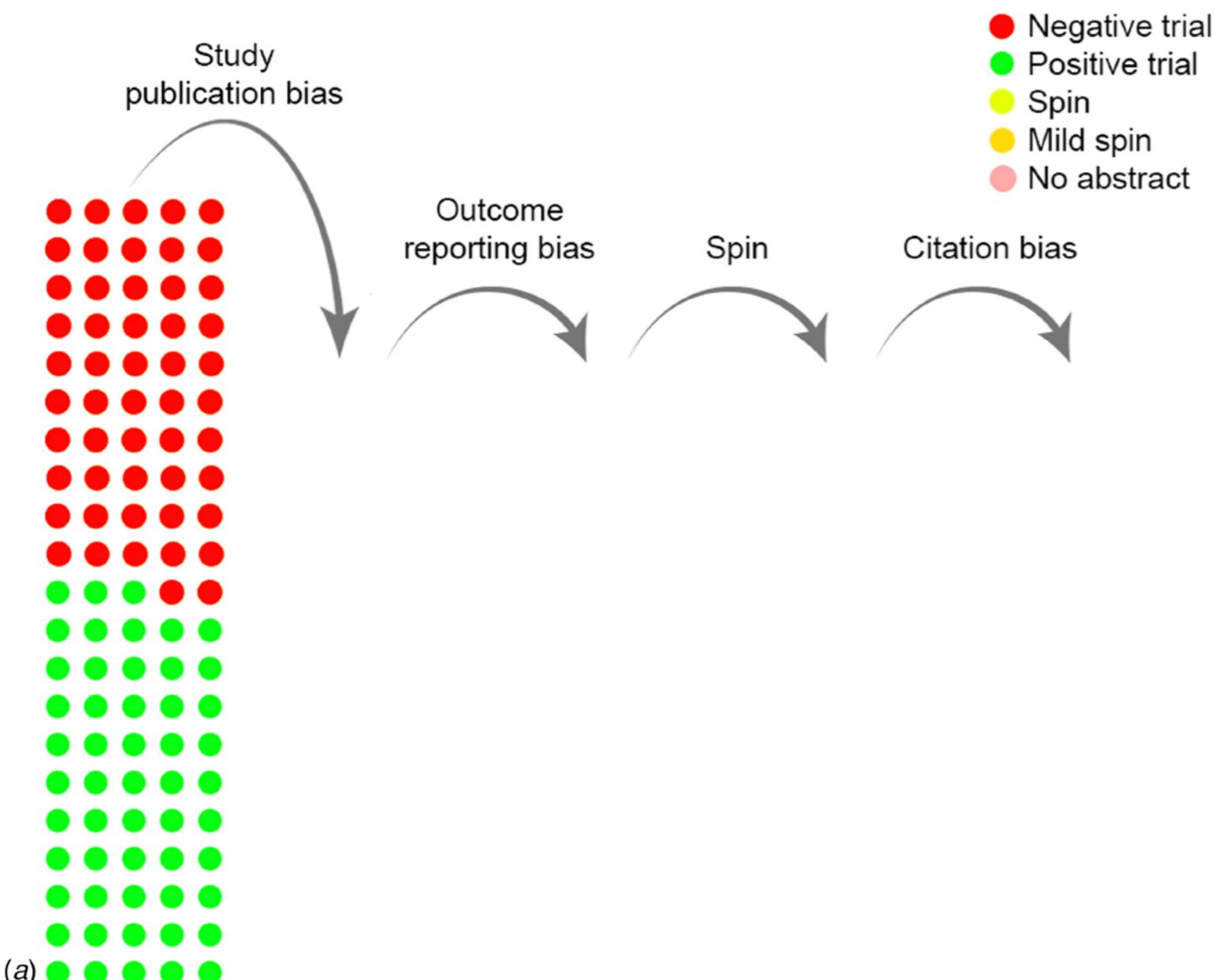


Trials preregistered at FDA*
38 positive, 36 negative



* Food and Drug Administration

More than just publication bias ...



- **Study publication bias:** non-publication of an entire study
- **Outcome reporting bias:** non-publication of negative outcomes within a published article or switching the status of (non-significant) primary and (significant) secondary outcomes
- **Spin:** authors conclude that the treatment is effective despite non-significant results on the primary outcome
- **Citation bias:** Studies with positive results receive more citations than negative studies

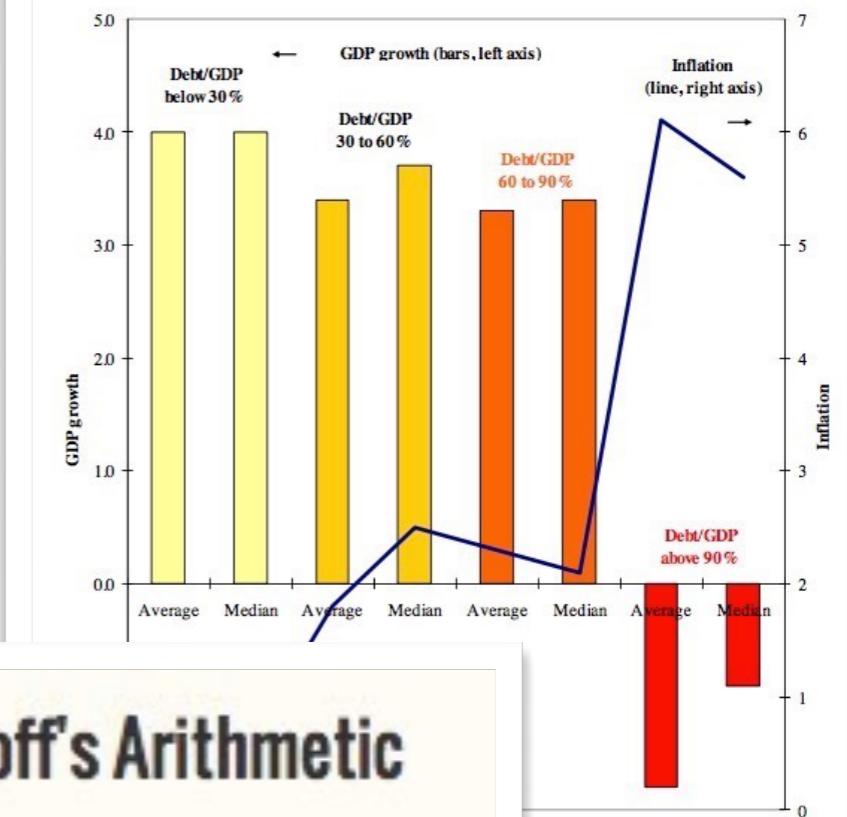
Honest mistakes

A Decade of Replications: Lessons from the Quarterly Journal of Political Science

Posted on December 9, 2014 by nicholaseubank

- Reproducible analysis code and open data required at submission - “inhouse checking” in review process
- **54% of all submissions had results in the paper that did not match the computed results from the code**
- wrong signs, wrong labeling of regression coefficients, errors in sample sizes, wrong descriptive stats

90%-Excel-Gate



How Much Unemployment Was Caused by Reinhart and Rogoff's Arithmetic Mistake?

Published: 16 April 2013

The Excel Depression

APRIL 18, 2013



Paul Krugman

In this age of information, math errors can lead to disaster. NASA's [Mars Orbiter crashed](#) because engineers forgot to convert to metric measurements; JPMorgan Chase's "[London Whale](#)" venture went bad in part because modelers divided by a sum instead of an average. So, did an [Excel coding error destroy the economies of the Western world?](#)

The most important point of the story: The original authors shared their raw data, which made it possible to correct the honest mistake!



Brief Report

At what sample size do correlations stabilize?

Felix D. Schönbrodt ^a✉, Marco Perugini ^b✉

Show more

<https://doi.org/10.1016/j.jrp.2013.05.009>

[Get rights and content](#)

Referred to by Felix D. Schönbrodt, Marco Perugini

[Corrigendum to "At what sample size do correlations stabilize?" \[J. Res. Pers. 47 ...\]](#)

Journal of Research in Personality, Volume 74, June 2018, Pages 194

The authors regret that a software bug was detected in the reproducible R code that accompanied the paper. The authors thank Dean Eckles and André Kretzschmar who independently detected and reported the bug.

The authors are sorry for this error, apologise for any inconvenience caused, and highlight the usefulness and necessity of open reproducible code. The corrected simulation code can be found at <https://osf.io/ydbwr/>.

p -hacking +
(uncorrectable) honest mistakes +
publication bias

=



Kahneman: Open Letter



Daniel Kahneman, Nobel prize 2002

I believe that you should
collectively do something
about this mess.

**I see a train wreck
looming.**

Replicability-Index

Improving the replicability of empirical research

A Meta-Scientific Perspective on “Thinking: Fast and Slow”

⌚ December 30, 2020 📄 Implicit Priming, Kahneman, r-index, Thinking Fast and Slow, Z-Curve

In conclusion, Daniel Kahneman is a distinguished psychologist who has made valuable contributions to the study of human decision making. His work with Amos Tversky was recognized with a Nobel Memorial Prize in Economics ([APA](#)). It is surely interesting to read what he has to say about psychological topics that range from cognition to well-being. However, his thoughts are based on a scientific literature with shaky foundations. Like everybody else in 2011, Kahneman trusted individual studies to be robust and replicable because they presented a statistically significant result. In hindsight it is clear that this is not the case.

