



~~The Metascience of Peer Review~~ The Metascience of Research Integrity

Tom Stafford,
Professor of Cognitive Science, University of Sheffield
Senior Research Fellow, Research on Research Institute

t.stafford@researchonresearch.org

2026-01-30



tomstafford.github.io



Peer review is central to the lives of researchers - it determines if our papers get accepted and if our funding proposals are successful. Metascience - also known as the science of science, or research on research - takes peer review as an object of study in its own right, and asks what we have learnt about how peer review works and how it might be improved. I'll discuss how we worked with a research funder to evaluate a trial of Distributed Peer Review (a system where applicants for a funding call review each others' applications), and talk about the topics in the psychology of expertise, judgement and decision making can be given new relevance in the study of peer review.



tomstafford.github.io

Research Integrity

These slides:
bit.ly/tomstafford

Universities?

Fostering Accountability for the
Integrity of Research Studies
(FAIRS), April 2025

<https://www.sjcfairsmeeting.com/>

A review of annual statements on research integrity from U.K. institutions in 2023–2024, with a focus on research fraud

Dorothy V. M. Bishop

Experimental Psychology, University of Oxford Medical Sciences Division, Oxford,
United Kingdom of Great Britain and Northern Ireland

Keywords: fabrication, falsification, institutional investigation, PubPeer, research integrity officer, transparency

ABSTRACT

Estimates of the frequency of research misconduct appear much higher in self-report surveys than would be expected from numbers of institutional investigations, but there is little hard data o

recent data fi
there
PubPe
49 co
and 21
comp.
multif
allega
institu
“In 117 institutional statements with
usable data, there were only 25
allegations of fabrication or
falsification involving 13
universities”

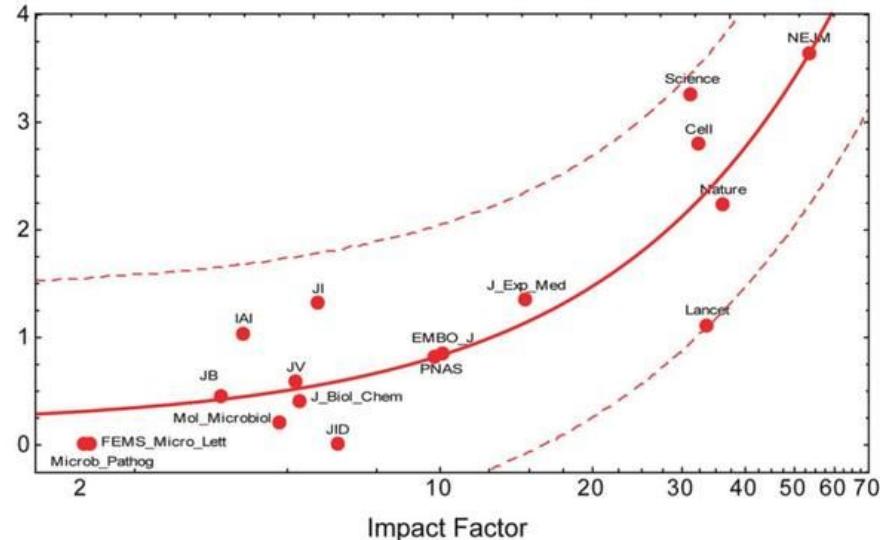
used proactively in investigations of misconduct, and that research integrity reports should be made openly available to increase confidence in the process.

Downloaded from <http://direct.mit.edu/qss/article-pdf/doi/10.1162/QSS.a.394/25579/qss.a.394.pdf> by guest.

<https://doi.org/10.1162/QSS.a.394> These slides:
bit.ly/tomstafford

Publishers?

Retraction Index



Paper-pushers

2

World, new scientific papers published, '000

■ Regular issue ■ Special issue

For-profit
open-access*

Selected traditional
publishers†



*Frontiers, Hindawi and MDPI

†Data only available for Nature, Springer and Wiley

Sources: "The strain on scientific publishing",
by Mark Hanson et al., 2024; *The Economist*

Economist 2024-11-20 [Scientific publishers
are producing more papers than ever](#)

Gáspár Jékely
https://jekelylab.github.io/COS_goes_FOSS_publishing.html

These slides:
bit.ly/tomstafford



The Research on Research
Institute Impact Report
(2023-2025)

[https://researchonresearch.org/
roris-impact-what-weve-achieve
d-and-where-were-going/](https://researchonresearch.org/roris-impact-what-weve-achieved-and-where-were-going/)





STIFTESSEN
DAM

Funders?



STIFTESSEN
DAM



Research
England

These slides:
bit.ly/tomstafford

Distributed Peer Review

Butters, A., Marshall, M. B.,
Pinfield, S., Stafford, T.,
Bondarenko, A., Neubauer, B., ...
Denecke, H. (2026, January 6).
Applicants as reviewers: A mixed
methods evaluation of the risks,
benefits, and potential of
distributed peer review for grant
funding allocations.
https://doi.org/10.31222/osf.io/t2p56_v1



These slides:
bit.ly/tomstafford

Distributed Peer Review @ VolkswagenStiftung

DPR



140 proposals submitted

18 proposals funded
3 recommended by both processes

60% overlap

47% overlap

Internal Shortlisting
70 shortlisted

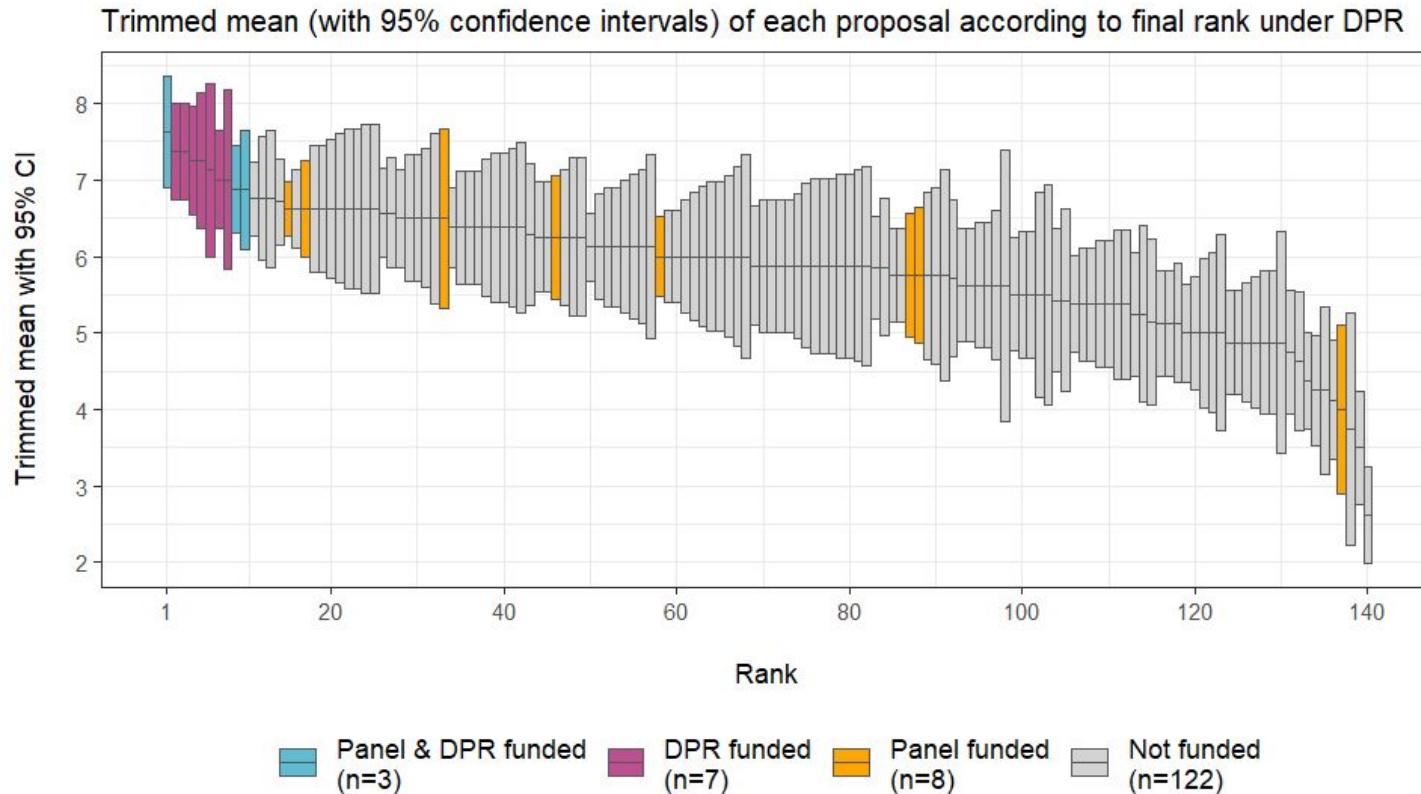
Quick Assessment
45 with 1+ A-, A, A+

Panel discussion
42 discussed

11 proposals recommended for funding

Panel Review

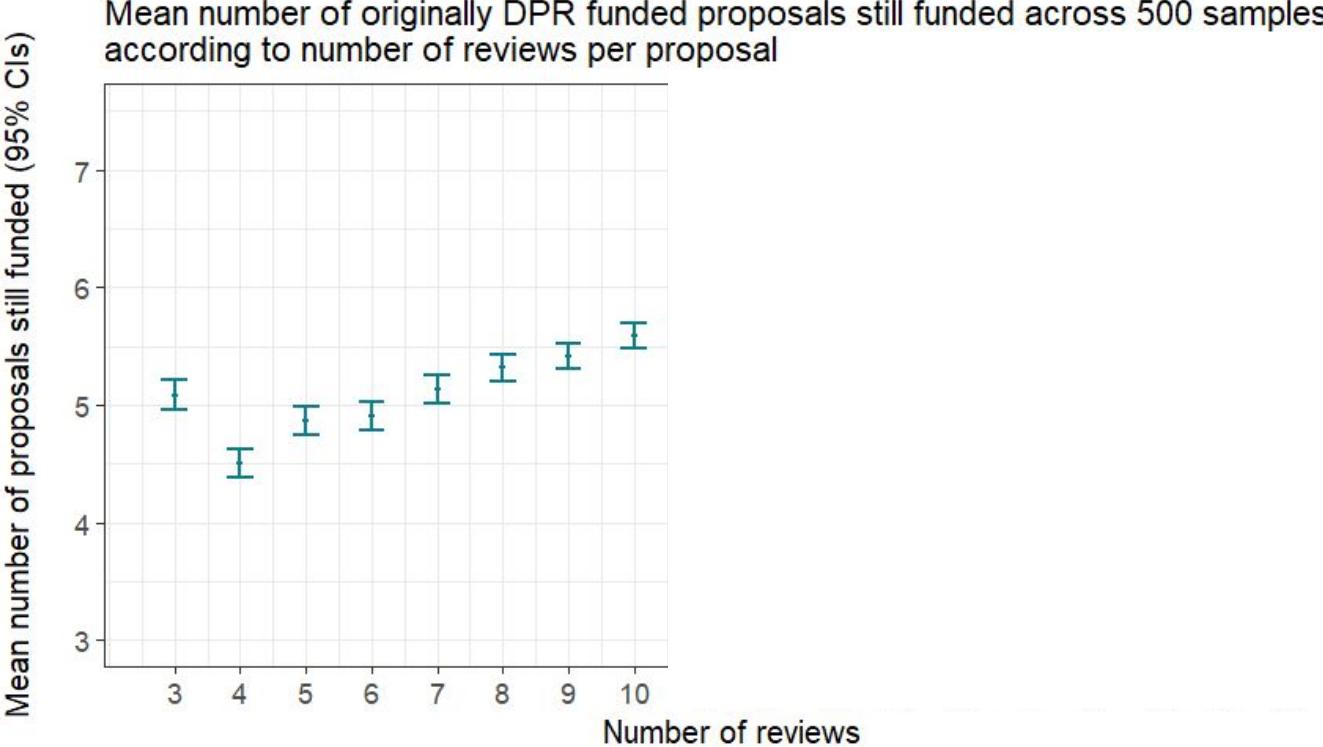
Panel selected proposals are found across the full range of DPR scores



Note: For the purpose of this visualization, where ranks were tied proposals were ordered by standard error.
Any remaining ties were broken by proposal order

These slides:
bit.ly/tomstafford

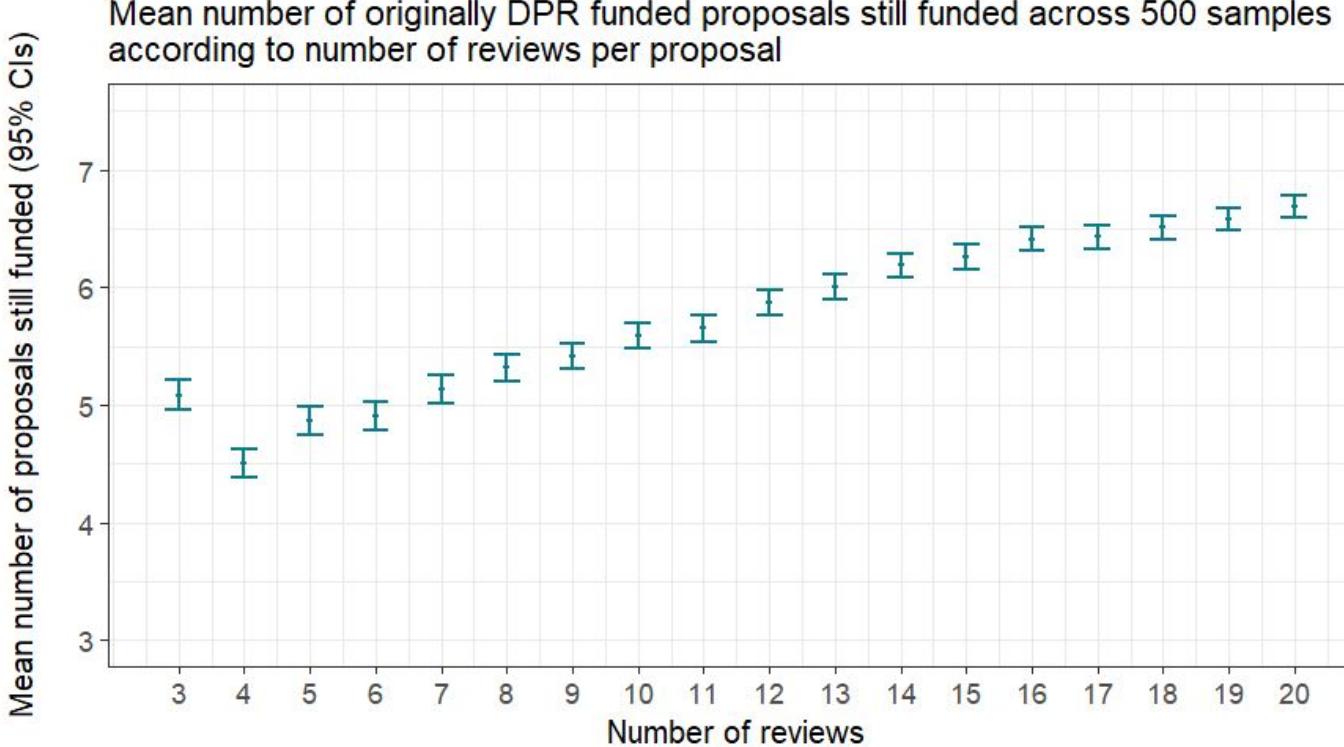
Stability increases with more reviews per proposal but no optimal number of reviews