Kahnemann's reply on the blog:



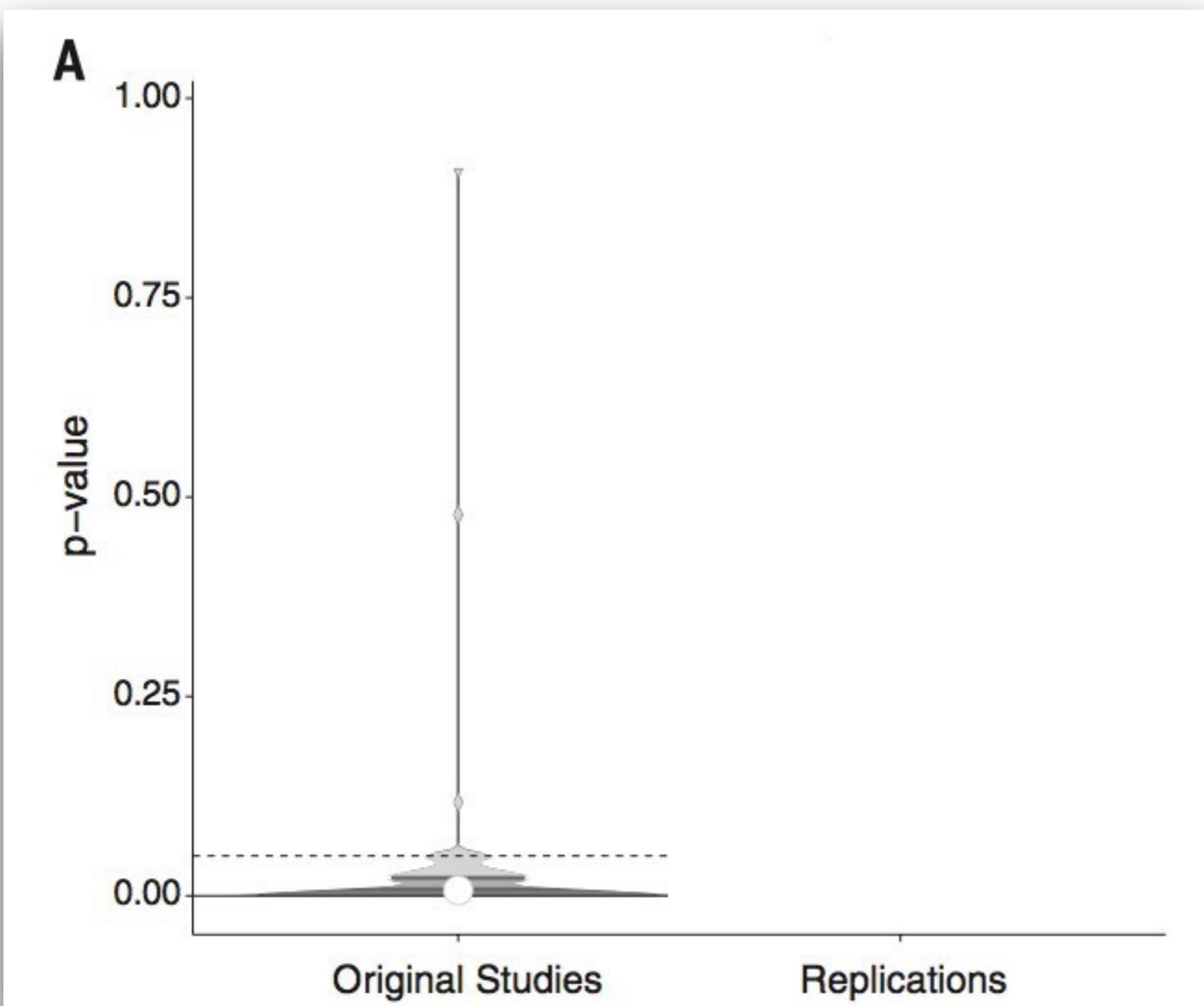
I accept the basic conclusions of this blog. [...]

What the blog gets absolutely right is that I placed too much faith in underpowered studies. [...]

My position when I wrote “Thinking, Fast and Slow” was that if a large body of evidence published in reputable journals supports an initially implausible conclusion, then scientific norms require us to believe that conclusion. Implausibility is not sufficient to justify disbelief, and belief in well-supported scientific conclusions is not optional. This position still seems reasonable to me – it is why I think people should believe in climate change. But the argument only holds when all relevant results are published.

I knew, of course, that the results of priming studies were based on small samples, that the effect sizes were perhaps implausibly large, and that no single study was conclusive on its own. What impressed me was the unanimity and coherence of the results reported by many laboratories. [...]

However, I now understand that my reasoning was flawed and that I should have known better. Unanimity of underpowered studies provides compelling evidence for the existence of a severe file-drawer problem (and/or p-hacking).



- 100 replications
- 36% of all replications were significant
 - PS - cog: 53%
 - JEP:LMC: 48%
 - PS - soc: 29%
 - JPSP - soc: 23%
- 83% of all effect sizes are smaller than the original:
 $M_O: r = .40$
 $M_R: r = .20$

Only psychology?
An outlook to other disciplines.



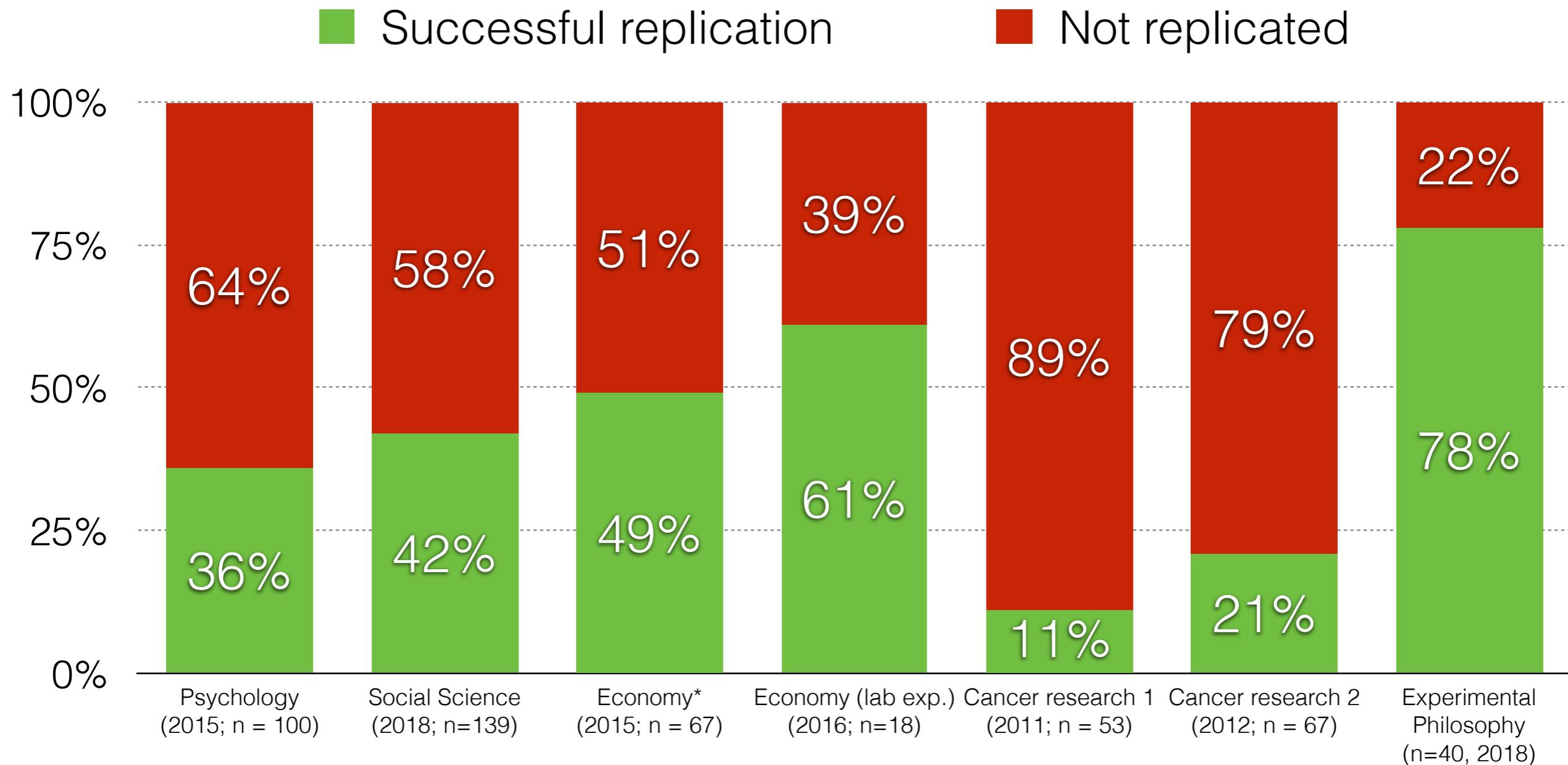
Many landmark findings in preclinical oncology research are not reproducible, in part because of inadequate cell lines and animal models.