*500 samples were simulated for each possible number of reviews

These slides:
bit.ly/tomstafford

Stability increases with more reviews per proposal but no optimal number of reviews

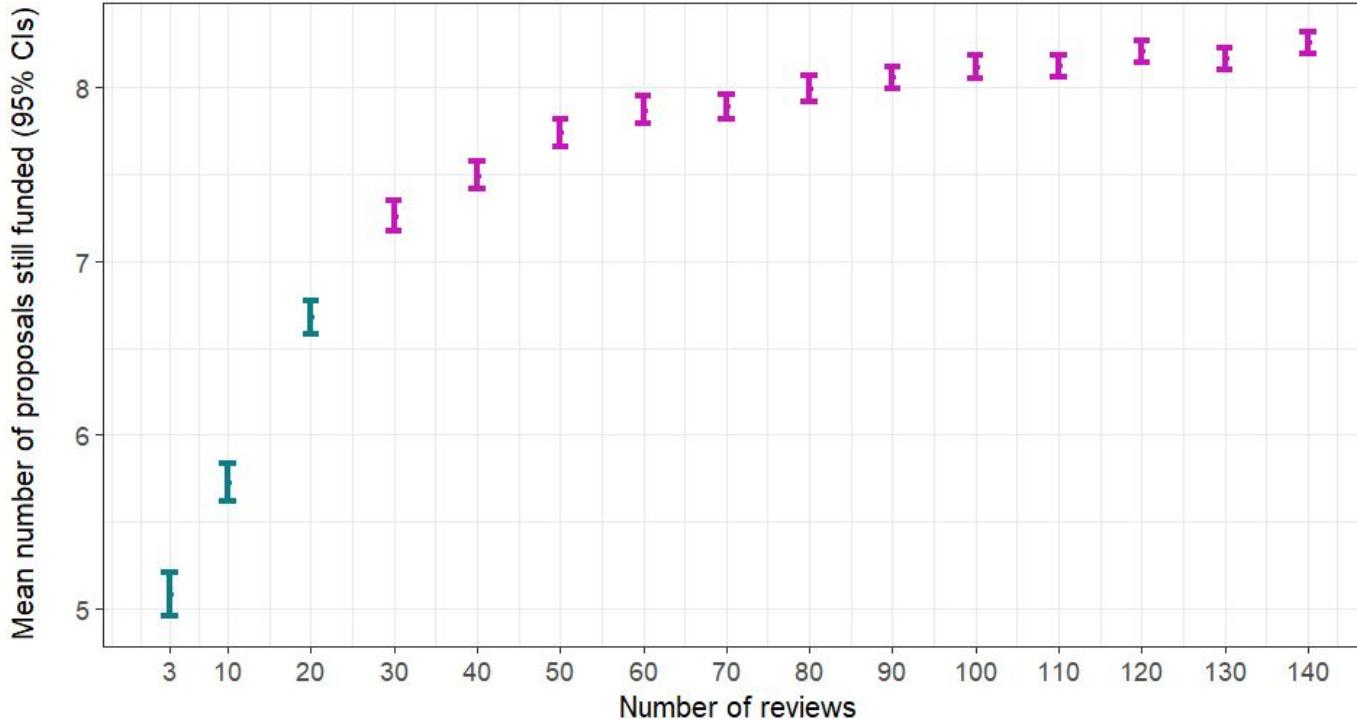


*500 samples were simulated for each possible number of reviews

These slides:
bit.ly/tomstafford

Increasing number of reviews per proposal increases stability to an extent

Mean number of originally DPR funded proposals still funded according to number of reviews per proposal



Do proposals matter?

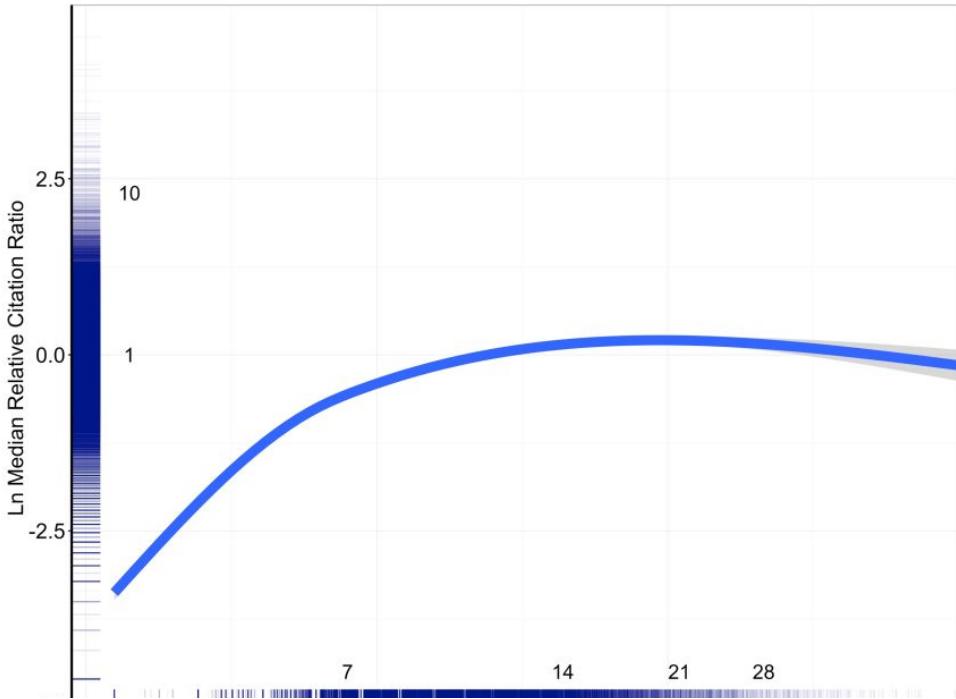
CV + abstract + short proposal

CV + abstract

Abstract

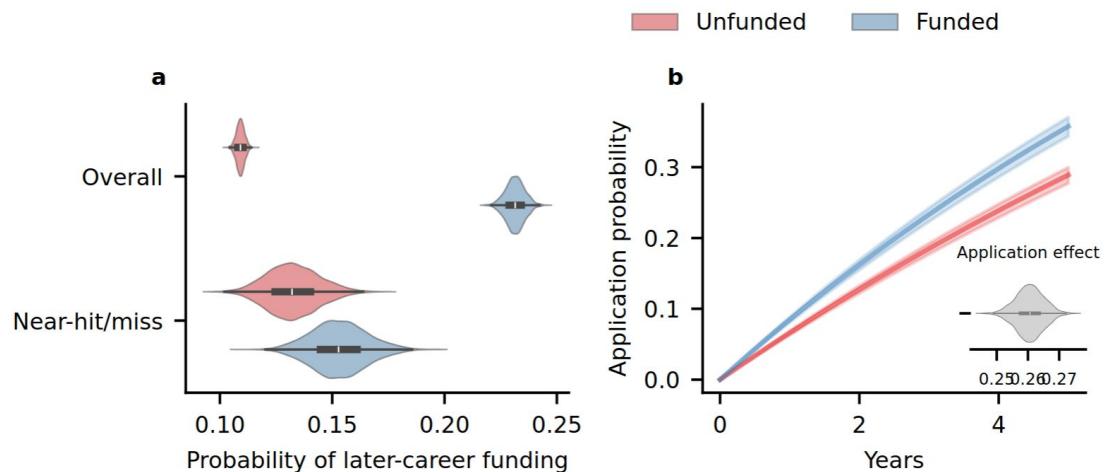
Scientists and funding agencies invest considerable resources in writing and evaluating grant proposals. But do grant proposal texts noticeably change panel decisions in single blind review? We report on a field experiment conducted by The Dutch Research Council (NWO) in collaboration with the authors in an early-career competition for awards of 800,000 euros of research funding. A random half of panelists were shown a CV and only a one-paragraph summary of the proposed research, while the other half were shown a CV and a full proposal. We find that withholding proposal texts from panelists did not detectably impact their proposal rankings. This result suggests that the resources devoted to writing and evaluating grant proposals may not have their intended effect of facilitating the selection of the most promising science.

Simsek, M., de Vaan, M., & van de Rijt, A. (2024). Do grant proposal texts matter for funding decisions? A field experiment. *Scientometrics*, 129(5), 2521-2532. <https://doi.org/10.1007/s11192-024-04968-7>



“In a cohort of over 71,000 unique scientists funded by NIH between 1996 and 2014 we analyzed the association of grant support (as measured by annual GSI) with 3 bibliometric outcomes, maximum Relative Citation Ratio (which arguably reflects a scientist’s most influential work), median Relative Citation Ratio, and annual weighted Relative Citation Ratio (which is more dependent on publication counts). We found that for all 3 measures marginal returns decline as annual GSI increases. Thus, we confirm prior findings of decreasing marginal returns with higher levels of research funding support.”

Lauer, M., Roychowdhury, D., Patel, K., Walsh, R., & Pearson, K. (2017). [Marginal returns and levels of research grant support among scientists supported by the national institutes of health](#). *BioRxiv*, 142554.



Traag, V., Brady, E., Vincent-Lamarre, P., Bidel, F., Lopes Bento, C., Andersen, J. P., & Bloch, C. (2025). The Matthew effect and early-career setbacks in research funding—a replication study.

Elife, 1–45.

<https://doi.org/10.6084/m9.figshare.29302004>



RoRI Working Paper No. 16

The Matthew effect and early-career setbacks in research funding—a replication study

Vincent Traag, Emer Brady, Philippe Vincent-Lamarre, Flavie Bidel, Cindy Lopes-Bento, Jens Peter Andersen and Carter Bloch

June 2025

Produced in partnership with:



Researchers?

Research question: “Are soccer players with dark skin tone more likely to receive red cards from referees than players with light skin tone?”

Four leagues: English, Germany, France and Spain

2012-13 Season

2,034 Players

3,147 Referees*

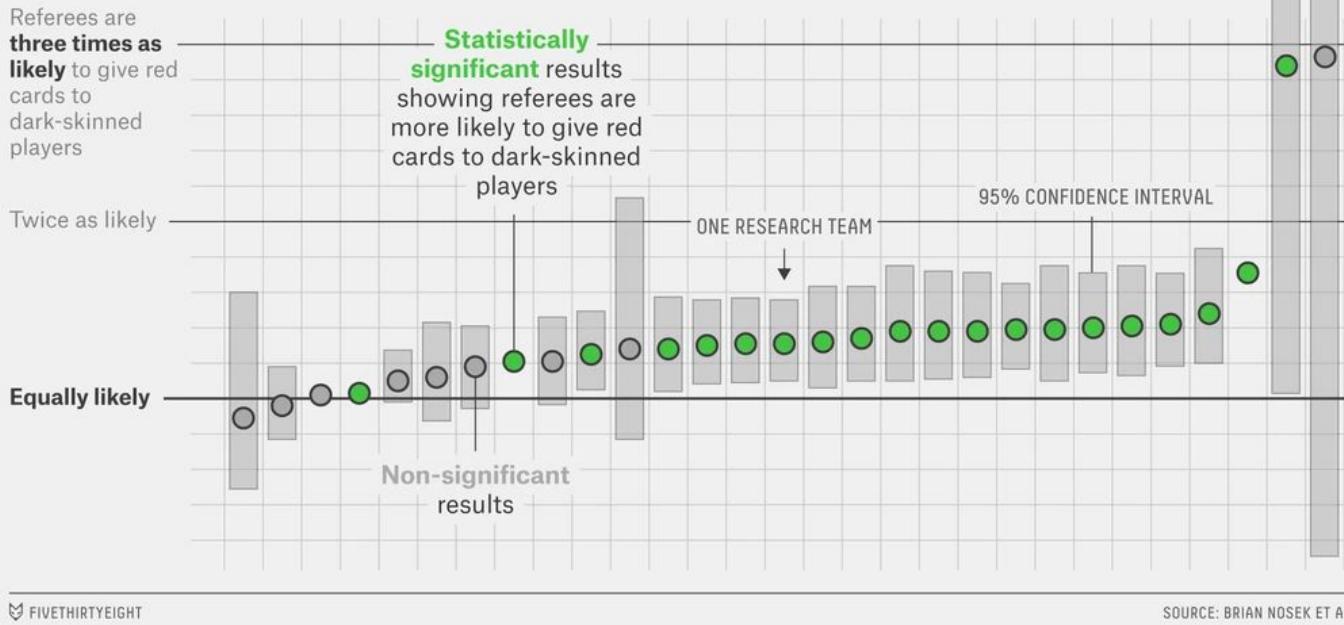


Silberzahn, R. et al (2018). [Many analysts, one dataset: Making transparent how variations in analytical choices affect results](#). *Advances in Methods and Practices in Psychological Science*, 1(3), 337-356.
<https://doi.org/10.1177/2515245917747646>

These slides:
bit.ly/tomstafford

Same Data, Different Conclusions

Twenty-nine research teams were given the same set of soccer data and asked to determine if referees are more likely to give red cards to dark-skinned players. Each team used a different statistical method, and each found a different relationship between skin color and red cards.



“These findings suggest that significant variation in the results of analyses of complex data may be difficult to avoid, even by experts with honest intentions.”

Silberzahn, R. et al (2018). [Many analysts, one dataset: Making transparent how variations in analytical choices affect results](#). *Advances in Methods and Practices in Psychological Science*, 1(3), 337-356.
<https://doi.org/10.1177/2515245917747646>

<https://fivethirtyeight.com/features/science-isnt-broken/#part1>

Iowa Gambling Task



You have won \$100!

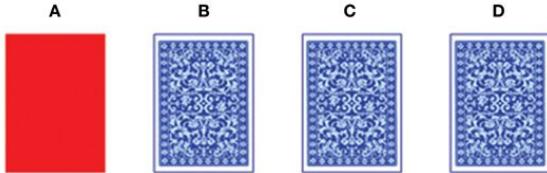


Figure 2, Bull, P. N., Tippett, L. J., & Addis, D. R. (2015). Decision making in healthy participants on the Iowa Gambling Task: new insights from an operant approach. *Frontiers in psychology*, 6, 391.

"To quantify methodological flexibility in the Iowa Gambling Task (IGT), a widely used measure of decision-making in clinical psychology, we conducted a meta-method review of a random sample of 100 studies. We found prevalent under-reporting and wide-spread modifications of procedural details. **We identified 244 distinct scores used to quantify IGT outcomes, 177 of which appeared only once in our sample.**"

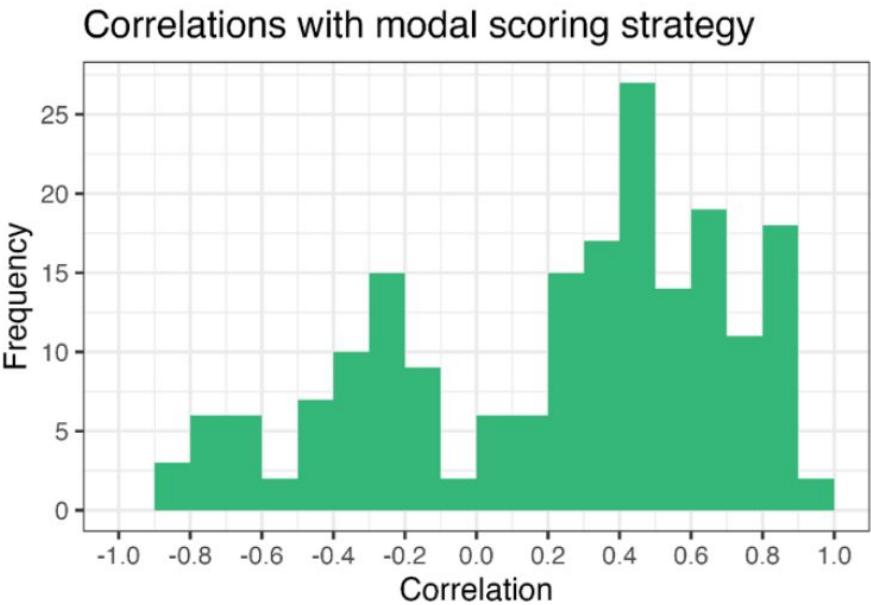
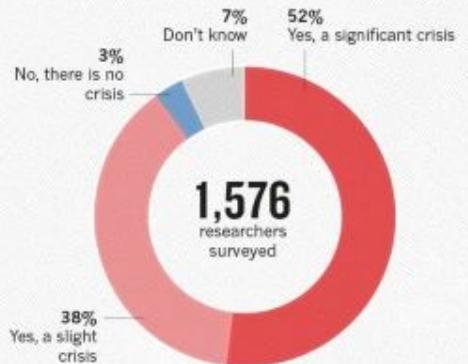


Figure 2, Kulpmann, A. I., Ries, J., Hussey, I., & Elson, M. (2026, January 24). Methodological Flexibility in the Iowa Gambling Task Undermines Interpretability: A Meta-method Review. https://doi.org/10.31234/osf.io/4g3vr_v1

Reproducibility

These slides:
bit.ly/tomstafford

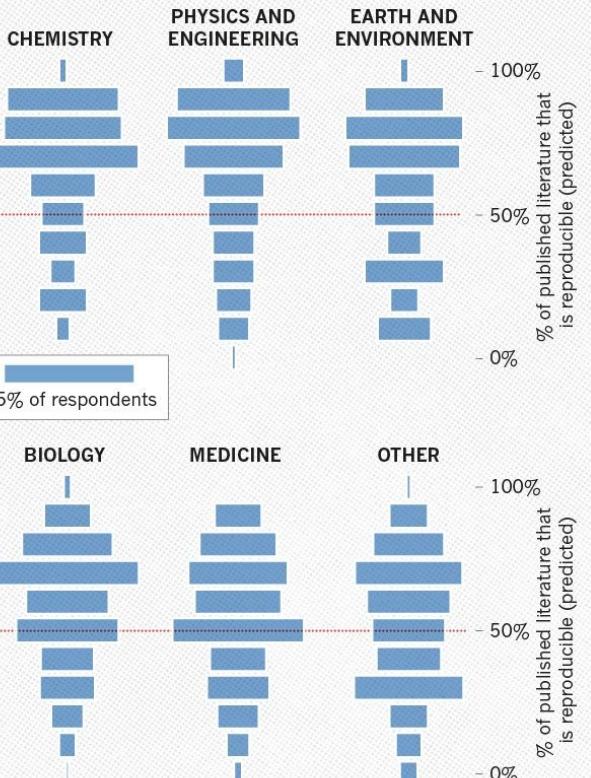
IS THERE A REPRODUCIBILITY CRISIS?



"1,500 scientists lift the lid on reproducibility."
Nature (2016)
<https://doi.org/10.1038/533452a>

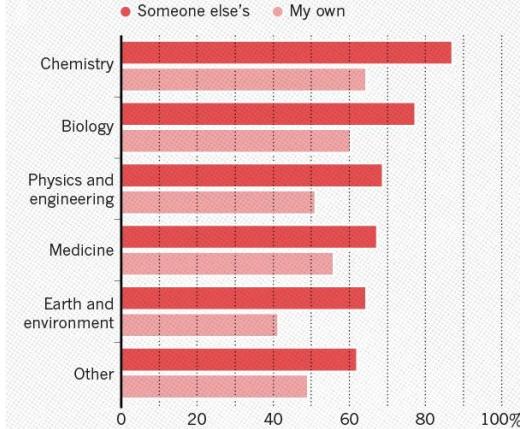
HOW MUCH PUBLISHED WORK IN YOUR FIELD IS REPRODUCIBLE?

Physicists and chemists were most confident in the literature.



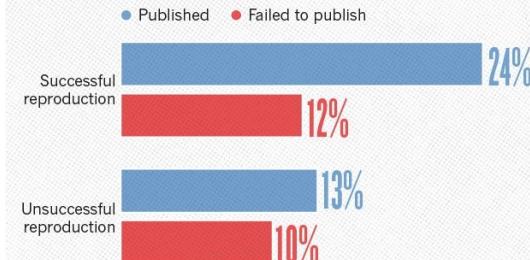
HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

Most scientists have experienced failure to reproduce results.



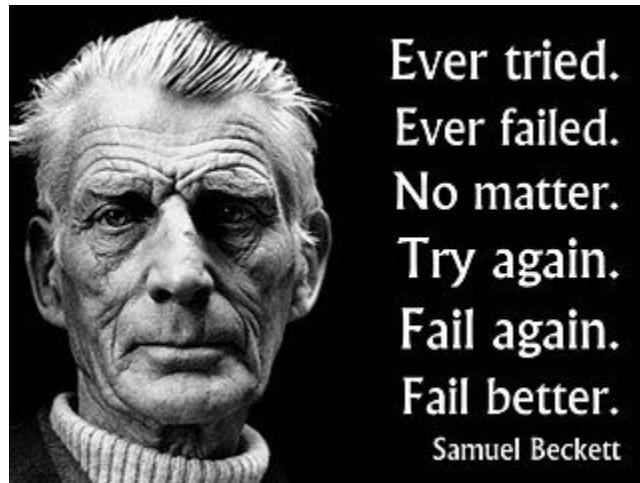
HAVE YOU EVER TRIED TO PUBLISH A REPRODUCTION ATTEMPT?

Although only a small proportion of respondents tried to publish replication attempts, many had their papers accepted.



Number of respondents from each discipline:
 Biology 703, Chemistry 106, Earth and environmental 95,
 Medicine 203, Physics and engineering 236, Other 233

The general lesson of the reproducibility crisis : the errors are free



Ever tried.
Ever failed.
No matter.
Try again.
Fail again.
Fail better.

Samuel Beckett

These slides:
bit.ly/tomstafford



“The UK Reproducibility Network (UKRN) is a national peer-led consortium that aims to ensure the UK retains its place as a centre for world-leading research. We do this by investigating the factors that contribute to robust research, promoting [training activities](#), and disseminating best practice. We seek to understand the factors that contribute to poor research reproducibility and replicability, and develop approaches to counter these, in order to [improve the trustworthiness and quality of research](#)”

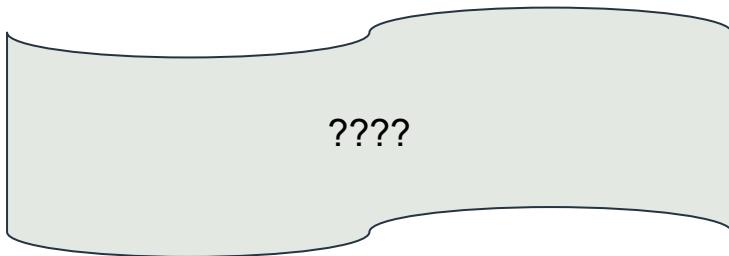


Open Research Programme scope

“to accelerate the uptake of high-quality open research practices”

- Research co-production
- Conducting open research consistent with relevant legal, ethical, and regulatory constraints
- Transparent Qualitative Data Analysis
- Defining the data, code, or other evidence on which your research findings will be based and how this will be managed and shared before the start of data collection and analysis
- Pre-registration of research protocols
- Using open source software created by others
- Creating your own open source software/analysis code to share with others
- Version control of research products
- Your own data analysis is computationally reproducible
- Sharing your own data, code, or other evidence according to the FAIR principles
- Guidelines for recognising the specific substantive contribution of everyone involved in research projects
- Declaring conflicts of interest
- Publishing pre-prints
- Ensuring accepted publications (and other final outputs of research) are Open Access

Process > Output

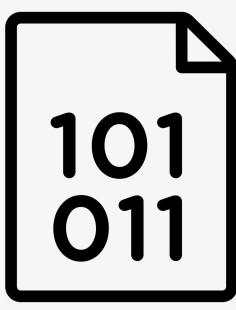


The actual work

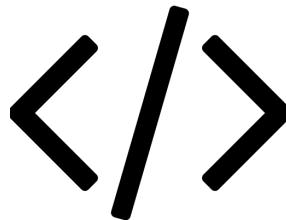


an advert

Your reproducible project



data



code



materials



report

- All shared!

Process quality > output quality

Reproducibility ...?

More repeatable

More reusable, extendable

More translatable

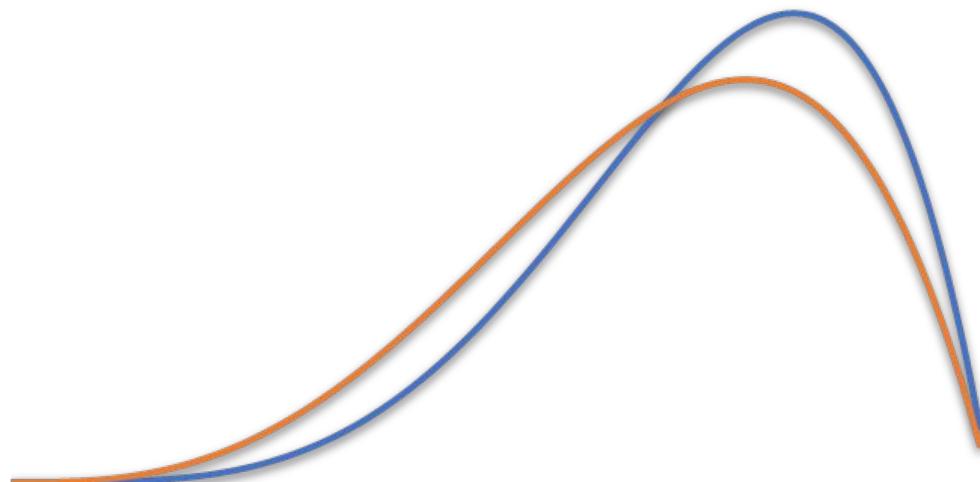
Better documented

Better archived

More secure

These slides:

MacLeod strategy: Research Improvement



These slides:
bit.ly/tomstafford

AI & Research Productivity

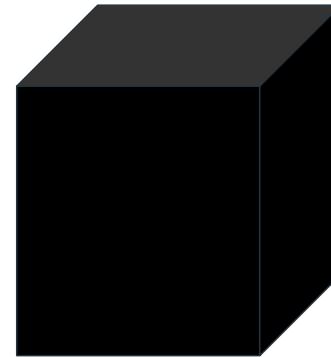
These slides:
bit.ly/tomstafford

AI : Some novelty

AI is novel & changing - our intuitions are poorly calibrated

Specific areas of uncertainty

- Fluency
- Stochasticity
- Hallucinations
- Inscrutability
- Bias



AI : Some constants

Intensification of competition

Profit extraction by big business

Irreproducibility is easy

Productivity

...of what?