Raise standards for preclinical cancer research

C. Glenn Begley and Lee M. Ellis propose how methods, publications and incentives must change if patients are to benefit.

- 53 ‘landmark studies’, not randomly selected: fresh approaches targeted for future drug development
- “scientific findings were confirmed in only 6 (**11%**) cases. Even knowing the limitations of preclinical research, this was a shocking result.”
- Bayer Healthcare: 67 target-validation projects in oncology, women’s health, and cardiovascular medicine. Only 14 (**21%**) could be reproduced.

Which part of published findings can be independently replicated?



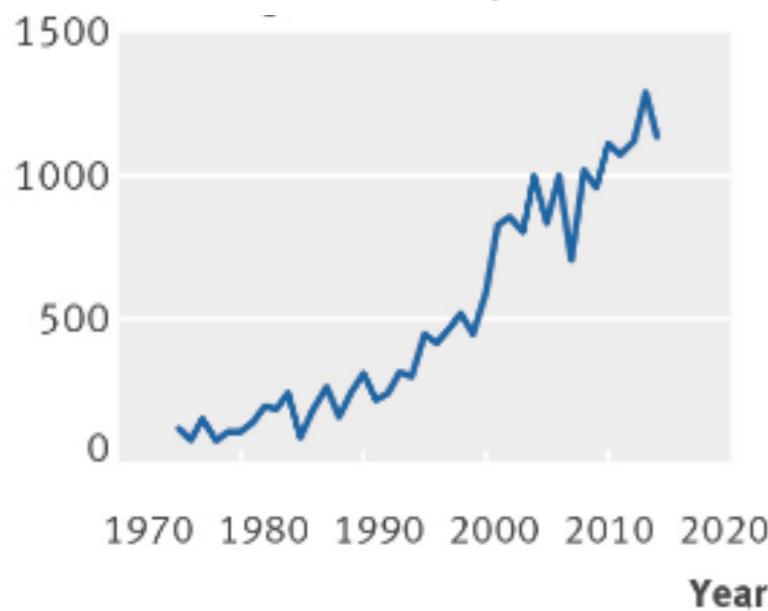
* The data on economics is about *reproducibility*; i.e. the attempt to get the same results if you apply the original data analysis on the original data set.

Buzzwords in scientific abstracts

+880% from 1974- 2014



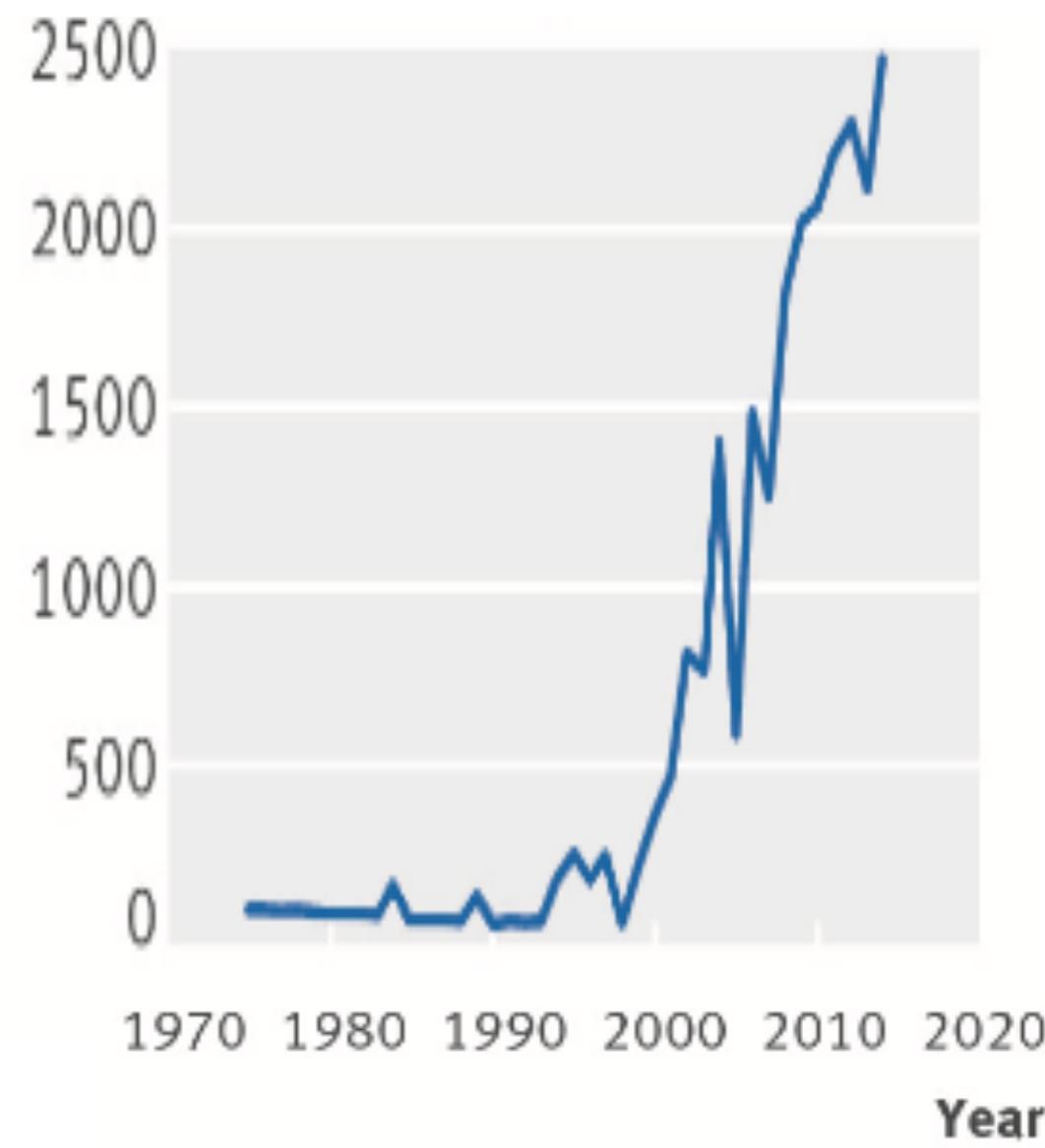
Amazing!!



Enormous!!



Groundbreaking!!!



Thesis:

Our current incentives foster questionable research practices, which decrease the truth value of our shared knowledge.

What is good for the individual careers of researchers leads to a collective fiasco.

Researchers who do it right (i.e., high power, no QRPs, transparency) have a clear competitive disadvantage.

Anti-Thesis:

Society pays for us that we generate valid and robust knowledge.

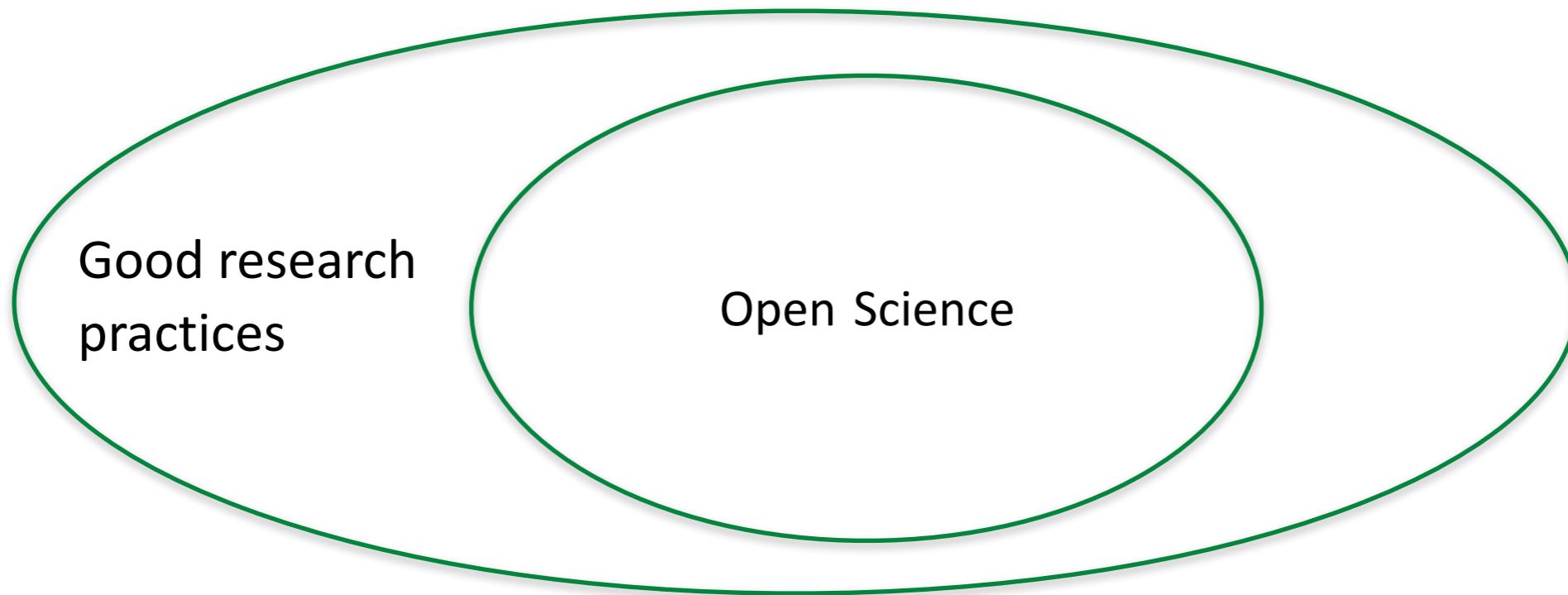
Our incentives should be chosen in a way that they foster good science.

Researchers who do it right should be supported and promoted.

The new way of doing research:

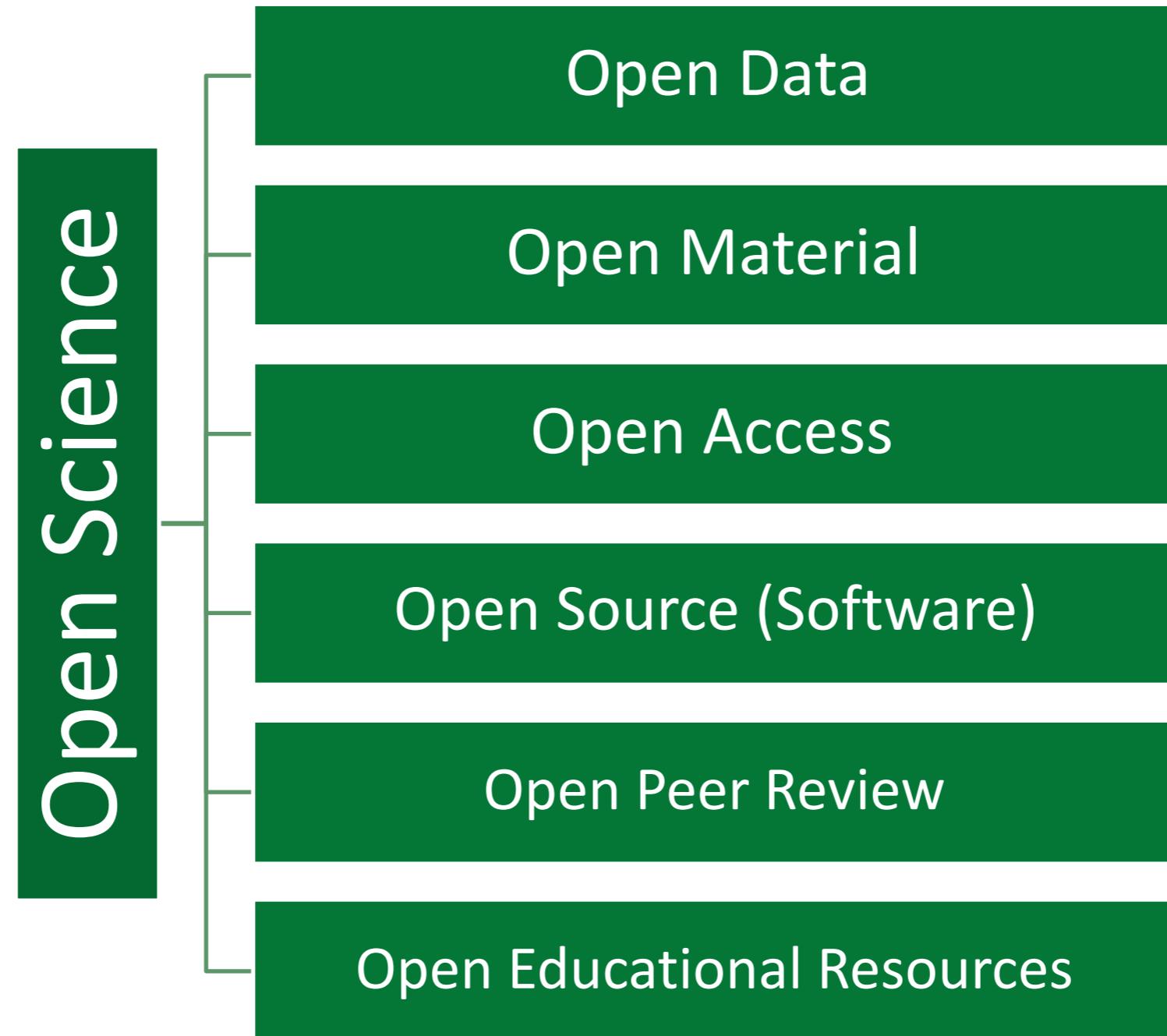
A scientific framework
for the 21. century

Open Science ∈ Good Science



Good research >> avoiding bad research !

Pillars of Open Science







Data & analysis script availability (prevalence estimates)

	Data	Analysis scripts
Psychology (2014-2017) ¹	2% [1-4%]*	1% [0-1%]
Social Sciences (2014-2017) ²	7% [2-13%]	1% [0-3%]

¹Hardwicke et al. (2021)

*[95% confidence intervals]

²Hardwicke et al. (2020)

Data availability on request (selected studies)

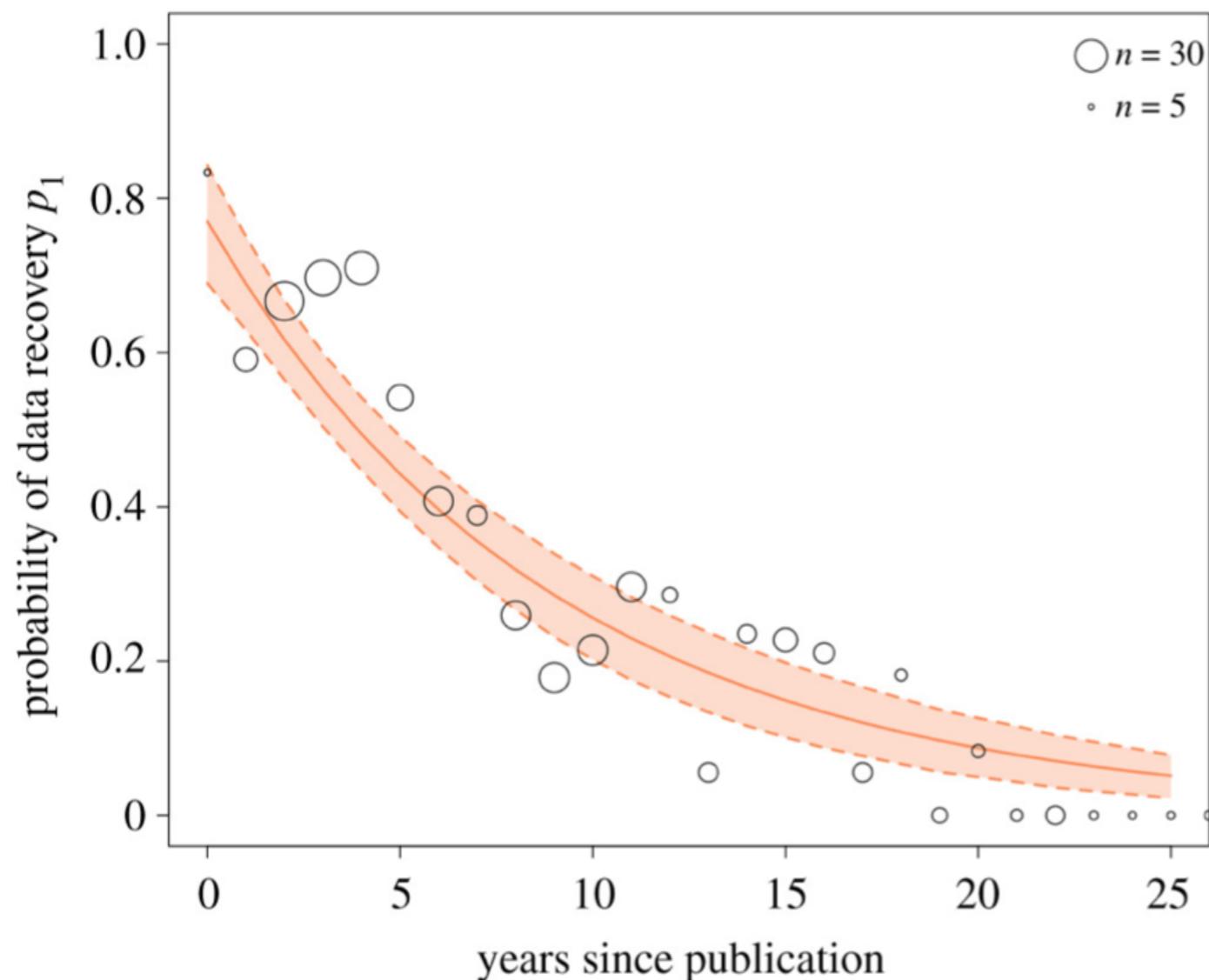
	Data shared
141 articles published in four major APA journals (2004) ³	27%
516 ecology articles published (1991-2011) ⁴	20%
111 most highly-cited psychology & psychiatry articles (2006-2016) ⁵	14%

³Wicherts et al. (2006)

⁴Vines et al. (2014)

⁵Hardwicke & Ioannidis (2018)

Data decay



- „probability of finding material for any publication **halves every 5.7 years**“
- „probability of recovering data for studies > 20 years ago is **close to zero**“

Funders demand it



„We expect our researchers to maximise the availability of research data, software and materials with as few restrictions as possible. **As a minimum, the data underpinning research papers should be made available to other researchers at the time of publication. [...]**