"I don't get it. We had all those meetings."

These slides:
bit.ly/tomstafford

Old cures for the new disease

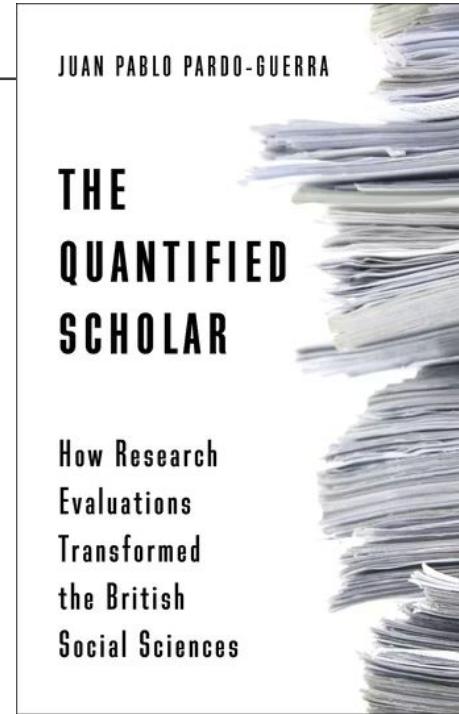
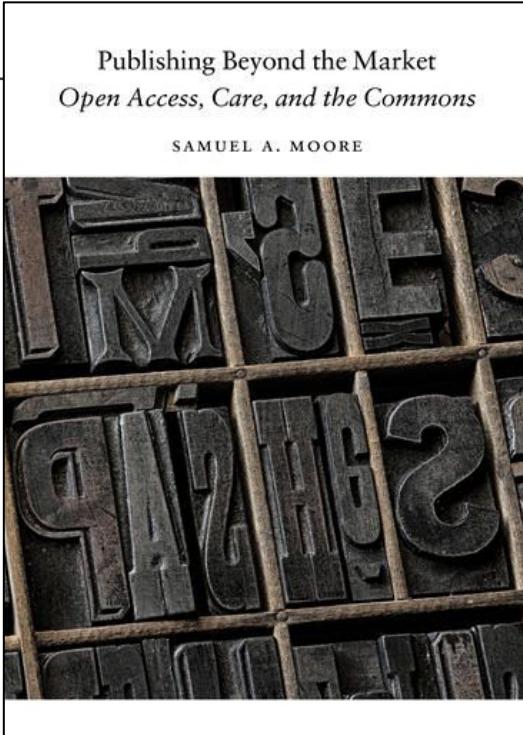
Research improvement

Openness

Process over output

Credibility signals

We can define quality



Points of light

“Action springs not from thought, but from a readiness for responsibility.” — Dietrich Bonhoeffer (attrib.)



These slides:
bit.ly/tomstafford

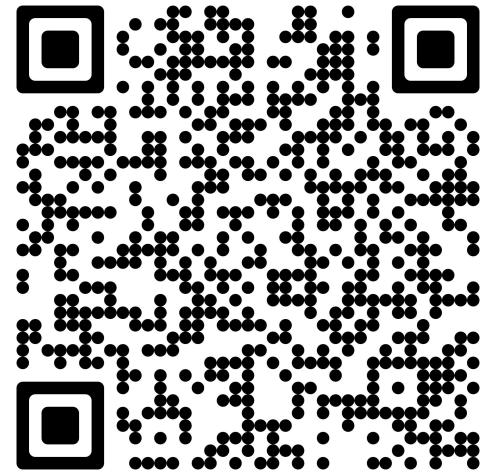
Research Improvement is a system problem

The lesson of the reproducibility crisis is
that the errors are too easy

Process needs to be valued over outputs

Researchers need to lead on defining
quality

The same cures hold in the age of AI



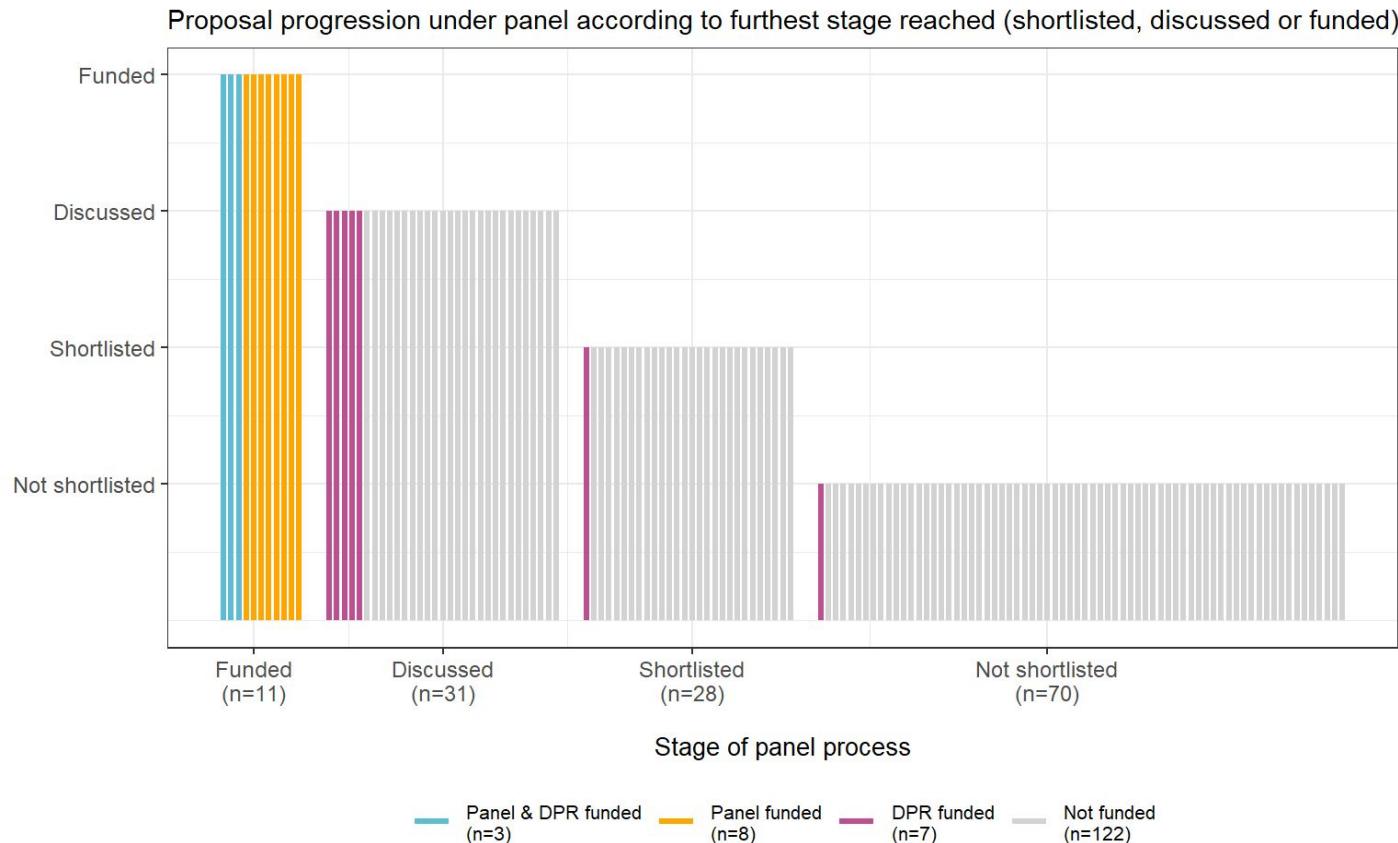
tomstafford.github.io

**RESERVE SLIDES
FOLLOW
(not for show)**

DPR

These slides:
<http://bit.ly/toms-talks>

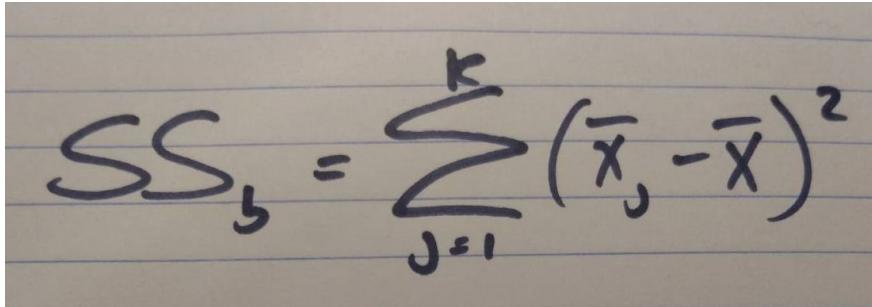
DPR selected proposals are found across all Panel stages



Reproducibility

These slides:
<http://bit.ly/toms-talks>

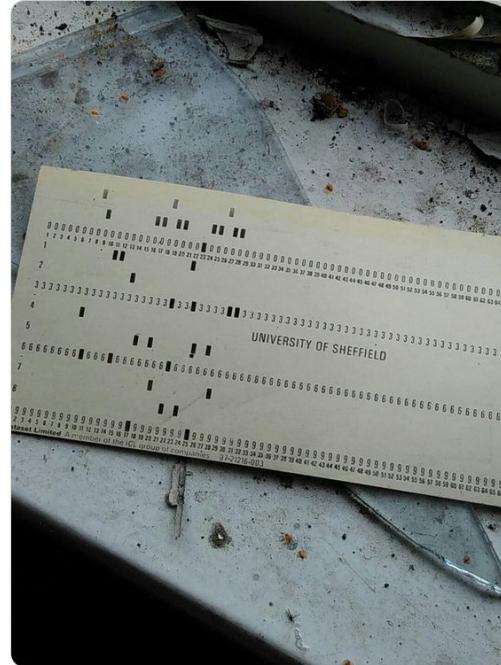
Digital disruption


$$SS_s = \sum_{j=1}^k (\bar{x}_j - \bar{x})^2$$



Tom Stafford
@tomstafford

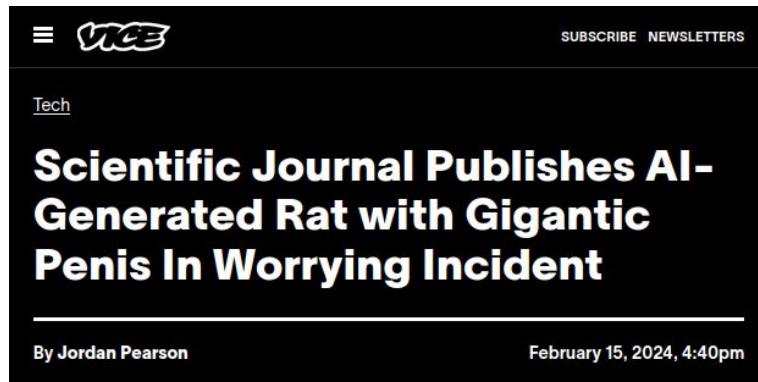
I'm in the attic of the old psychology department. Is this a punch card for an early computer?



12:24 PM · Oct 19, 2016

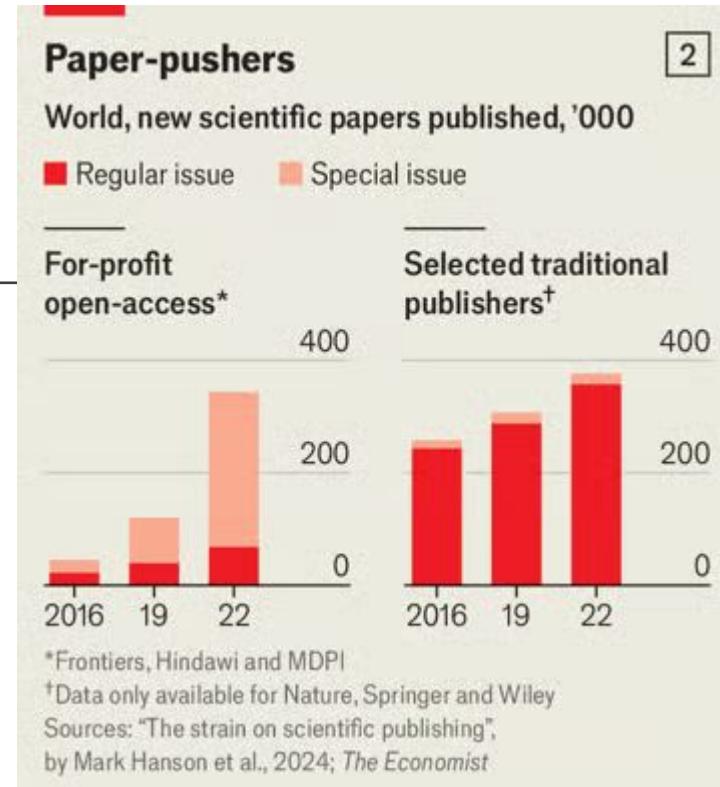
These slides:
bit.ly/tomstafford

Publishers?



A screenshot of a Vice news article. The header reads "Tech". The main title is "Scientific Journal Publishes AI-Generated Rat with Gigantic Penis In Worrying Incident". Below the title, it says "By Jordan Pearson" and "February 15, 2024, 4:40pm". The Vice logo is in the top left corner.

Vice 2024-02-15 [Scientific Journal Publishes AI-Generated Rat with Gigantic Penis In Worrying Incident](#)



Economist 2024-11-20 [Scientific publishers are producing more papers than ever](#)

These slides:
bit.ly/tomstafford

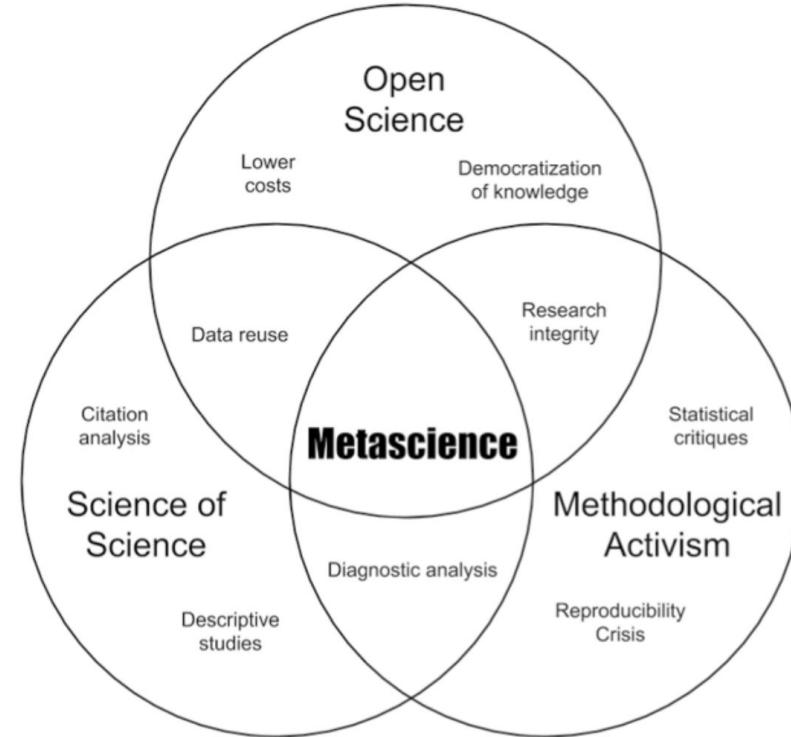
Research on Research

These slides:
<http://bit.ly/toms-talks>

Research on research
Metaresearch

Metascience

Science of science
Scientometrics
Science and Technology Studies



Wilsdon, J., Brasil, A., Waltman, L., & Steyn, B. (2025). The past, present and future of UK metascience.(Annex to '[A Year in Metascience](#)').

Peterson, D., & Panofsky, A. (2023).
Metascience as a scientific social movement.
Minerva, 61(2), 147-174.

\$2 trillion is invested globally in research every year

Source: McKinsey (2020) Building an R&D strategy for modern times

These slides:
<http://bit.ly/tom-talks>

\$2 trillion is invested globally in research every year

Image: pagetutor.com

[What does one TRILLION dollars look like?](#)

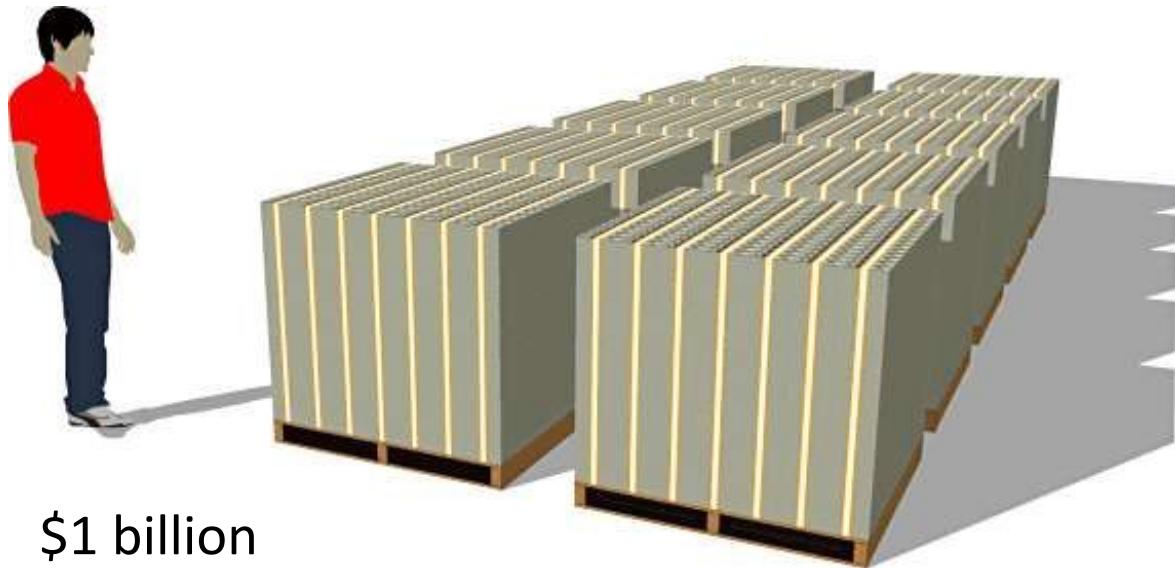


\$10,000

Source: McKinsey (2020) Building an R&D strategy for modern times

These slides:
<http://bit.ly/tom-talks>

\$2 trillion is invested globally in research every year



\$1 billion

Image: pagetutor.com

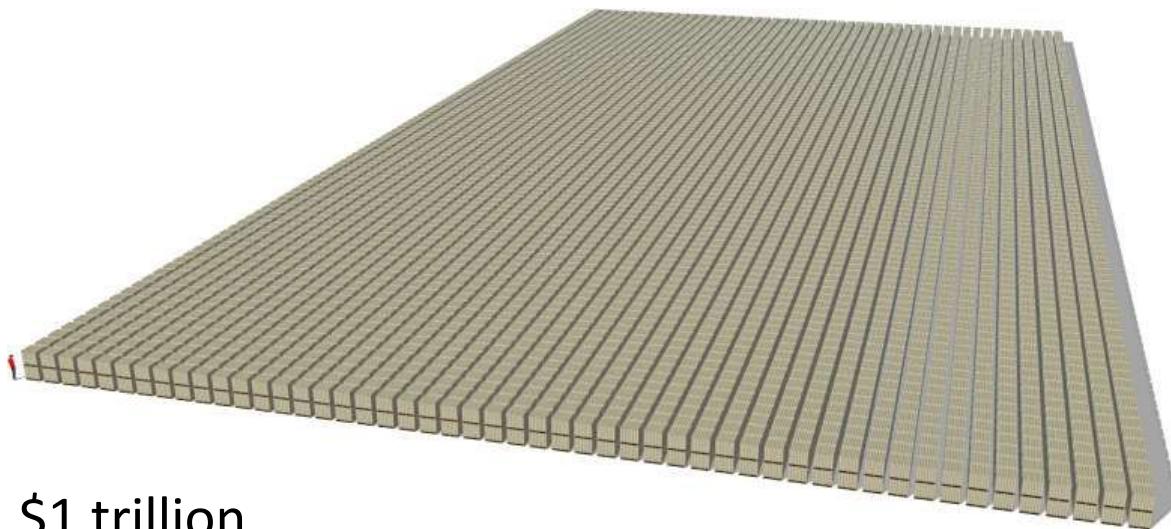
What does one
TRILLION dollars look
like?

These slides:
<http://bit.ly/tom-talks>

\$2 trillion is invested globally in research every year

Image: pagetutor.com

[What does one TRILLION dollars look like?](#)

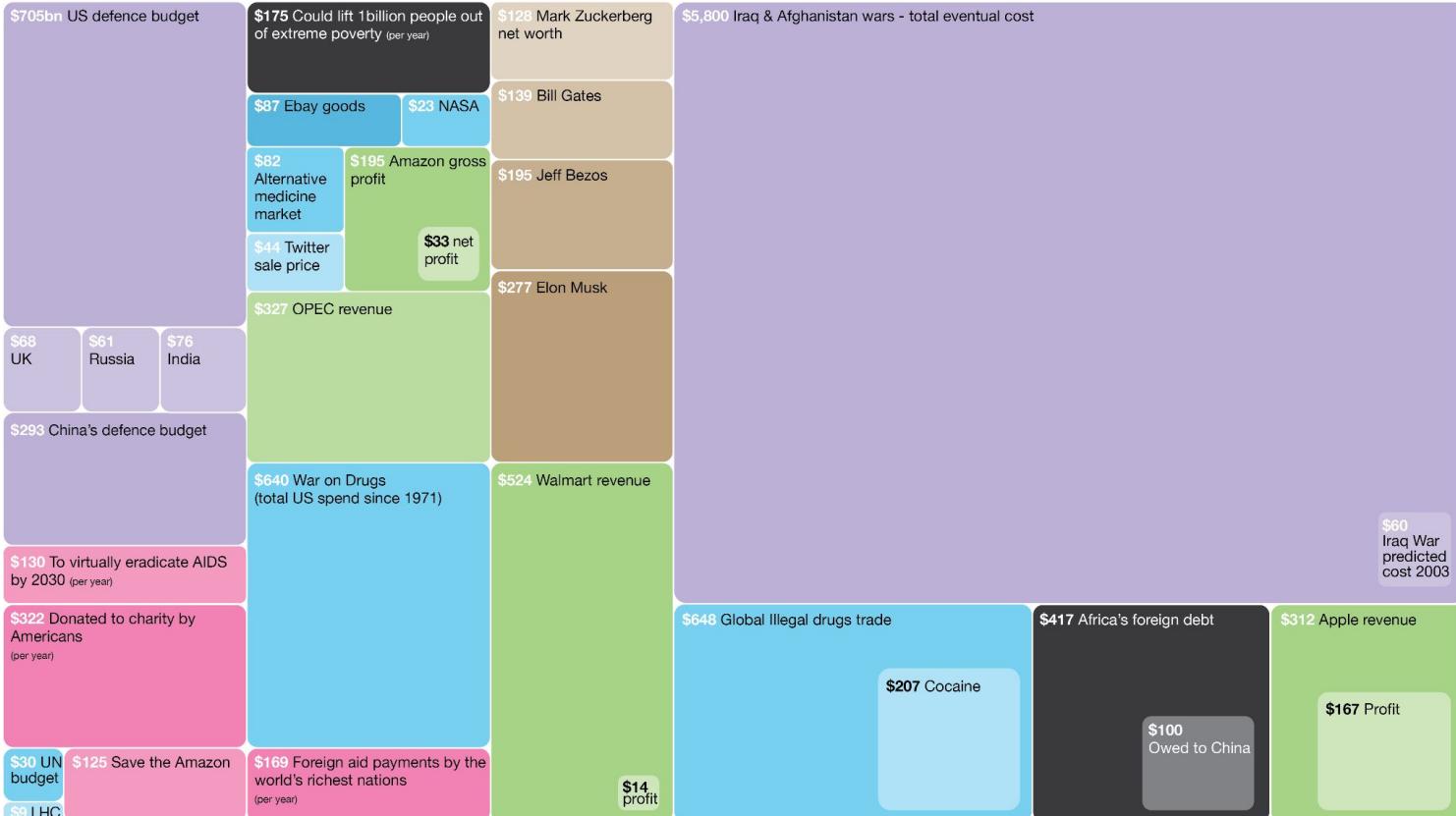


\$1 trillion

These slides:
<http://bit.ly/tom-talks>

The Billion Dollar O Gram

■ accumulating ■ earning ■ fighting ■ giving ■ owing ■ spending

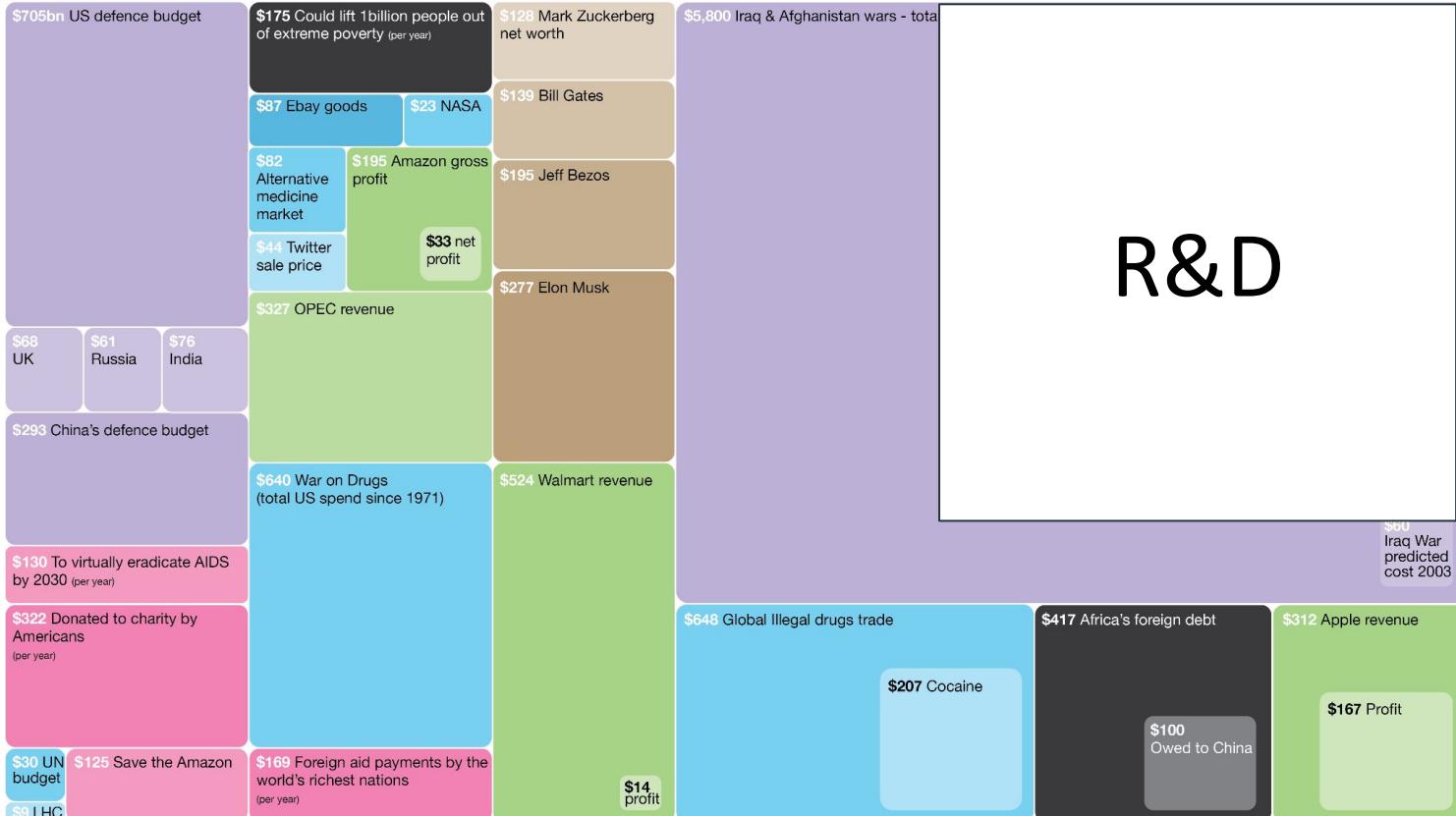


david mccandless
informationisbeautiful

updated Jun 2022 // data bit.ly/billions2022
sources United Nations, Guardian, CNBC, Wikipedia and news reports