Wellcome will also consider whether researchers have managed and shared their research outputs in line with our requirements, as a critical part of the end of grant reporting process“



„The NIH expects and supports the **timely release and sharing of final research data** from NIH-supported studies for use by other researchers. [...] ... are expected to include a plan for data sharing or state why data sharing is not possible.“



„Ebenso erwartet der SNF, dass Daten, die während der Forschungsarbeiten produziert wurden, künftig auf öffentlich zugänglichen, digitalen Datenbanken archiviert werden, sofern dem keine rechtlichen, ethischen, urheberrechtlichen oder andere Klauseln entgegenstehen.“



„**It is recommended to make all research data, code and software created within a research project available for reuse**, for example under Creative Commons, GNU, MIT or another relevant licence.“



„FAIR (Findable, Accessible, Interoperable and Re-usable data) and open data sharing should **become the default for the results of EU-funded scientific research.**“

Journals demand it: <https://topfactor.org/>

Journal	Total	Data Citation	↓	Data Transparency	Analysis Code Transparency
Meta-Psychology 	27	3		3	3
LNU Open					
Archives of Scientific Psychology 	3	0		2	0
American Psychological Association					
Journal of Research in Personality 	19	0		2	2
Elsevier					
Social Psychological Bulletin 	18	1		2	1
PsychOpen					
Collabra 	20	2		2	2
University of California Press					
Social Cognition 	13	2		2	2
Guilford Press					
Personality Science 	24	3		2	2
Cortex 	23	3		2	2
Elsevier					
Royal Society Open Science 	14	2		2	2
Royal Society Publishing					
Advances in Methods and Practices in Psychological Science 	25	2		2	2
SAGE					
Science 	11	2		2	2
AAAS					

- **TOP Level 3** = Data must be posted to a trusted repository, and reported analyses will be reproduced independently prior to publication.
- **TOP Level 2** = Data must be posted to a trusted repository. Exceptions must be identified at article submission.
- As of Dec 2020, the TOP factor website lists for psychology:
 - 1 journal with data transparency level 3, and
 - 14 journals with level 2

Not only open, but FAIR

The FAIR principles



Findable



Accessible



Interoperable



Reusable

Good research data management

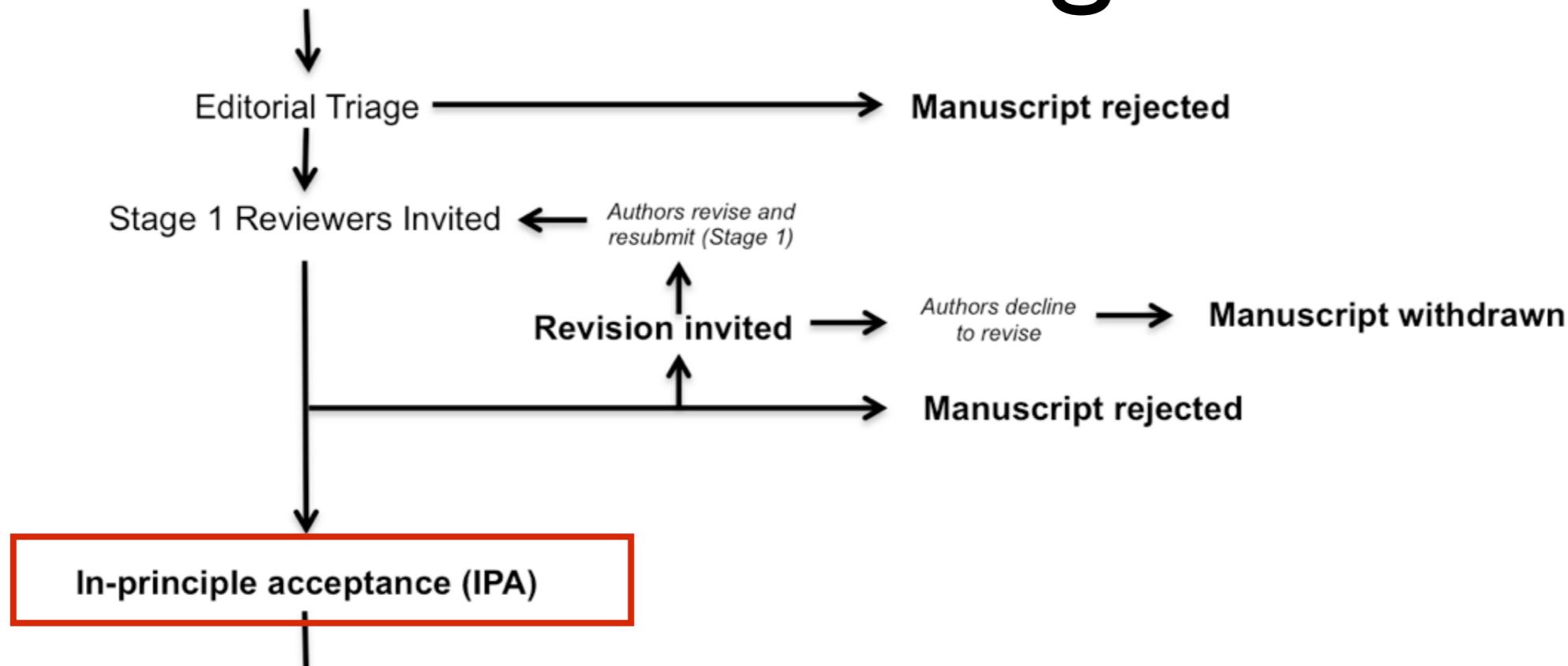


Stage 1 Registered Report

Peer review of Introduction, Method, Proposed Analyses, and Pilot Data (if applicable)

Registered Reports

<https://cos.io/rr/>

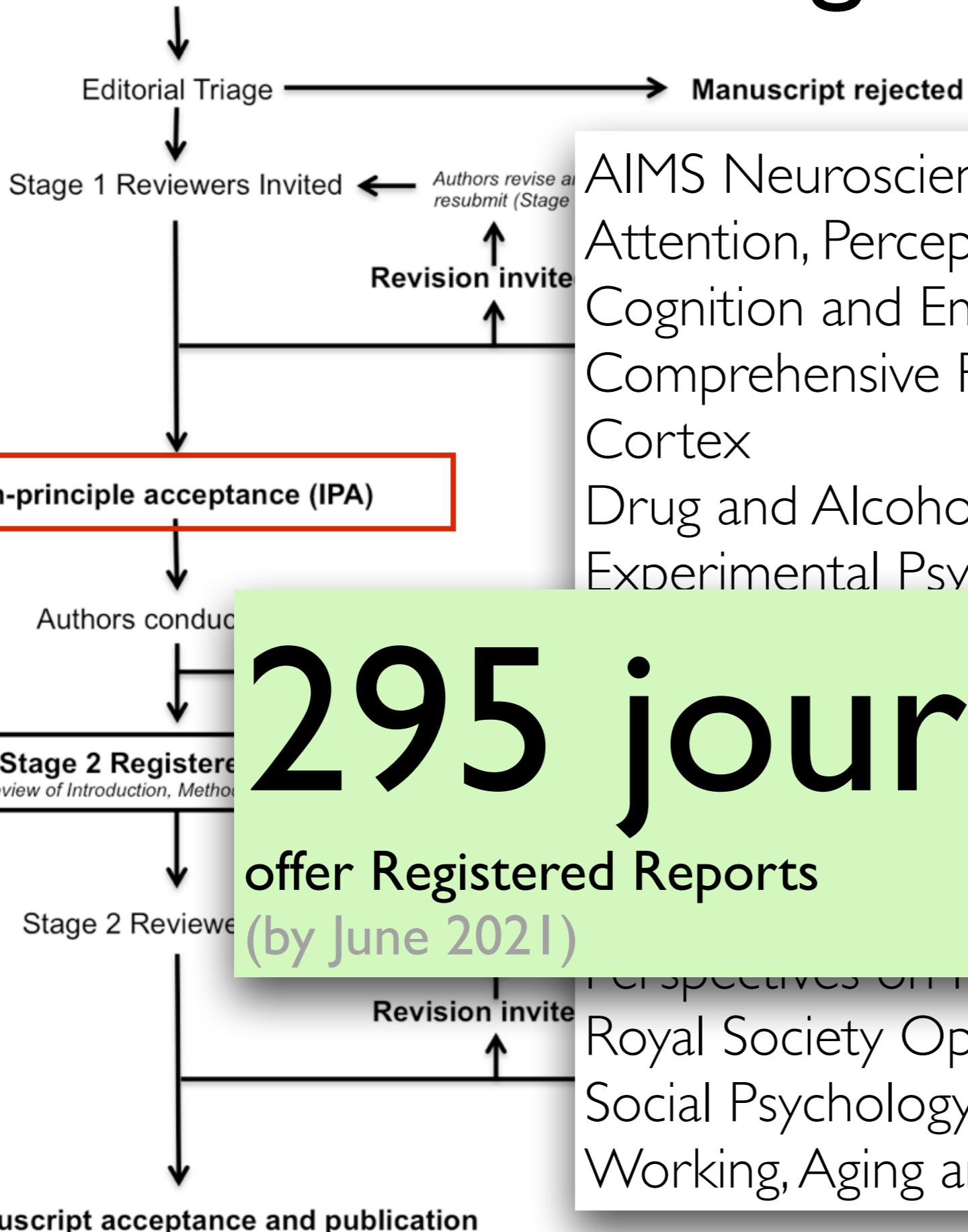


Stage 1 Registered Report

Peer review of Introduction, Method, Proposed Analyses, and Pilot Data (if applicable)

Registered Reports

<https://cos.io/rr/>



AIMS Neuroscience

Attention, Perception, and Psychophysics

Cognition and Emotion

Comprehensive Results in Social Psychology

Cortex

Drug and Alcohol Dependence

Experimental Psychology

295 journals

offer Registered Reports

(by June 2021)

Perspectives on Psychological Science

Royal Society Open Science

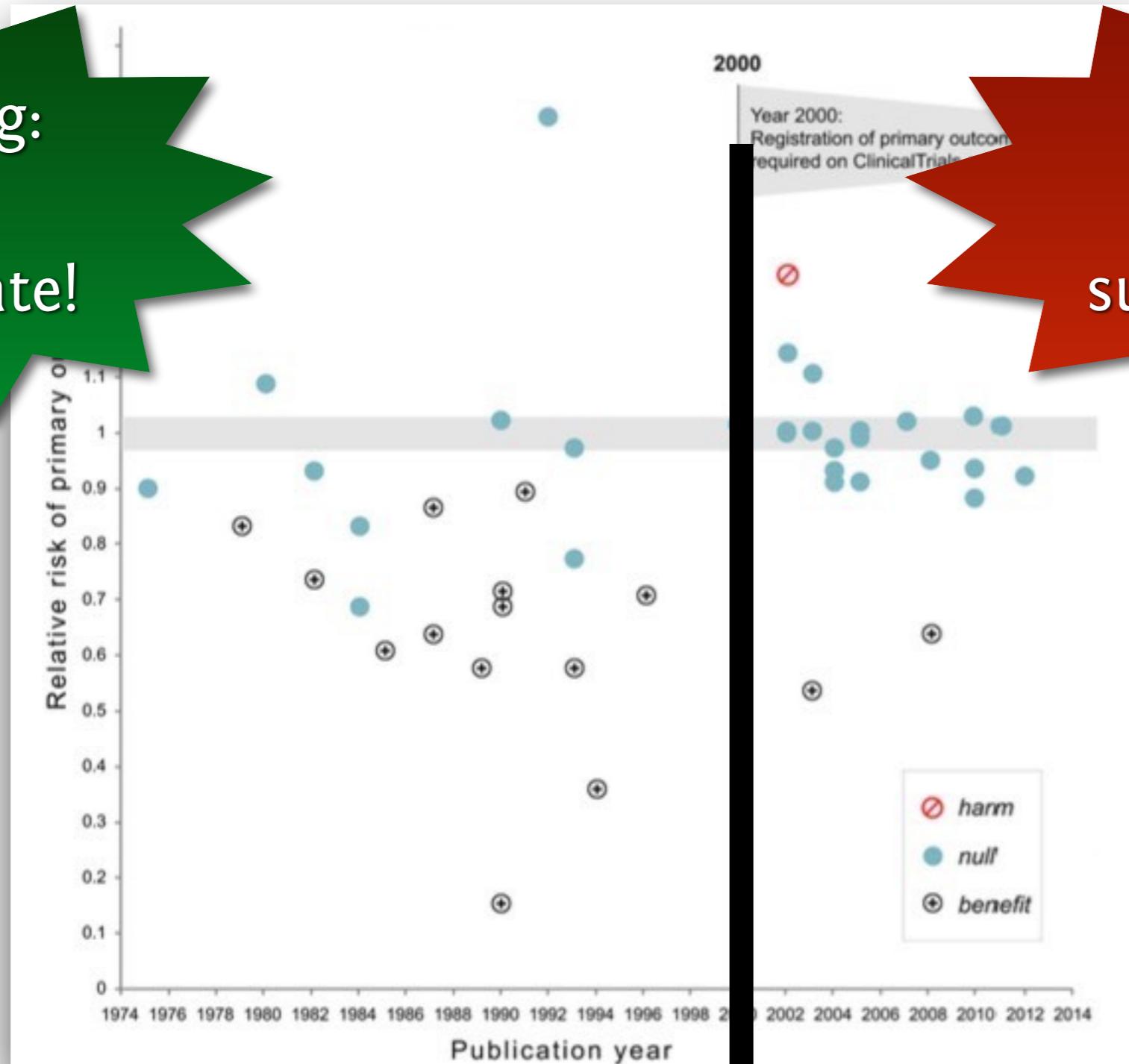
Social Psychology

Working, Aging and Retirement

Pre-registration causes medicines to stop working!

no prereg:
57%
success rate!

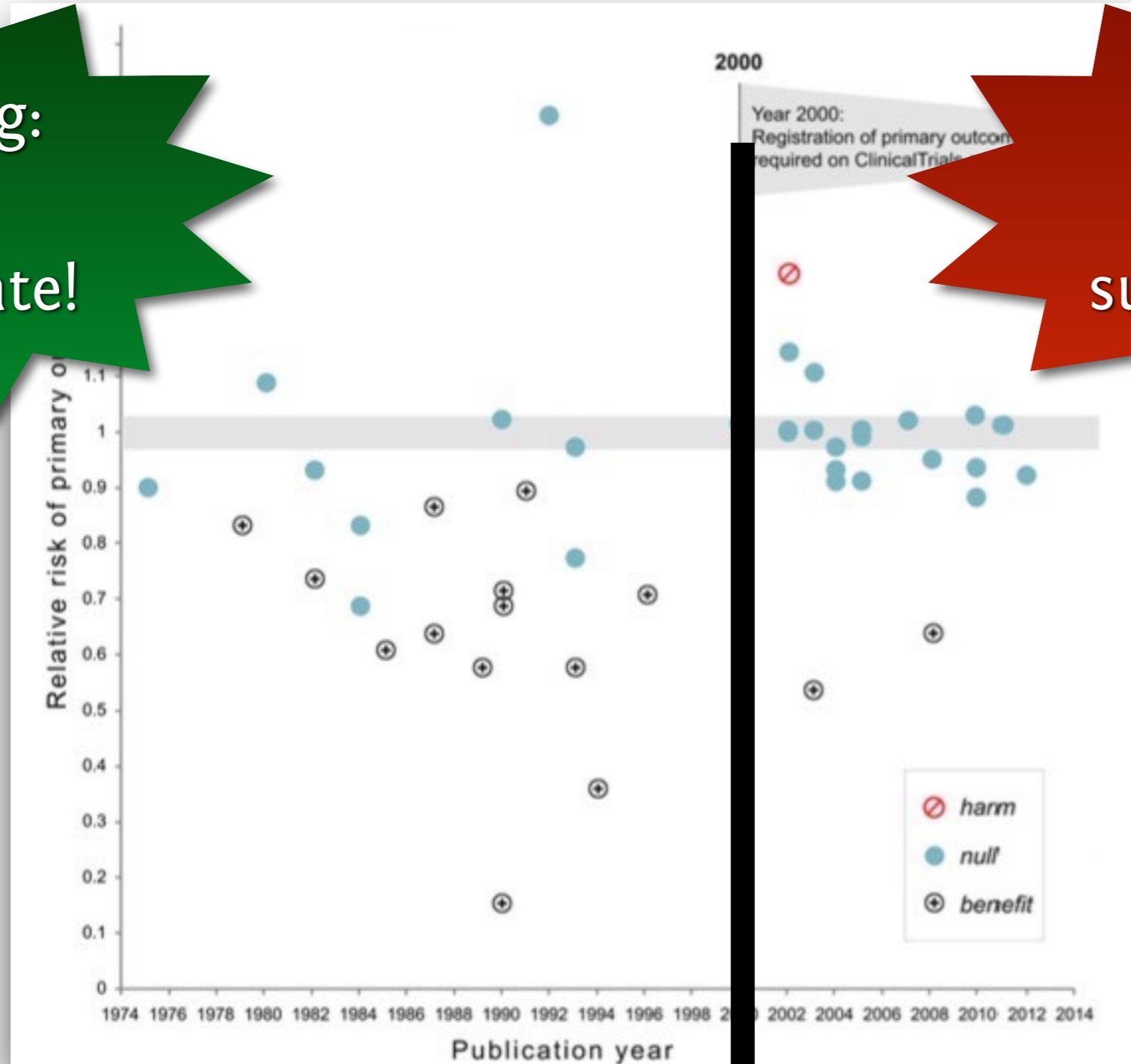
prereg:
8%
success rate...



Pre-registration causes p-hacking ~~medicines~~ to stop working!