The Billion Dollar O Gram

■ accumulating ■ earning ■ fighting ■ giving ■ owing ■ spending



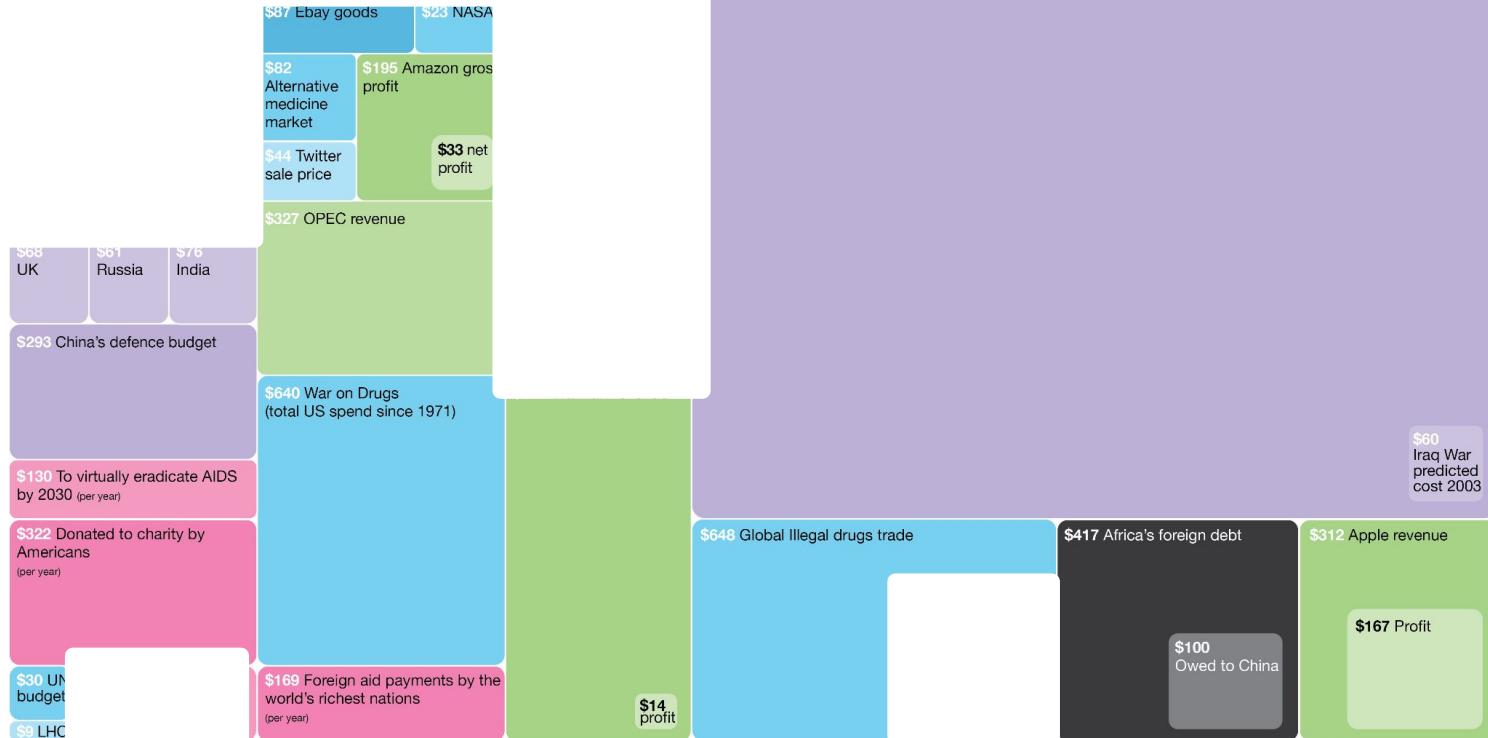
david mccandless
informationisbeautiful

updated Jun 2022 // data bit.ly/billions2022
sources United Nations, Guardian, CNBC, Wikipedia and news reports

The Billion Dollar O Gram

█ accumulating █ earning █ fighting █ giving █ owing █ spending

5,800 Iraq & Afghanistan wars - total eventual cost



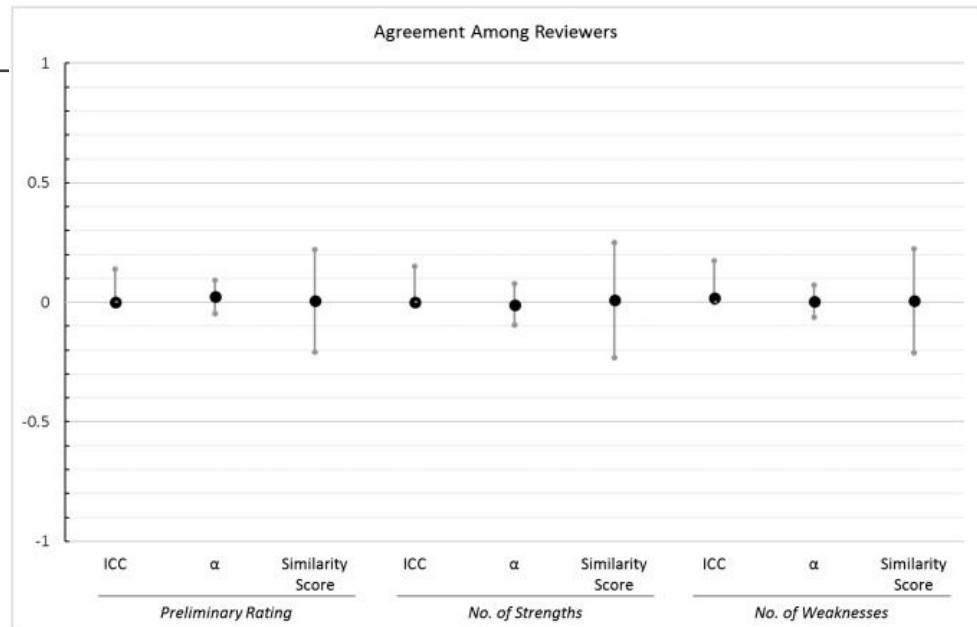
david mccandless
informationisbeautiful

updated Jun 2022 // data bit.ly/billions2022
sources United Nations, Guardian, CNBC, Wikipedia and news reports

How much of this
money and effort is
wasted?

We don't know

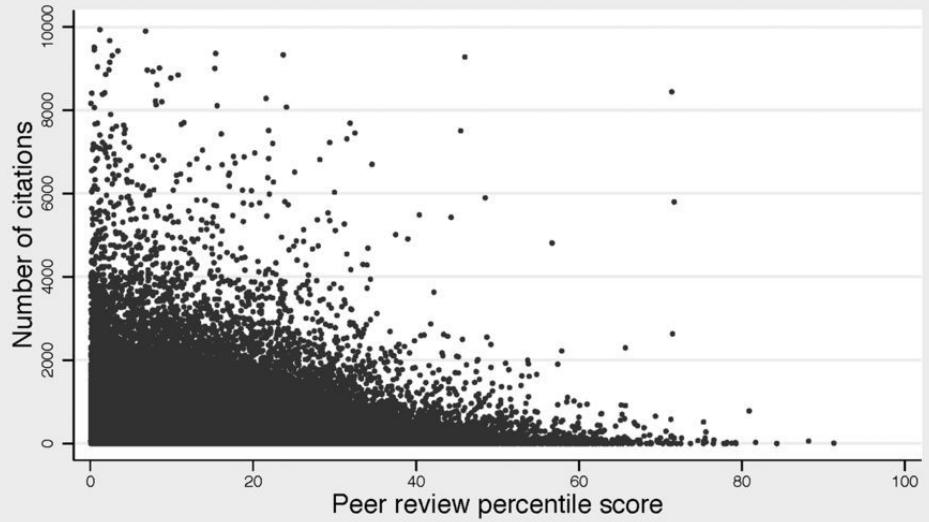
Reviewers may not agree on a common signal



Pier, E. L., Brauer, M., Filut, A., Kaatz, A., Raclaw, J., Nathan, M. J., ... & Carnes, M. (2018). [Low agreement among reviewers evaluating the same NIH grant applications](#). *Proceedings of the National Academy of Sciences*, 115(12), 2952-2957.

Funder ranking is not predictive

130,000 NIH grants, funded 1980-2008



Li, D., & Agha, L. (2015). [Big names or big ideas: Do peer-review panels select the best science proposals?](#) *Science*, 348(6233), 434-438.

ABSTRACT

Formulae display: **MathJax** [?](#)

Peer-review is widely used throughout academia, most notably in the publication of journal articles and the allocation of research grants. Yet peer-review has been subject to much criticism, including being slow, unreliable, subjective and potentially prone to bias. This paper contributes to this literature by investigating the consistency of peer-reviews and the impact they have upon a high-stakes outcome (whether a research grant is funded). Analysing data from 4,000 social science grant proposals and 15,000 reviews, this paper illustrates how the peer-review scores assigned by different reviewers have only low levels of consistency (a correlation between reviewer scores of only 0.2). Reviews provided by 'nominated reviewers' (i.e. reviewers selected by the grant applicant) appear to be overly generous and do not correlate with the evaluations provided by independent reviewers. Yet a positive review from a nominated reviewer is strongly linked to whether a grant is awarded. Finally, a single negative peer-review is shown to reduce the chances of a proposal being funded from around 55% to around 25% (even when it has otherwise been rated highly).

KEYWORDS: [Peer-review](#) [consistency](#) [grant funding](#)

Jerrim, J., & Vries, R. D. (2020). [Are peer-reviews of grant proposals reliable? An analysis of Economic and Social Research Council \(ESRC\) funding applications](#). *The Social Science Journal*, 1-19. These slides: <http://bit.ly/toms-talks>

Review scoring punishes weak aspects

Erosheva et al
(2020) 140k
NIH reviews

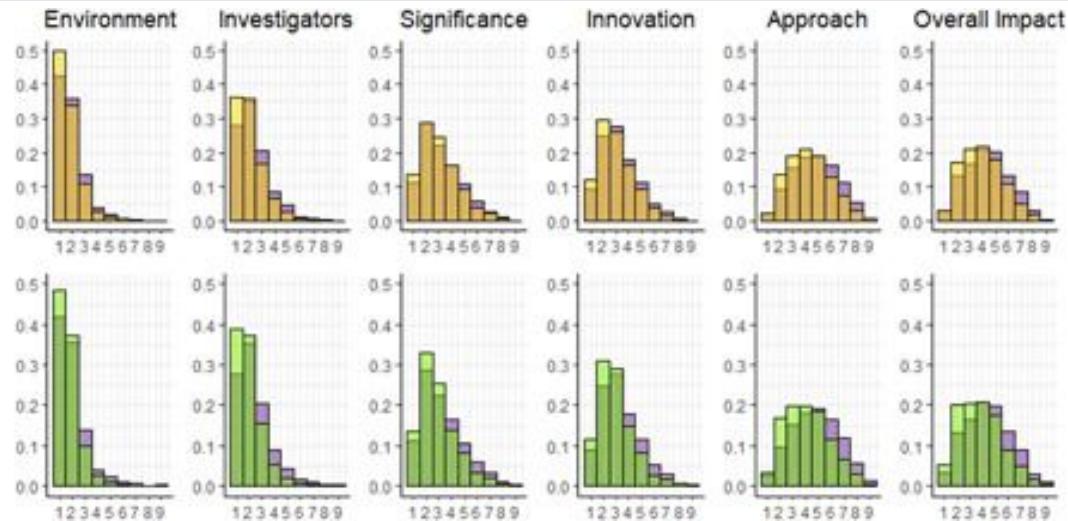
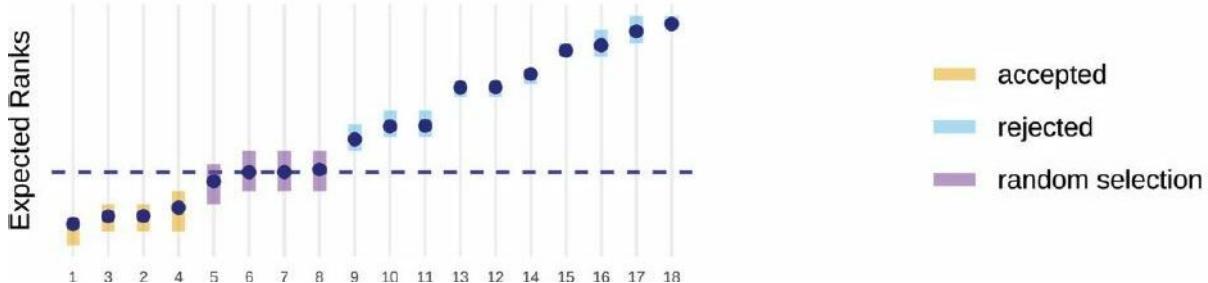


Fig. 2 Frequency histograms for the five preliminary criterion scores and the preliminary overall impact score.

Erosheva, Elena A., Sheridan Grant, Mei-Ching Chen, Mark D. Lindner, Richard K. Nakamura, and Carole J. Lee. "NIH peer review: Criterion scores completely account for racial disparities in overall impact scores." *Science Advances* 6, no. 23 (2020): eaaz4868.