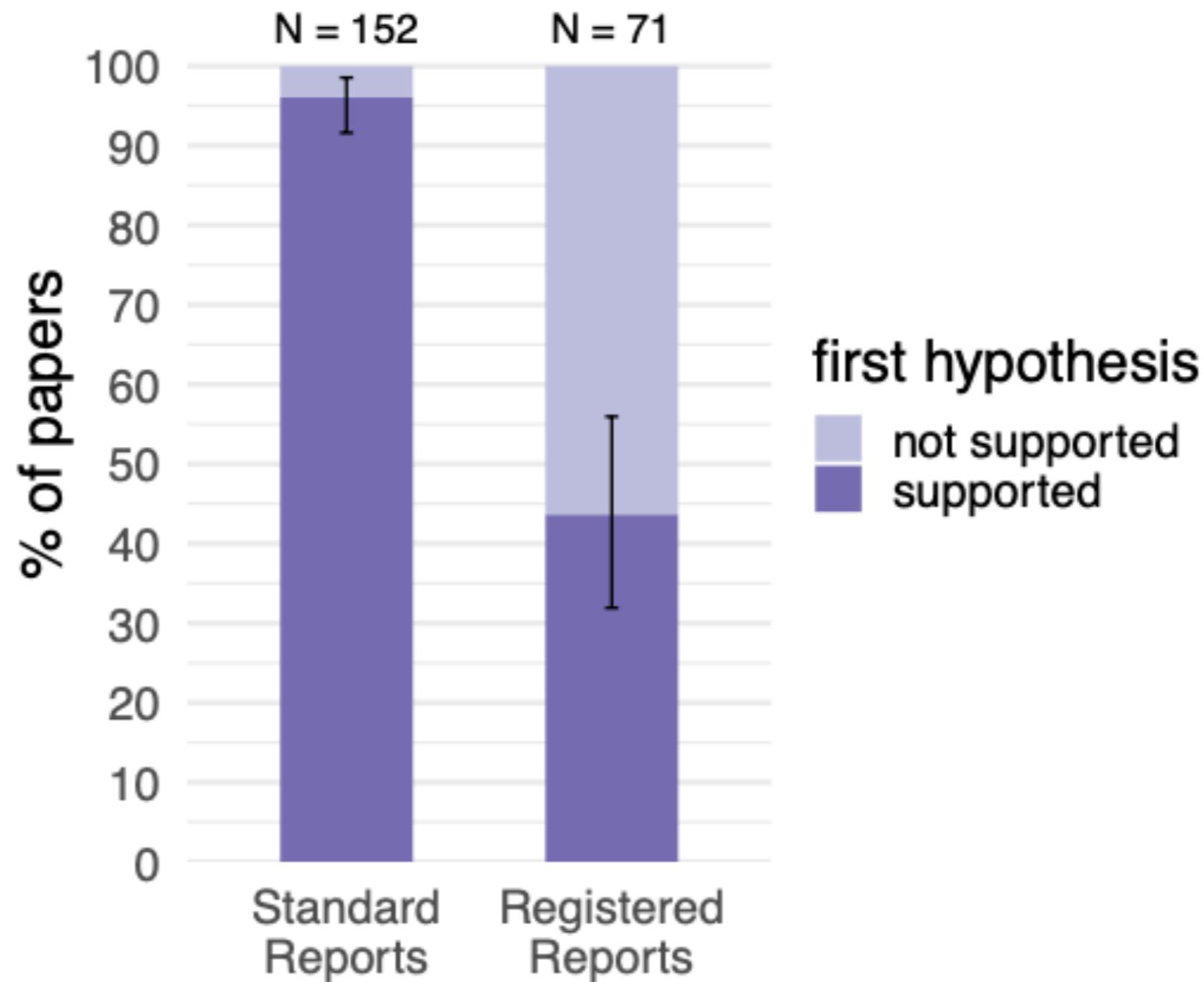
no prereg:
57%
success rate!

prereg:
8%
success rate...





?



Open Access



Publish (nearly) all of your papers as green Open Access

 SHERPA/ROMEO

Search - Publisher copyright policies & self-archiving

One journal found when searched for: **0036-8075**

Journal:	Science (ISSN: 0036-8075, ESSN: 1095-9203)
RoMEO:	This is a RoMEO green journal
Author's Pre-print:	author can archive pre-print (ie pre-refereeing)
Author's Post-print:	author can archive post-print (ie final draft post-refereeing)
Publisher's Version/PDF:	author cannot archive publisher's version/PDF
General Conditions:	<ul style="list-style-type: none">Pre-print may be considered prior publicationPre-print on not-for-profit preprint servers where allowed, please contact editors for clarificationCannot archive until publicationAuthors retain copyrightOn author's personal website or institutional repositoryPublisher's version/PDF cannot be usedMust link to publisher versionSet statement must accompany post-print (see policy)Published source must be acknowledged with DOIAuthors covered by funding agency rules, may post author's post-print in PubMed Central or fundedAuthors covered by funding agency rules, must state on submission, for article to be released in P
Mandated OA:	Compliance data is available for 23 funders
Copyright:	License Agreement (pdf) - PMC policy - Pre-print comment - General Information for Authors
Updated:	22-Apr-2016 - Suggest an update for this record
Link to this page:	http://www.sherpa.ac.uk/romeo/issn/0036_8075/

<http://www.sherpa.ac.uk/romeo/index.php>

Open Science: The first steps

Step 1: Provide open material for an existing project (first steps)

- Create free account on the *Open Science Framework* (OSF; <https://osf.io>)
 - OSF = non-profit organization; has preservation fund that ensures availability of the uploaded material for at least 50 (?) years
- Publish open data, open material, reproducible analysis code, preprints, postprints, supplemental material
- Get a persistent URL (add that to your paper!) and even a doi.
- Easy start (example for psychology): Upload ...
 - the questionnaire items that you used in your study
 - The pictures you used as stimuli
 - your R scripts for data analysis

The screenshot shows a project page on the Open Science Framework (OSF). The title of the project is "Replicability, Robustness, and Reproducibility in Psychological Science". Below the title, it lists contributors: Brian A. Nosek, Tom E Hardwicke, Hannah Moshontz, Aurélien Allard, Katherine S. Corker, Anna Dreber Almenberg, Fiona Fidler, Joseph Hilgard, and Anne M. Scheel, Laura Scherer, Felix Schönbrodt, Simine Vazire. It also mentions affiliated institutions: University of Virginia, Center For Open Science. The date created was 2020-06-11 01:43 PM, and it was last updated on 2021-02-09 01:26 PM. The category is "Project". The description states: "Piece for the Annual Review of Psychology, completed February 2021". The license is CC-By Attribution 4.0 International. A note indicates that there are supplemental materials for the project on PsyArXiv. The page includes sections for "Wiki" and "Files". The "Wiki" section contains a brief description of the project as a repository for a paper prepared for the Annual Review of Psychology. The "Files" section shows a file named "RRR.AnnualReview.pdf" uploaded on 2021-02-09 01:25 PM and a file named "all_replication_studies.csv" uploaded on 2021-03-11 09:29 PM. On the right side, there are sections for "Citation" and "Components", which list various subprojects and their details.

Next steps

- Make your analyses reproducible
 - <https://the-turing-way.netlify.app>
 - Wilson et al. (2017). Good enough practices in scientific computing. doi:10.1371/journal.pcbi.1005510
- Make your manuscripts computable
 - Jupiter notebooks, rmarkdown
 - Limited journal space? Share supplemental analyses, robustness checks, etc. on OSF
- Share your research data
 - Working with human data? Respect GDPR, add open data to your consent form
 - Go for publicly funded, non-profit repositories
 - Make FAIR (findable, accessible, interoperable, reusable)
- Preregister your research



Guide for Reproducible Research

This guide covers topics related to skills, tools and best practices for research reproducibility.

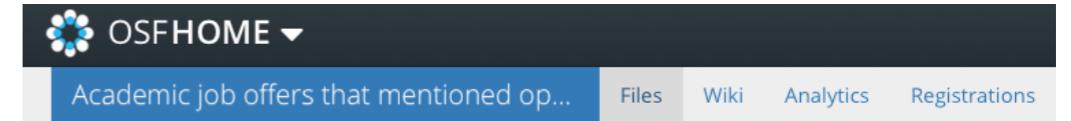
The Turing Way defines reproducibility in data research as data and code being available to fully rerun the analysis.

There are several definitions of reproducibility in use, and we discuss these in more detail in the Definitions section of this chapter. While it is absolutely fine for us each to use different words, it will be useful for you to know how The Turing Way defines reproducibility to avoid misunderstandings when reading the rest of the handbook.



Your benefit

- You will be a *more competitive job candidate*: More and more job listings have open science as desirable or essential job criterion (<https://osf.io/7jbnt/>)
- You will be a *more competitive grant applicant*: More and more funders require open science now
- More citations: Papers with open access versions get more citations (49% higher Altmetric Attention Score and 36% more citations; Fu & Hughey, 2019). Registered Reports get cited more often
- **You will do better science!**



Academic job offers that mentioned open science

The Department of Psychology encourages transparent and replicable research and pursues these objectives with open data, open material and pre-registration. Please state clearly in your application how you have followed these objectives in the past and how you intend to do so in the future.

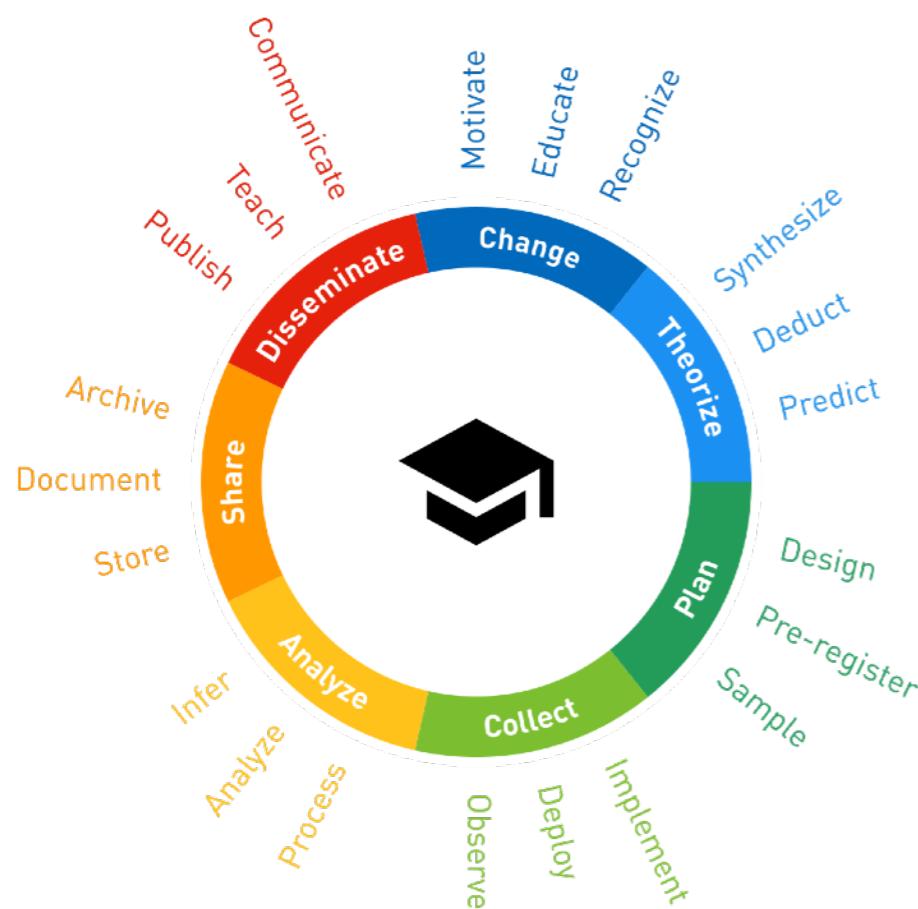


Open science is a policy priority for the European Commission and the standard method of working under its research and innovation funding programmes as it improves the quality, efficiency and responsiveness of research.

Resources

Resources

Open Scholarship Knowledge Base (OSKB): <https://www.oercommons.org/hubs/OSKB>



Resources



The screenshot shows the FORRT website. At the top, there is a dark navigation bar with the FORRT logo (a stylized castle tower icon) and the word "FORRT". To the right of the logo are five menu items: "ABOUT FORRT", "FORRT", "EDUCATIONAL NEXUS", "PEDAGOGIES", and a search icon. Below the navigation bar, the main content area features the text "Framework for Open and Reproducible Research Training" on the left and the FORRT logo with the word "FORRT" on the right.

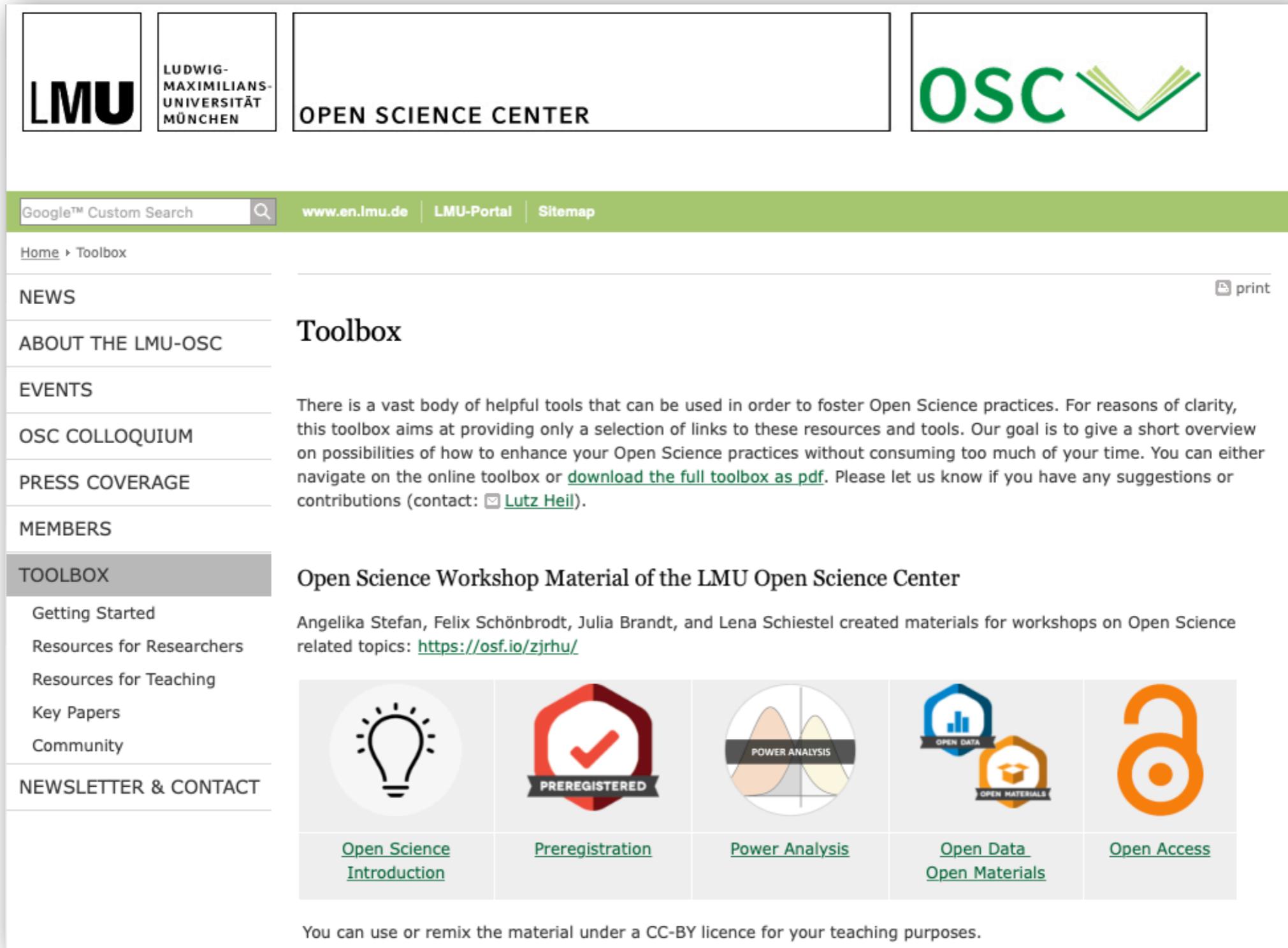
Framework for
Open and
Reproducible
Research
Training



FORRT

<https://forrt.org>

Resources



The screenshot shows the homepage of the LMU Open Science Center (OSC) Toolbox. At the top, there are three logos: LMU (Ludwig-Maximilians-Universität München), OPEN SCIENCE CENTER, and OSC. Below the logos is a green header bar with links for Google™ Custom Search, www.en.lmu.de, LMU-Portal, and Sitemap. The main content area has a sidebar on the left with links for Home, NEWS, ABOUT THE LMU-OSC, EVENTS, OSC COLLOQUIUM, PRESS COVERAGE, MEMBERS, and TOOLBOX. The TOOLBOX link is highlighted with a grey background. Under TOOLBOX, there are links for Getting Started, Resources for Researchers, Resources for Teaching, Key Papers, and Community. Below the sidebar is a section titled "Open Science Workshop Material of the LMU Open Science Center" created by Angelika Stefan, Felix Schönbrodt, Julia Brandt, and Lena Schiestel. It includes five icons: a lightbulb (Open Science Introduction), a red hexagon with a checkmark (Preregistration), a circle divided into four quadrants (Power Analysis), a blue hexagon with a bar chart (Open Data) and an orange hexagon with a graduation cap (Open Materials), and an orange padlock (Open Access). At the bottom, a note states: "You can use or remix the material under a CC-BY licence for your teaching purposes." The URL <https://www.osc.uni-muenchen.de/toolbox/> is also present.

Resources

- Gilad Feldman: *Endorsing open-science and supporting a science “credibility revolution”: Challenges, benefits, and practical tools and tips for early career and experienced researchers* (https://www.dropbox.com/sh/i8hy9kfjddleupl/AAA7abj8cj1_KHcvxpFAleyaa?dl=0)
- Dan Quintana: *Five things that every researcher should know about open science* <https://www.youtube.com/watch?v=0uCG3FI6ugE>
- Priya Silverstein: *Easing into open science: No time like the present.* <https://www.youtube.com/watch?v=owJaD3UiSeQ>
- Felix Schönbrodt: *Open Science Crash Course (everything in 5 hours).* <https://osf.io/qvfp8/>
- Crüwell, S., van Doorn, J., Etz, A., Makel, M. C., Moshontz, H., Niebaum, J., ... Schulte-Mecklenbeck, M. (2018, November 16). 7 easy steps to open science: An annotated reading list. <https://doi.org/10.31234/osf.io/cfzyx>