STEM



Rethinking the funding line: random selection at the Swiss National Science Foundation

Marco Bieri and Rachel Heyard



Heyard, R., Ott, M., Salanti, G., & Egger, M. (2022).
[Rethinking the Funding Line at the Swiss National Science Foundation: Bayesian Ranking and Lottery](#).
Statistics and Public Policy, (just-accepted), 1-27.

These slides:
<http://bit.ly/toms-talks>

Signal - what to reviewers judge?

- “The strongest evidence around effectiveness indicates a bias against innovative research”

Guthrie, S., Ghiga, I., & Wooding, S. (2017). What do we know about grant peer review in the health sciences?. *F1000Research*, 6.

Lee, Carole J. “Commensuration Bias in Peer Review.” *Philosophy of Science* 82 (2015): 1272-1283

Does name matter?

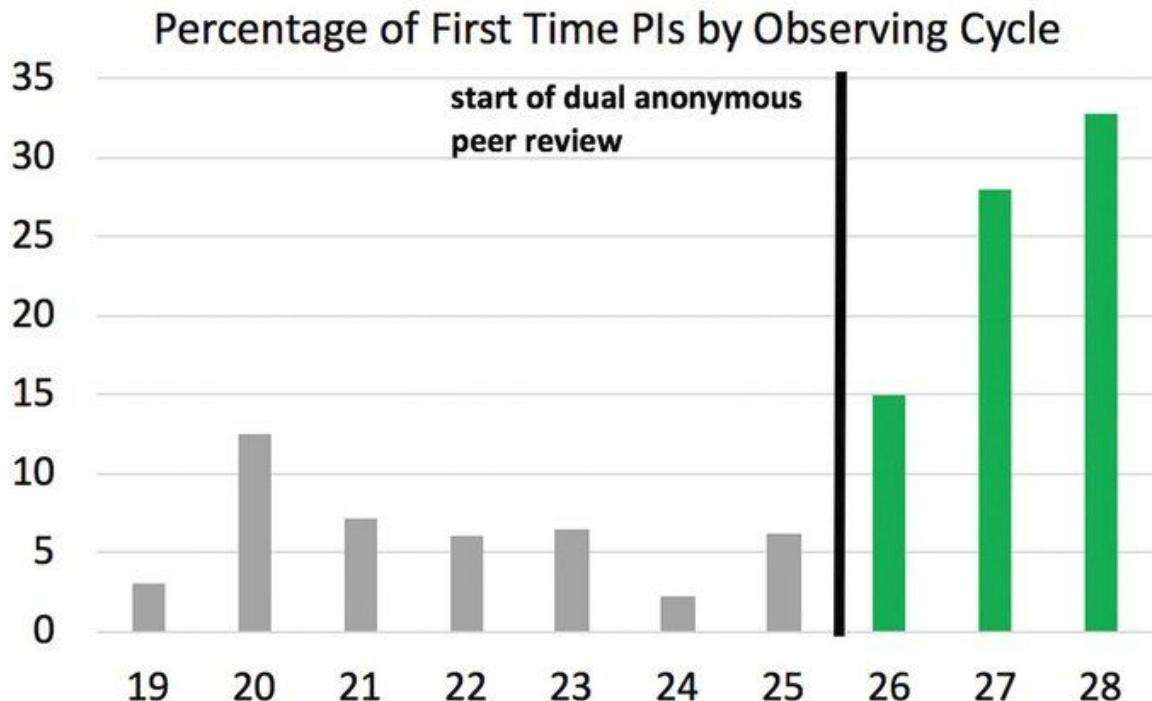


Image: Rachel Osten [Space Telescope Science Institute](#). See also

Strolger, L., & Natarajan, P. (2019). [Doling out Hubble time with dual-anonymous evaluation](#). *Physics Today*.

These slides:
<http://bit.ly/toms-talks>

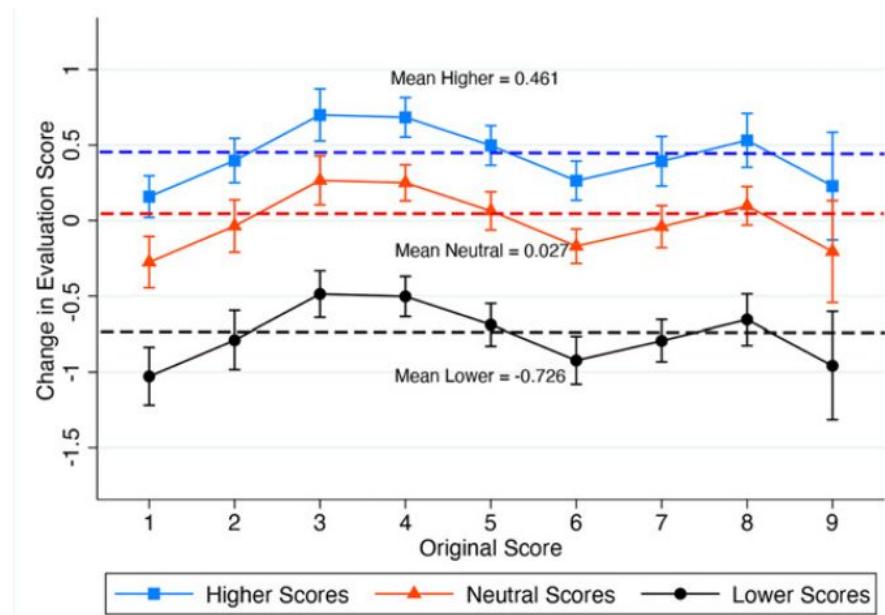
Experimentally varied what reviewers saw from others - positive or negative reviews

Reviewers revised their scores down more than up, when they saw others' evaluations

Lane, J. N., Teplitskiy, M., Gray, G., Ranu, H., Menietti, M., Guinan, E., & Lakhani, K. R. (2021).

Conservatism Gets Funded? A Field Experiment on the Role of Negative Information in Novel Project Evaluation. Management Science.

Figure 2. Margins Plot of Change in Evaluation Score and Treatment Scores Valence by Original Score with 95% CIs



These slides:
<http://bit.ly/toms-talks>

What would work better?



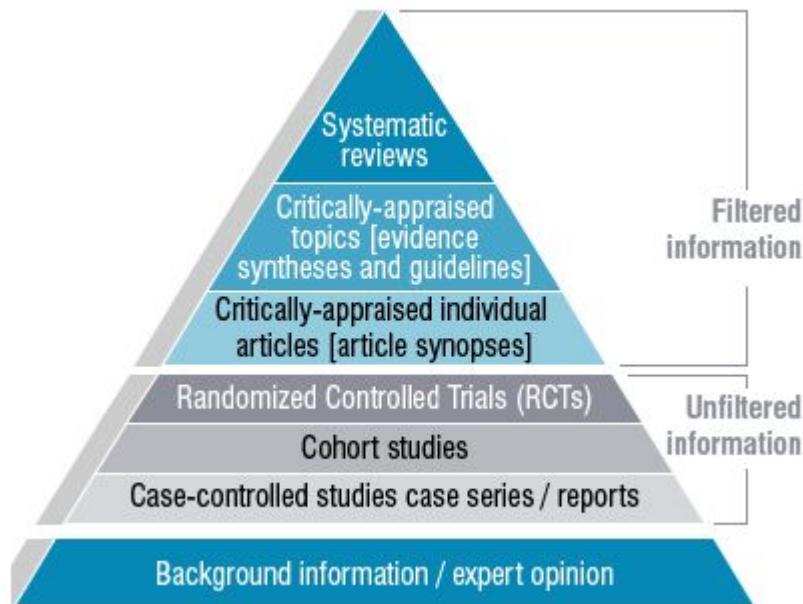
We don't know

Nielsen, M., & Qui, K. (2022). A vision of metascience: An engine of improvement for the social processes of science. *The Science++ Project*.
<https://scienceplusplus.org/metascience/>

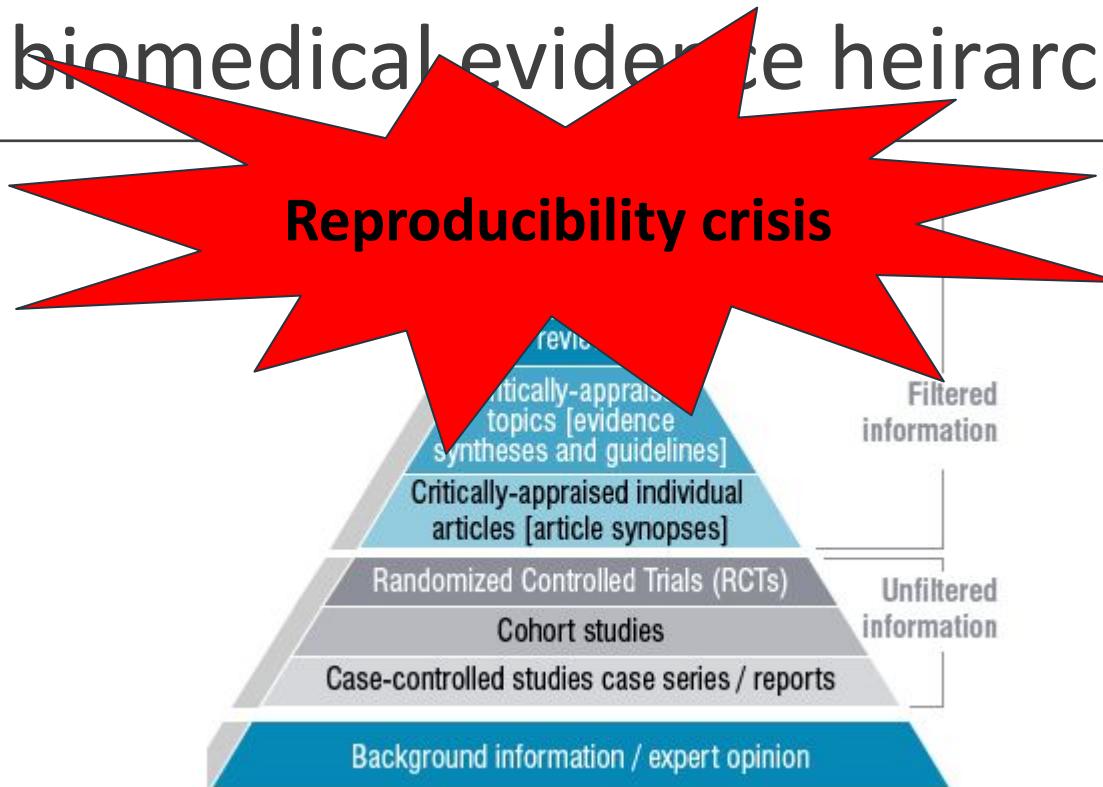
Reproducibility Crisis

These slides:
<http://bit.ly/toms-talks>

The biomedical evidence hierarchy

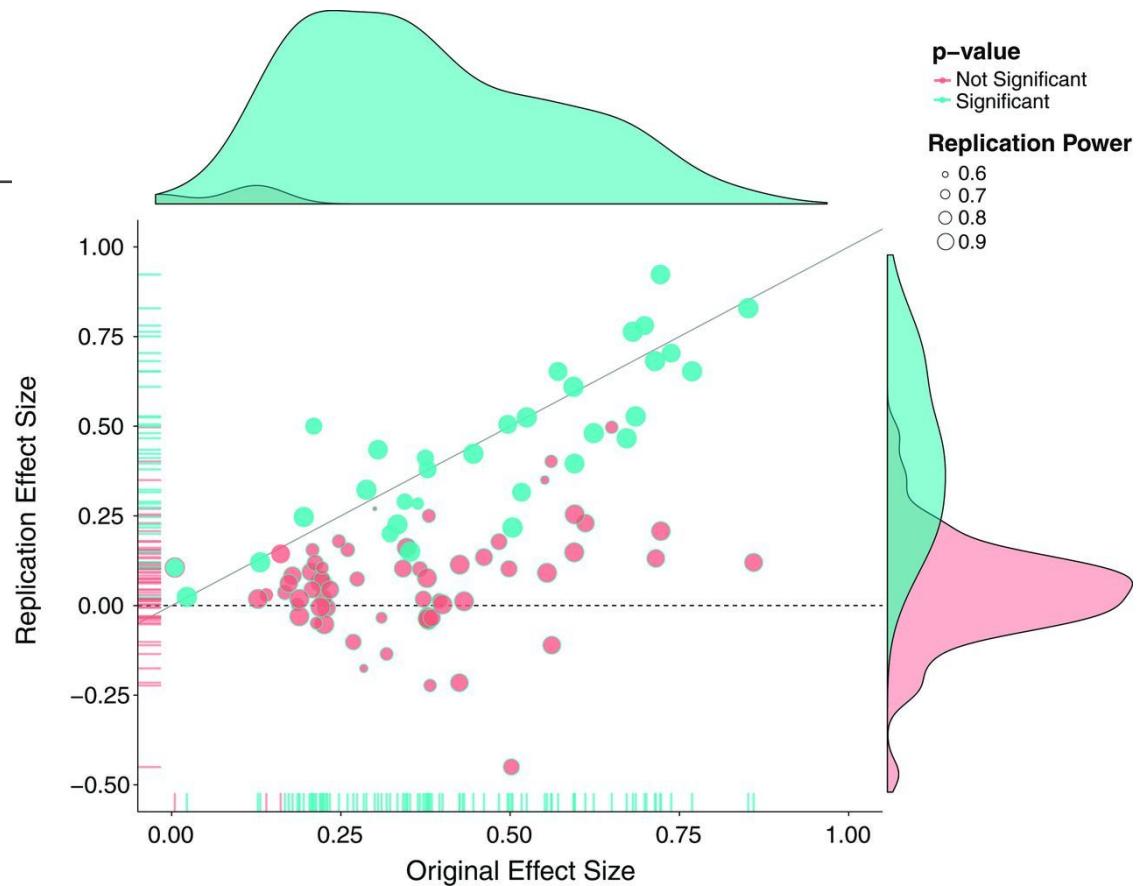


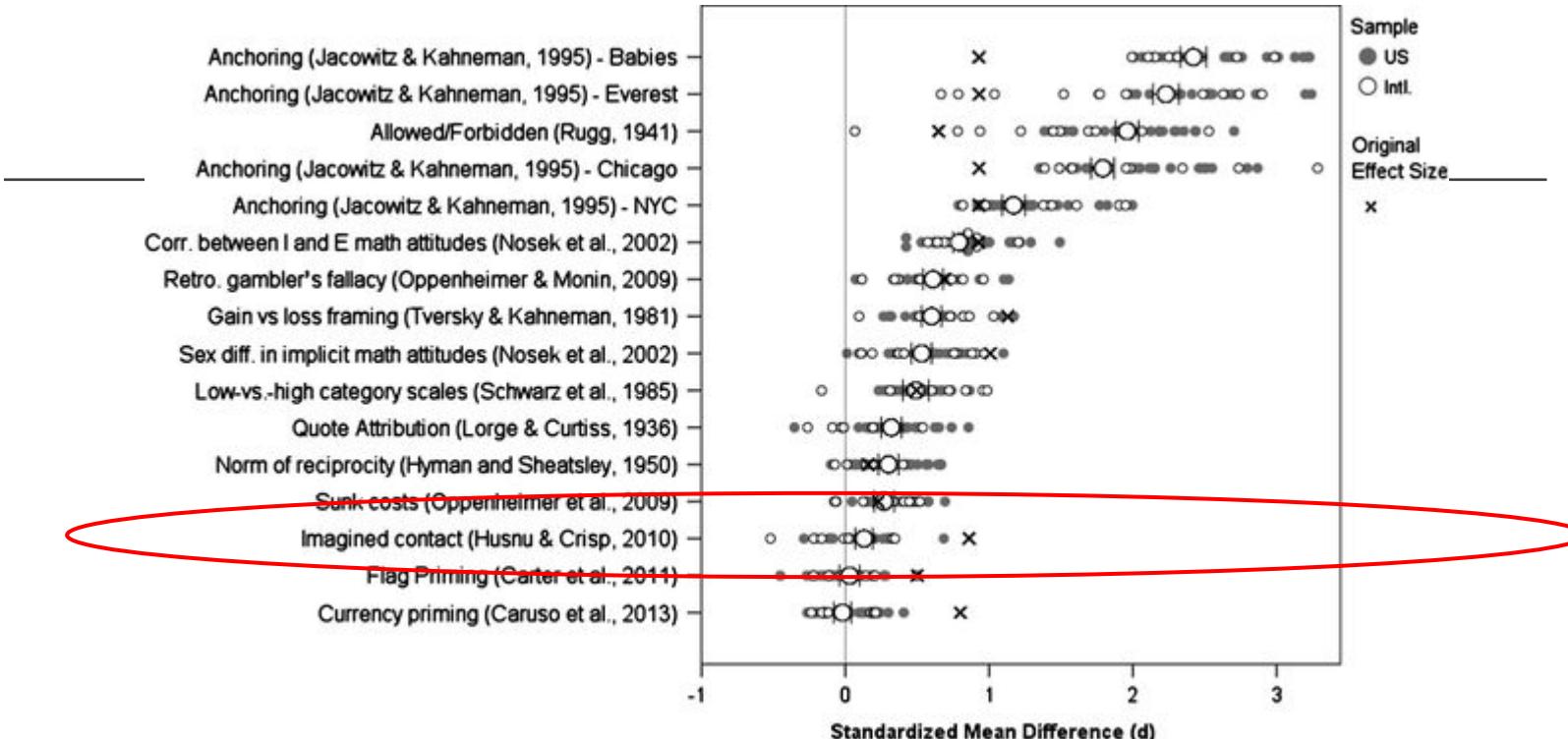
The biomedical evidence hierarchy



Replication of 100 studies from 2008 articles of three important psychology journals: *Psychological Science (PSCI)*, *Journal of Personality and Social Psychology (JPSP)*, and *Journal of Experimental Psychology: Learning, Memory, and Cognition (JEP:LMC)*

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716.





Klein Richard, A., Ratliff, K. A., Vianello, M., Adams Jr, R. B., Bahník, Š., Bernstein, M. J., ... & Cemalcilar, Z. (2014). Investigating Variation in Replicability: A "Many Labs" Replication Project. *Social Psychology*, 45(3), 142-152.

Ego depletion

Ego Depletion: Is the Active Self a Limited Resource?

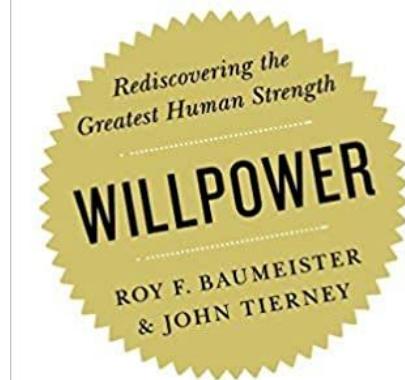
Roy F. Baumeister, Ellen Bratslavsky, Mark Muraven, and Dianne M. Tice
Case Western Reserve University

Choice, active response, self-regulation, and other volition may all draw on a common limited resource. In Experiment 1, people who forced themselves to eat radishes instead of ice cream subsequently quit faster on unsolvable puzzles than people who had not had to overeat. In Experiment 2, making a meaningful personal choice rather than a trivial one caused behavior to cause a similar decrement in persistence. In Experiment 3, self-control was used in a subsequent drop in performance of solvable anagrams. In Experiment 4, an effort to exert self-control made people more passive (i.e., more prone to favor the passive alternative). These results suggest that the self's capacity for active volition is limited. The findings also suggest that seemingly different, unrelated acts share a common resource.

6000+ citations!

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource?. *Journal of personality and social psychology*, 74(5), 1252.

NEW YORK TIMES BESTSELLER



"An immensely rewarding book, filled with ingenious research, wise advice and insightful reflections on the human condition."
—STEVEN PINKER, THE NEW YORK TIMES BOOK REVIEW

Google scholar hits “ego depletion”

2020 - 3,400 papers

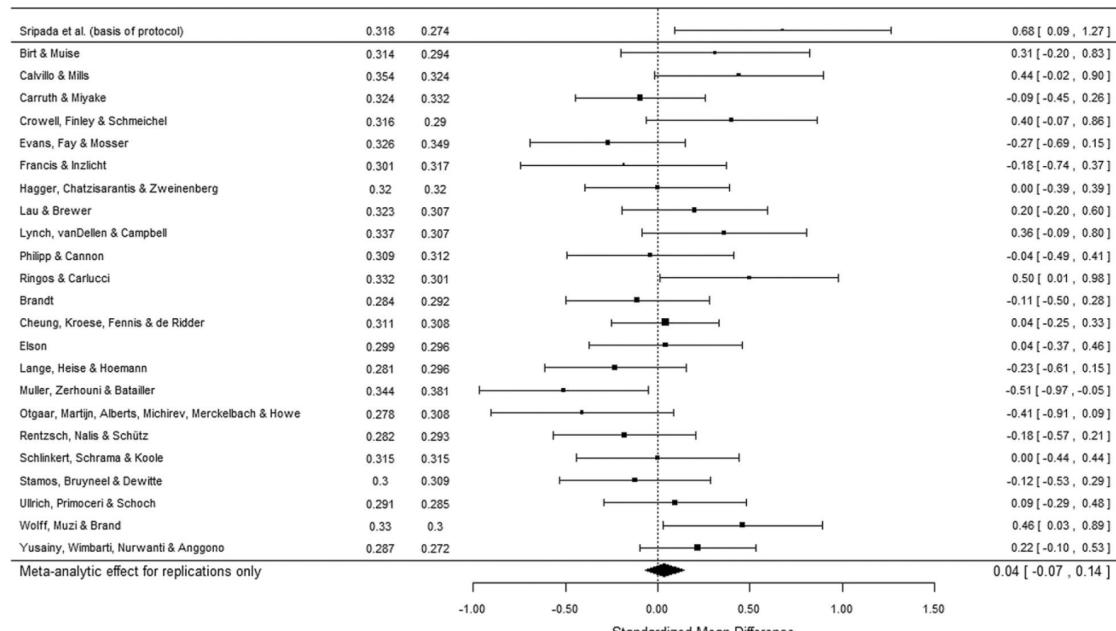
2010 meta-analysis, $d = 0.62$ ($n = 198$)

Hagger, M. S., Wood, C., Stiff, C., and Chatzisarantis, N. L. D. (2010). Ego depletion and the strength model of self-control: a meta-analysis. *Psychol. Bull.* 136, 495–525. doi: 10.1037/a0019486

Ego depletion many labs 1

Multiple laboratories ($k = 23$, total $N = 2,141$) conducted replications of a standardized ego-depletion protocol based on a sequential-task paradigm by Sripada et al. Meta-analysis of the studies revealed that the size of the ego-depletion effect was small with 95% confidence intervals (CIs) that encompassed zero ($d = 0.04$, 95% CI $[-0.07, 0.15]$).

Hagger, M. S., Chatzisarantis, N. L., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., ... & Calvillo, D. P. (2016). A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science*, 11(4), 546-573.



Edo depletion many labs 2

"A preregistered, multi-lab project (N=3531) to assess the size & robustness of ego depletion.. d=0.06..Bayesian..found 4x more likely under null"

Vohs, K., Schmeichel, B., Lohmann, S., Gronau, Q. F., Finley, A. J., whenyoup, I., ... Albarracín, D. (2021). A Multi-Site Preregistered Paradigmatic Test of the Ego Depletion Effect. *Psychological Science*.
<https://doi.org/10.1177/0956797621989733>

Reproducibility

These slides:
<http://bit.ly/toms-talks>

Better research

more....

leads to...

“Reproducible
Research”?

Replicable (reliable
effects)

Repeatable
(transparent methods)

Scalable &
Translatable

Trusted
(audit ready,
inclusive,
co-produced?)

Faster and further
collaboration

Knowledge exchange

Reuse & extension

Greater dissemination /
citation

Credibility Revolution

Robust statistics

Open Data

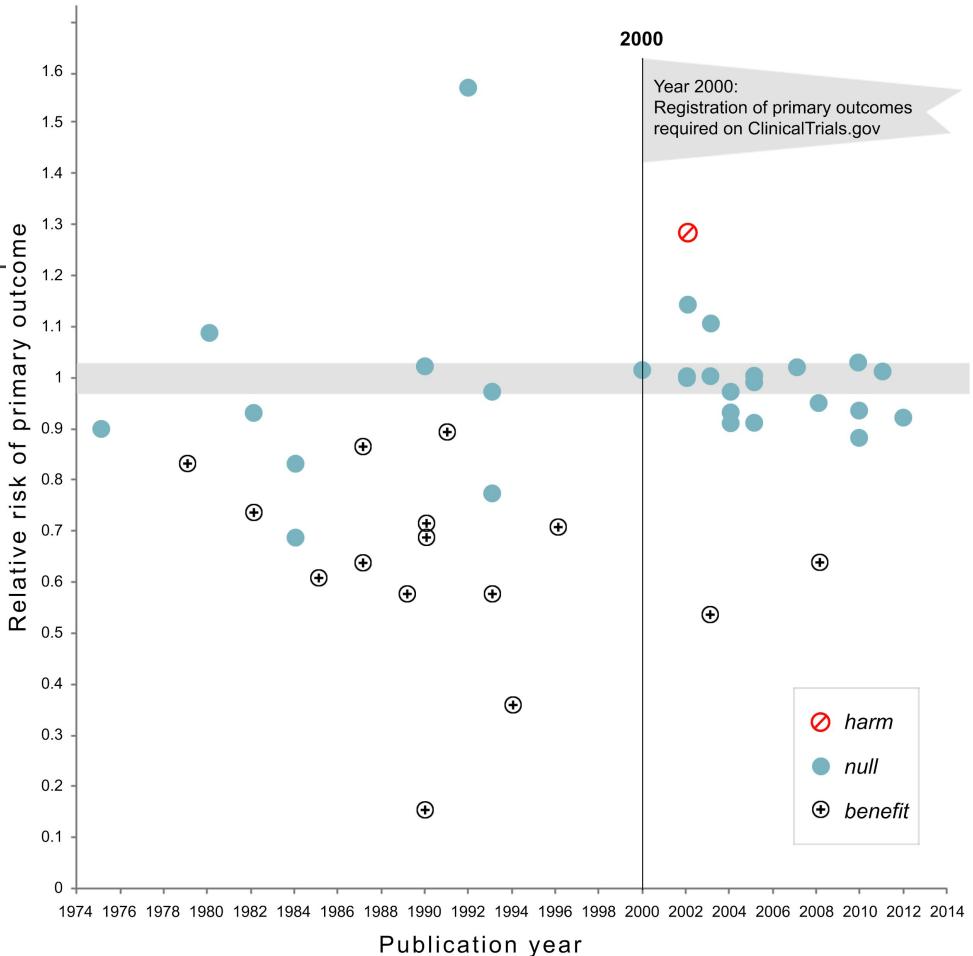
Pre-registration

Causal Inference

Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., Percie du Sert, N., ... & Ioannidis, J. (2017). A manifesto for reproducible science. *Nature human behaviour*, 1(1), 1-9. <https://doi.org/10.1038/s41562-016-0021>

Preregistration

Kaplan, R. M., & Irvin, V. L. (2015). Likelihood of null effects of large NHLBI clinical trials has increased over time. *PLoS One*, 10(8), e0132382.
<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0132382>



Principles vs practices, an example

PRACTICE: preregistration

publicly archived trial design, outcome measures and statistical models

EXCEPTIONS

exploratory research, feasibility/pilot studies,

PRINCIPLE: avoid (self) deception in results narration

formal record of what you intended when you started

alertness to researcher bias

These slides:

Principles vs practices, example II

PRACTICE: open data

public, permanent, archive of study data

EXCEPTIONS:

commercially or clinically sensitive data, DOUS (data of unusual size), lack of consent

PRINCIPLE: reproducibility, audit, reuse

share post-processed data at earlier possible point

share metadata for discovery, synthesis

These slides: