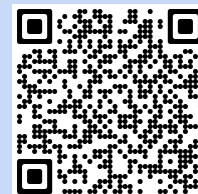# Experiments with the funding of research

Tom Stafford,
Professor of Cognitive Science & University Research Practice Lead
Senior Research Fellow, Research on Research Institute

t.stafford@sheffield.ac.uk

2024-09-09

tomstafford.github.io

TODAY

0. Research funding is important
1. Better evidence in funding
2. Funding as a decision problem
3. RoRI's AFIRE project

Research funding is big business

$2 trillion is invested globally in research every year

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

RoЯi RESEARCH ON RESEARCH INSTITUTE

$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
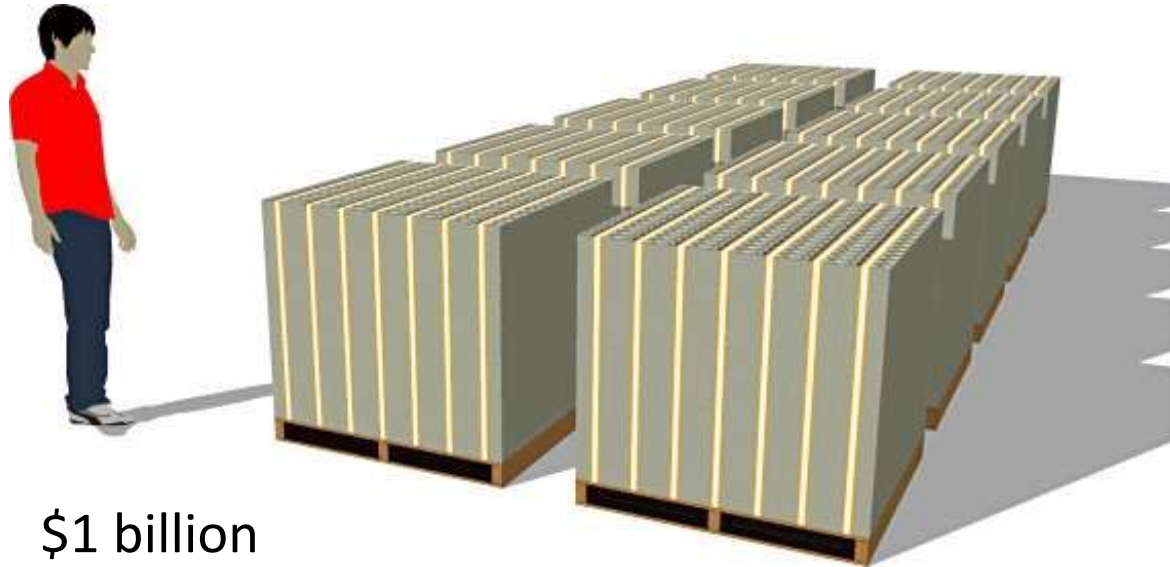
RoЯi RESEARCH ON RESEARCH INSTITUTE

$2 trillion is invested globally in research every year

Image: pagetutor.com

What does one TRILLION dollars look like?

$1 billion

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

RoЯI RESEARCH ON RESEARCH INSTITUTE

$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

RoЯi RESEARCH ON RESEARCH INSTITUTE
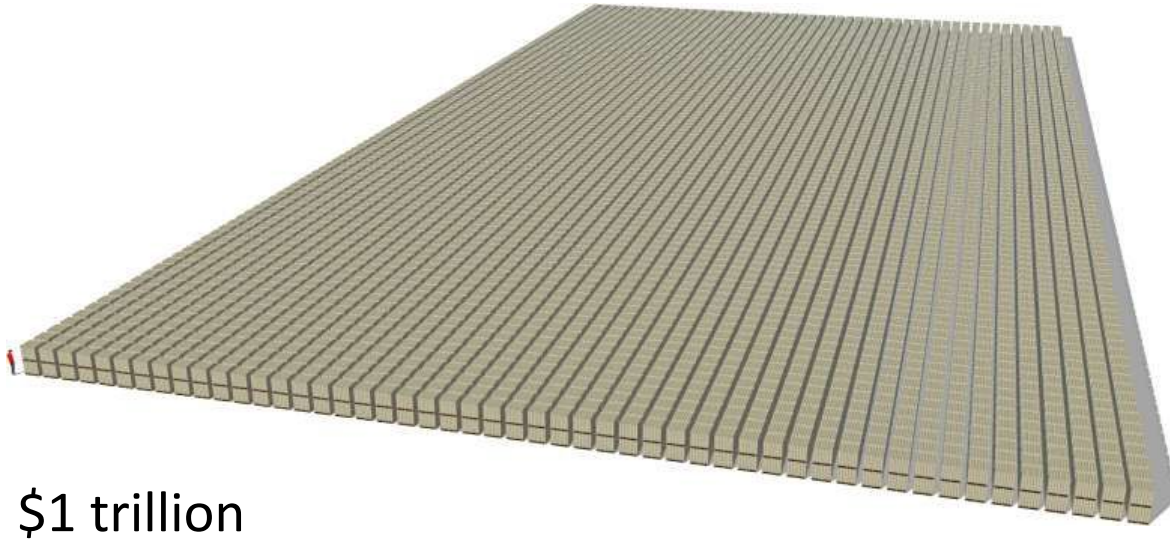
Grant writing ~5 hours per week

Abstract

Many previous time allocation studies treat work as a single activity and examine trade-offs between work and other activities. This paper investigates the at-work allocation of time among teaching, research, grant writing and service by science and engineering faculty at top US research universities. We focus on the relationship between tenure (and promotion) and time allocation, and we find that tenure and promotion do affect the allocation of time. The specific trade-offs are related to particular career paths. For example, full professors spend increasing time on service at the expense of teaching and research while longer-term associate professors who have not been promoted to full professor spend significantly more time teaching at the expense of research time. Finally, our results suggest that women, on average, allocate more hours to university service and less time to research than do men.

Image CC BY-SA 3.0, Wikimedia commons

Link, A. N., Swann, C. A., & Bozeman, B. (2008). A time allocation study of university faculty. *Economics of education review*, 27(4), 363-374.

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

"We surveyed a representative sample of Australian researchers and found that preparing new proposals for the National Health and Medical Research Council's project grants took an average of 38 working days; resubmitted ones took 28 days on average. **Extrapolating this to all 3,727 submitted proposals gives an estimated 550 working years of researchers' time** (95% confidence interval, 513–589), equivalent to a combined annual salary cost of Aus$66 million (US$68 million)"

Image CC BY-SA 3.0, Wikimedia commons

Herbert, D. L., Barnett, A. G., & Graves, N. (2013). Australia's grant system wastes time. *Nature*, *495*(7441), 314-314.

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

Reviewing is burdensome

NSF: in 2015, 16,255 reviewers evaluated 51,588 proposals - 360 person-years

ERC: 2017, 2375 reviewers evaluated 8,000 proposals

DFG: 2018, requested 22,500 reviews from 14,900 reviewers

Bendiscioli, S. (2019). The troubles with peer review for allocating research funding: Funders need to experiment with versions of peer review and decision‐making. *EMBO reports*, *20*(12), e49472.

RoЯı RESEARCH ON RESEARCH INSTITUTE

How much of this money and effort is wasted?

We don't know

What would work better?
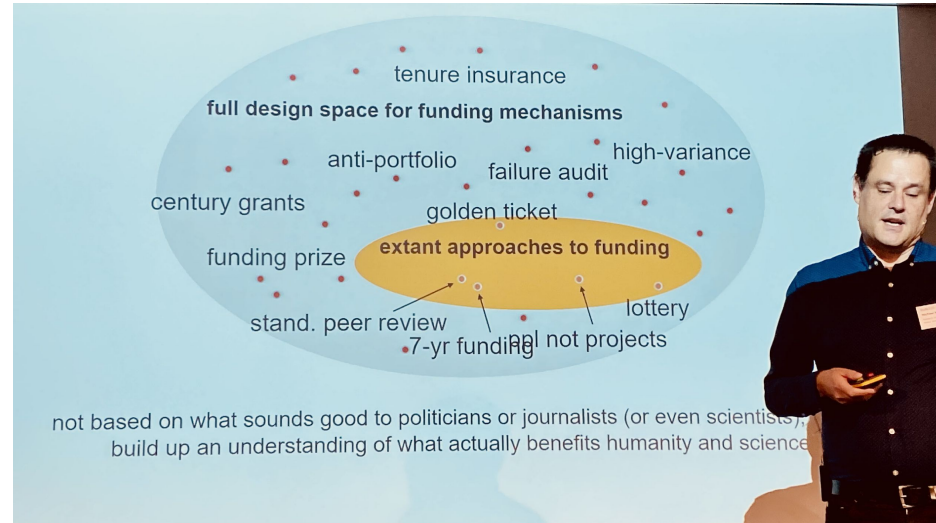
We don't know

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/

Evidence quality

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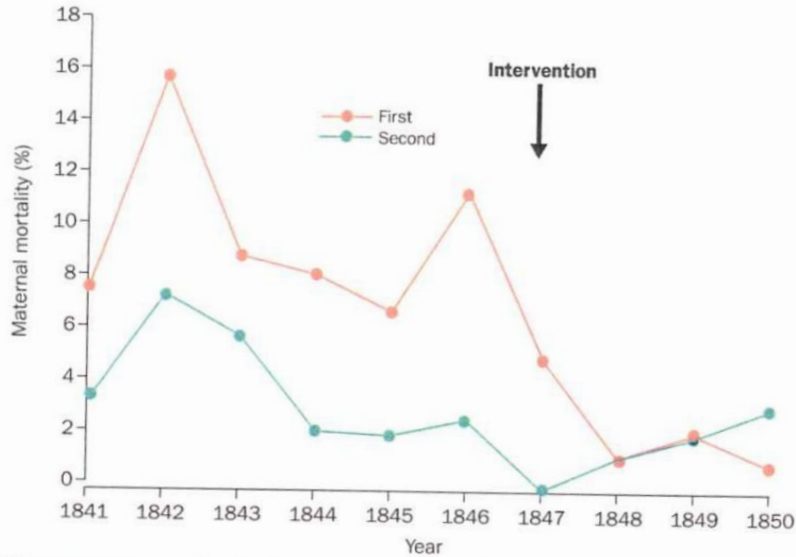
Evidence > Experience

Figure 2. Maternal mortality rates in the First and Second Clinic at the Lying-In Women's Hospital, Vienna, before and after hand hygiene in chlorinated lime had been introduced in May, 1847. Rates have been calculated according to numbers given in reference 22.

Pittet, D., & Boyce, J. M. (2001). Hand hygiene and patient care: pursuing the Semmelweis legacy. *The Lancet Infectious Diseases*, *1*, 9-20.

RoRI RESEARCH ON RESEARCH INSTITUTE

Experiments > Observation

out of 52 claims about nutrition based on observational studies, none replicated in randomised trials (and 5 trials showed effects in the opposite direction)
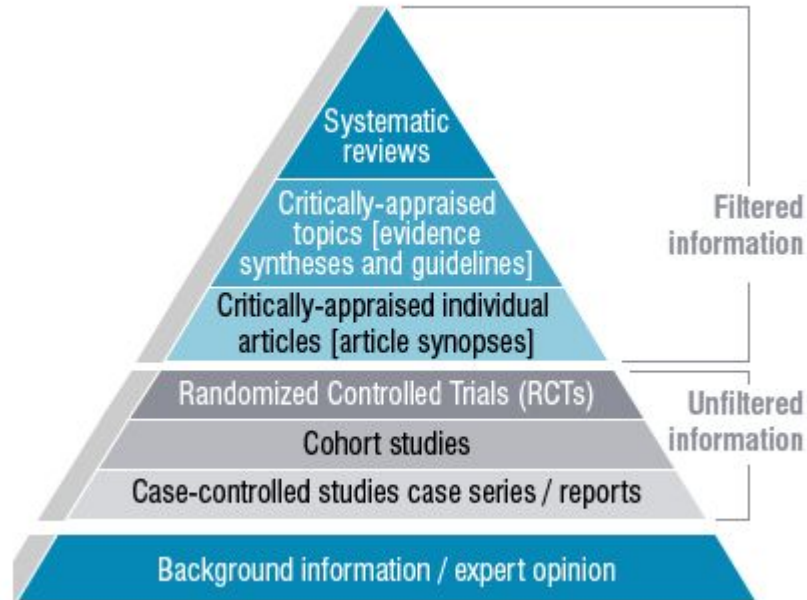
Table 1. We have found 12 papers in which claims coming from observational studies were tested in randomised clinical trials. Many of the trials are quite large. In most of the observational studies multiple claims were tested, often in factorial designs, e.g. vitamin D and calcium individually and together along with a placebo group. Note that none of the claims replicated in the direction claimed in the observational studies and that there was statistical significance in the opposite direction five times

ID no.	Pos.	Neg.	No. of claims	Treatment(s)	Reference
1	0	1	3	Vit E, beta-carotene	*NEJM* 1994; **330**: 1029–1035
2	0	3	4	Hormone Replacement Ther.	*JAMA* 2003; **289**: 2651–2662, 2663–2672, 2673–2684
3	0	1	2	Vit E, beta-carotene	*JNCI* 2005; **97**: 481–488
4	0	0	3	Vit E	*JAMA* 2005; **293**: 1338–1347
5	0	0	3	Low Fat	*JAMA*. 2006; **295**: 655–666
6	0	0	3	Vit D, Calcium	*NEJM* 2006; **354**: 669–683
7	0	0	2	Folic acid, Vit B6, B12	*NEJM* 2006; **354**: 2764–2772
8	0	0	2	Low Fat	*JAMA* 2007; **298**: 289–298
9	0	0	12	Vit C, Vit E, beta-carotene	*Arch Intern Med* 2007; **167**: 1610–1618
10	0	0	12	Vit C, Vit E	*JAMA* 2008; **300**: 2123–2133
11	0	0	3	Vit E, Selenium	*JAMA* 2009; **301**: 39–51
12	0	0	3	HRT + Vitamins	*JAMA* 2002; **288**: 2431–2440
Totals	0	5	52		

Deming, data and observational studies (2011)
https://rss.onlinelibrary.wiley.com/doi/epdf/10.1111/j.1740-9713.2011.00506.x

RoRi RESEARCH ON RESEARCH INSTITUTE

The biomedical evidence hierarchy

The biomedical evidence heirarchy

Reproducibility crisis
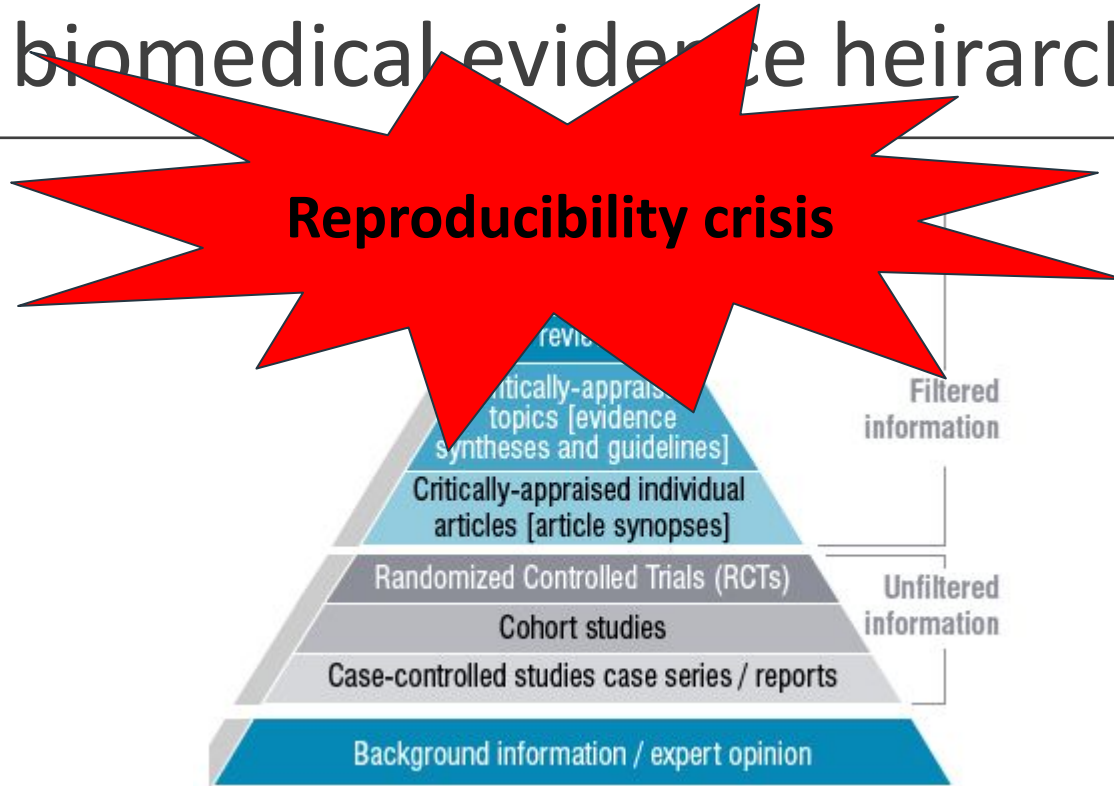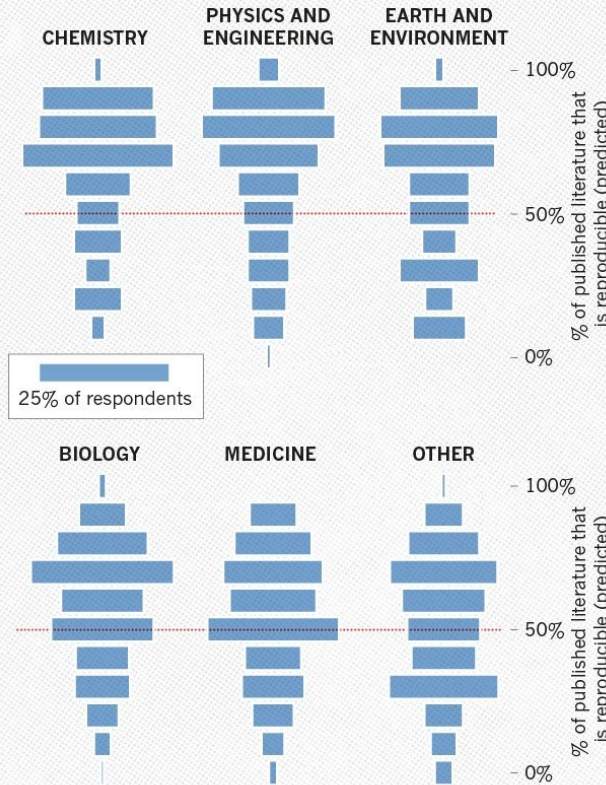
Baker, Monya. "1,500 scientists lift the lid on reproducibility." (2016). https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970

HOW MUCH PUBLISHED WORK IN YOUR FIELD IS REPRODUCIBLE?
Physicists and chemists were most confident in the literature.

CHEMISTRY   PHYSICS AND ENGINEERING   EARTH AND ENVIRONMENT

% of published literature that is reproducible (predicted)

100%
50%
0%

25% of respondents

BIOLOGY   MEDICINE   OTHER

% of published literature that is reproducible (predicted)
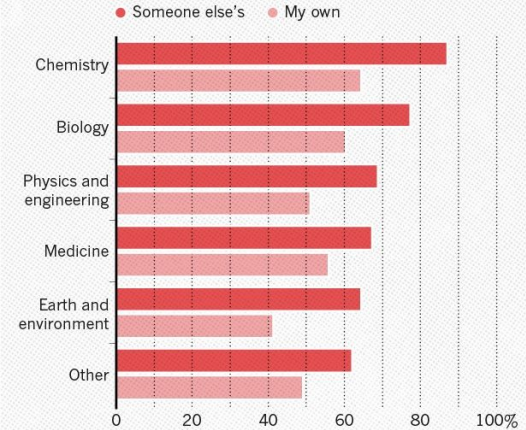
100%
50%
0%

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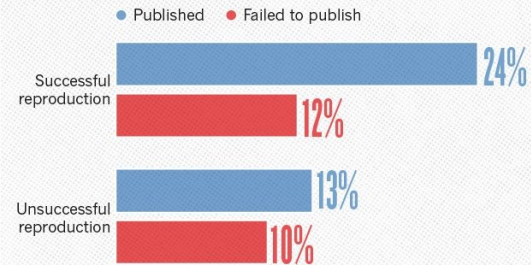
©nature

HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?
Most scientists have experienced failure to reproduce results.

● Someone else's   ● My own

Chemistry
Biology
Physics and engineering
Medicine
Earth and environment
Other

0   20   40   60   80   100%

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.

● Published   ● Failed to publish

Successful reproduction   24%   12%

Unsuccessful reproduction   13%   10%

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

©nature

RoᴙI RESEARCH ON RESEARCH INSTITUTE

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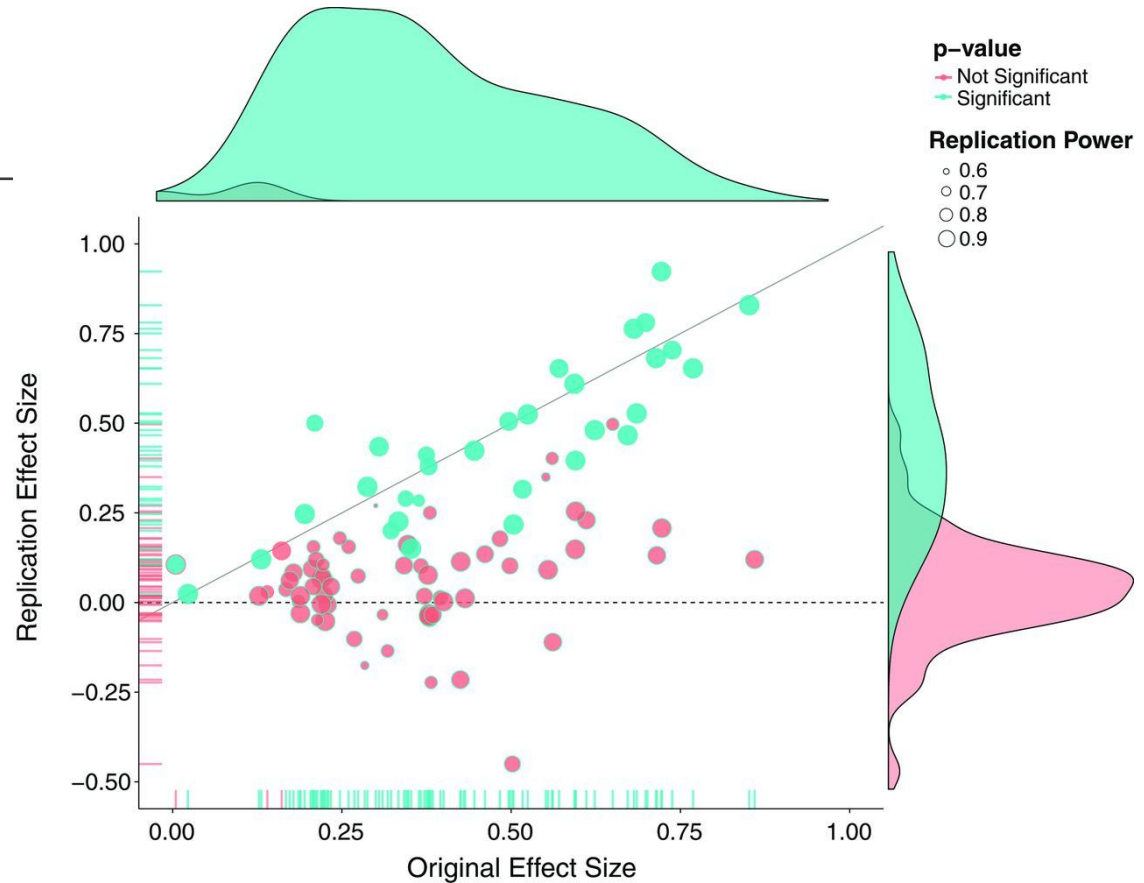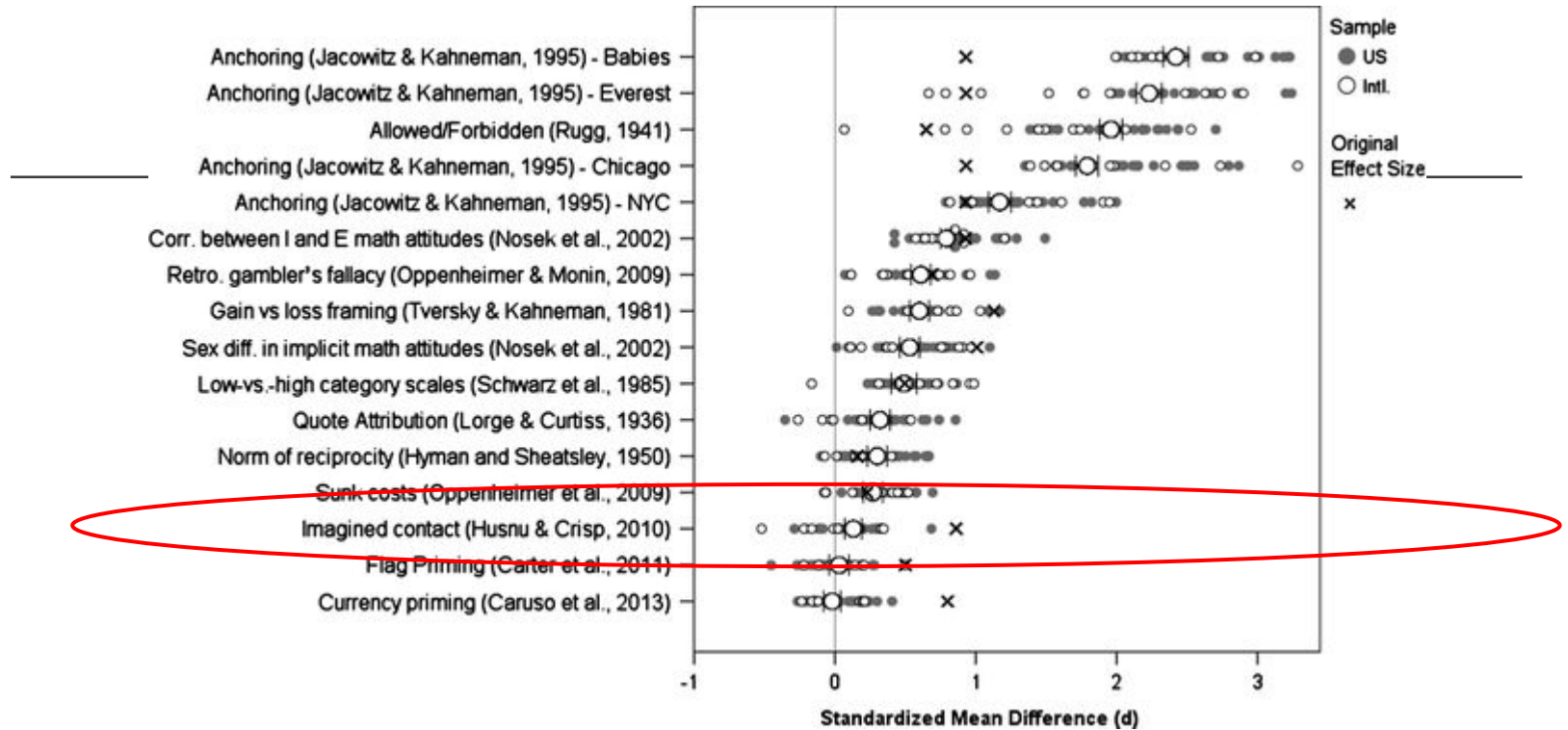
Klein Richard, A., Ratliff, K. A., Vianello, M., Adams Jr, R. B., Bahník, Š., Bernstein, M. J., ... & Cemalcilar, Z. (2014). Investigating Variation in Replicability: A "Many Labs" Replication Project. *Social Psychology*, *45*(3), 142-152.

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 c... er res... In Experiment 1, people who forced themselves to eat radishes instead of t... subsequently quit faster on unsolvable puzzles than people who had not had to over eating. In Experiment 2, making a meaningful personal choi... behavior caused a similar decrement in persistence. In Experiment 3, su... subsequent drop in performance of solvable anagrams. In Experiment 4, a... high self-regulation made people more passive (i.e., more prone to favor the passi... These results suggest that the self's capacity for active volition is lim... 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

Rediscovering the Greatest Human Strength

WILLPOWER

ROY F. BAUMEISTER & JOHN TIERNEY

"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*

RoRi RESEARCH ON RESEARCH INSTITUTE
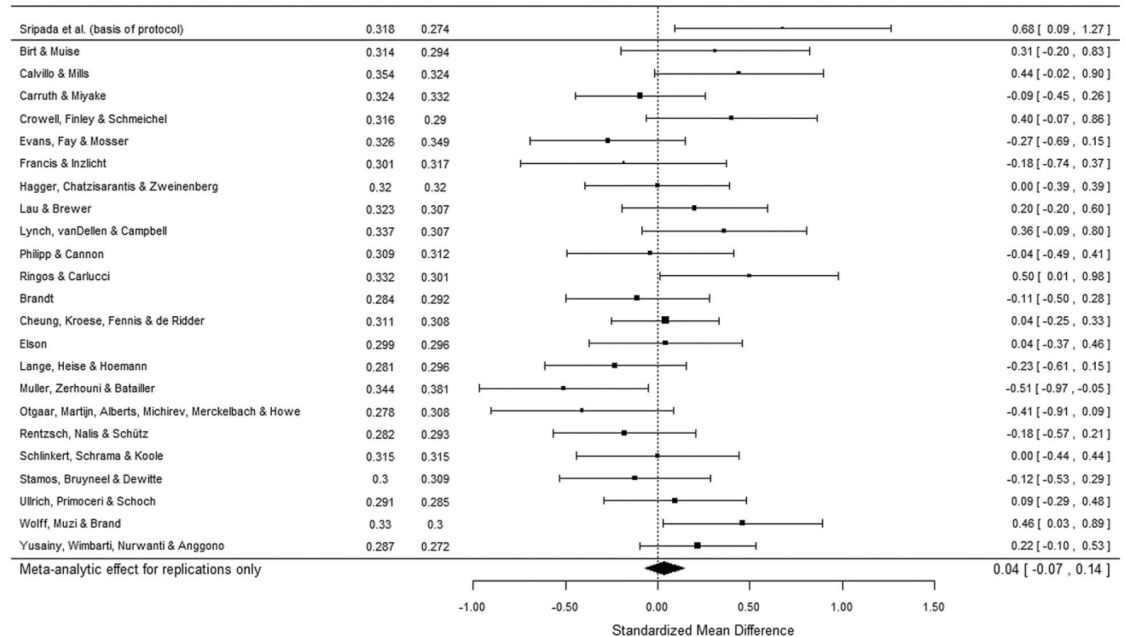
Google scholar hits "ego depletion"

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

RoЯI RESEARCH ON RESEARCH INSTITUTE

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.

Study			SMD [95% CI]
Sripada et al. (basis of protocol)	0.318	0.274	0.68 [0.09 , 1.27]
Birt & Muise	0.314	0.294	0.31 [−0.20 , 0.83]
Calvillo & Mills	0.354	0.324	0.44 [−0.02 , 0.90]
Carruth & Miyake	0.324	0.332	−0.09 [−0.45 , 0.26]
Crowell, Finley & Schmeichel	0.316	0.29	0.40 [−0.07 , 0.86]
Evans, Fay & Mosser	0.326	0.349	−0.27 [−0.69 , 0.15]
Francis & Inzlicht	0.301	0.317	−0.18 [−0.74 , 0.37]
Hagger, Chatzisarantis & Zweinenberg	0.32	0.32	0.00 [−0.39 , 0.39]
Lau & Brewer	0.323	0.307	0.20 [−0.20 , 0.60]
Lynch, vanDellen & Campbell	0.337	0.307	0.36 [−0.09 , 0.80]
Philipp & Cannon	0.309	0.312	−0.04 [−0.49 , 0.41]
Ringos & Carlucci	0.332	0.301	0.50 [0.01 , 0.98]
Brandt	0.284	0.292	−0.11 [−0.50 , 0.28]
Cheung, Kroese, Fennis & de Ridder	0.311	0.308	0.04 [−0.25 , 0.33]
Elson	0.299	0.296	0.04 [−0.37 , 0.46]
Lange, Heise & Hoemann	0.281	0.296	−0.23 [−0.61 , 0.15]
Muller, Zerhouni & Batailler	0.344	0.381	−0.51 [−0.97 , −0.05]
Otgaar, Martijn, Alberts, Michirev, Merckelbach & Howe	0.278	0.308	−0.41 [−0.91 , 0.09]
Rentzsch, Nalis & Schütz	0.282	0.293	−0.18 [−0.57 , 0.21]
Schlinkert, Schrama & Koole	0.315	0.315	0.00 [−0.44 , 0.44]
Stamos, Bruyneel & Dewitte	0.3	0.309	−0.12 [−0.53 , 0.29]
Ullrich, Primoceri & Schoch	0.291	0.285	0.09 [−0.29 , 0.48]
Wolff, Muzi & Brand	0.33	0.3	0.46 [0.03 , 0.89]
Yusainy, Wimbarti, Nurwanti & Anggono	0.287	0.272	0.22 [−0.10 , 0.53]
Meta-analytic effect for replications only			0.04 [−0.07 , 0.14]

Standardized Mean Difference

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

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

RoЯI RESEARCH ON RESEARCH INSTITUTE

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
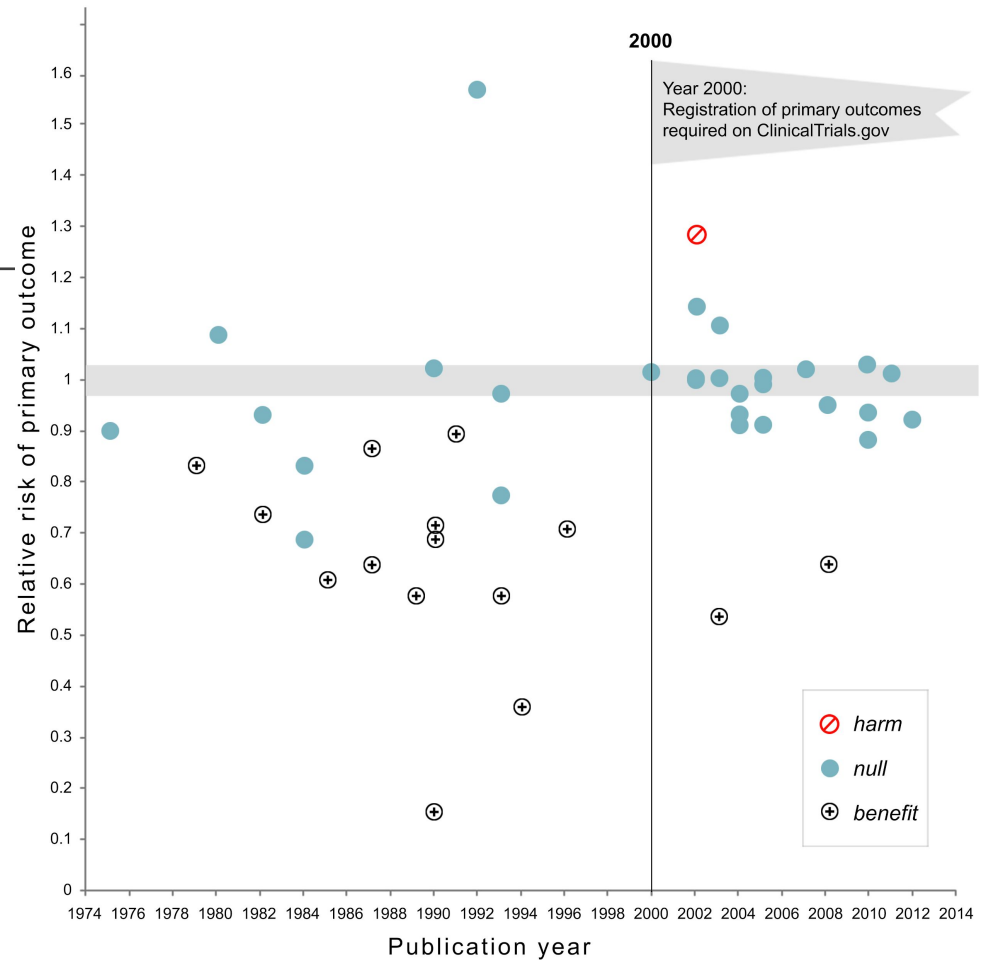
Causal Inference

Randomised Experiment

Natural Experiment

Regression Discontinuity

Difference in Difference

Instrumental Variable

How this applies to RoR

Most trials are observational

Research systems are conservative - lack of investment in evidence

Timespans are long

The "Dark matter" of Research on research

Research Funding as a decision problem

Psychophysics: the dawn of psychology
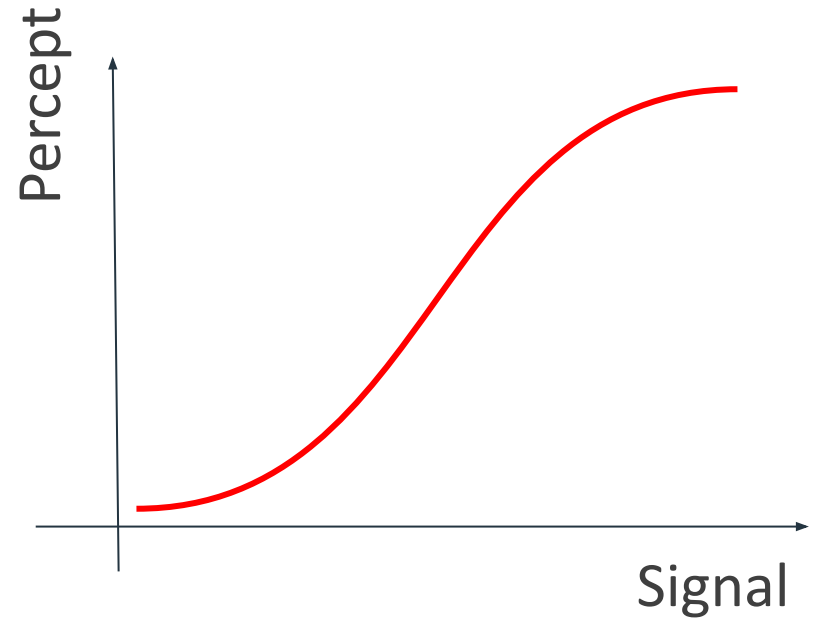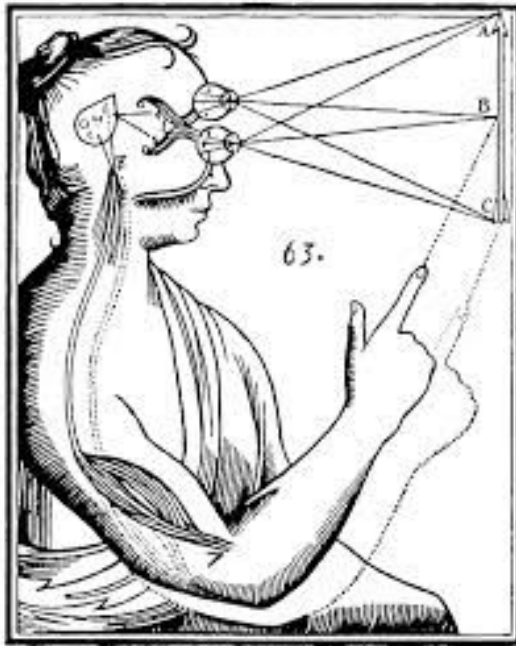
Signal Detection Theory

	Signal Present	Signal Absent
Observer Responded	Hit	False Alarm
Observer Did Not Responded	Miss	Correct Rejection

Judgement and Decision Making

Costs and Benefits

Expected Utility

Information Aggregation

Deliberation

Funding decisions are decisions

How sensitive are agencies and reviewers?

How much do they agree?

What signals are they detecting

Which direction does their bias go?

(which error types do they make?)

How does deliberation shifts decisions?

What is the value of diversity of opinions?

Reviewers may not agree on a common signal

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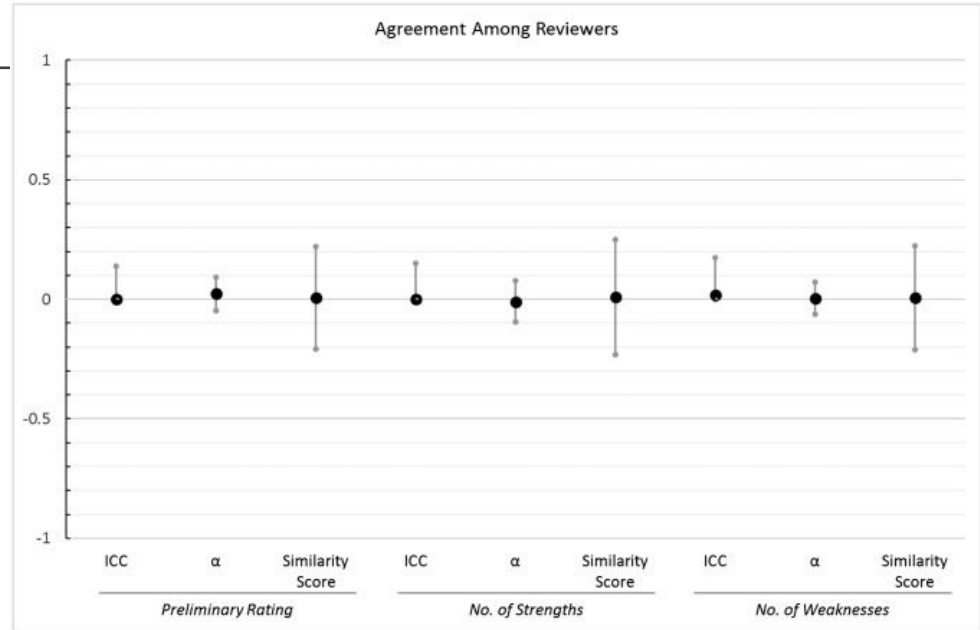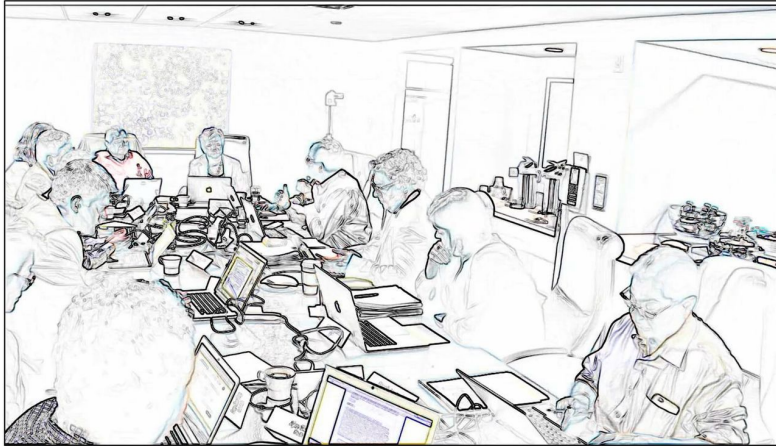
Pier, E. L., Brauer, M., Filut, A., Kaatz, A., Raclaw, J., Nathan, M. J., ... & Carnes, M. (2018). Low agreement among reviewers evaluating the same NIH grant applications. *Proceedings of the National Academy of Sciences*, *115*(12), 2952-2957.

Agency ranking are not predictive

130,000 NIH grants, funded 1980-2008

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 funding from around 55% to around 25% (even when it has otherwise been rated highly).

Q KEYWORDS: Peer-review   consistency   grant funding

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

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

RoRI RESEARCH ON RESEARCH INSTITUTE

Review scoring punishes weak aspects
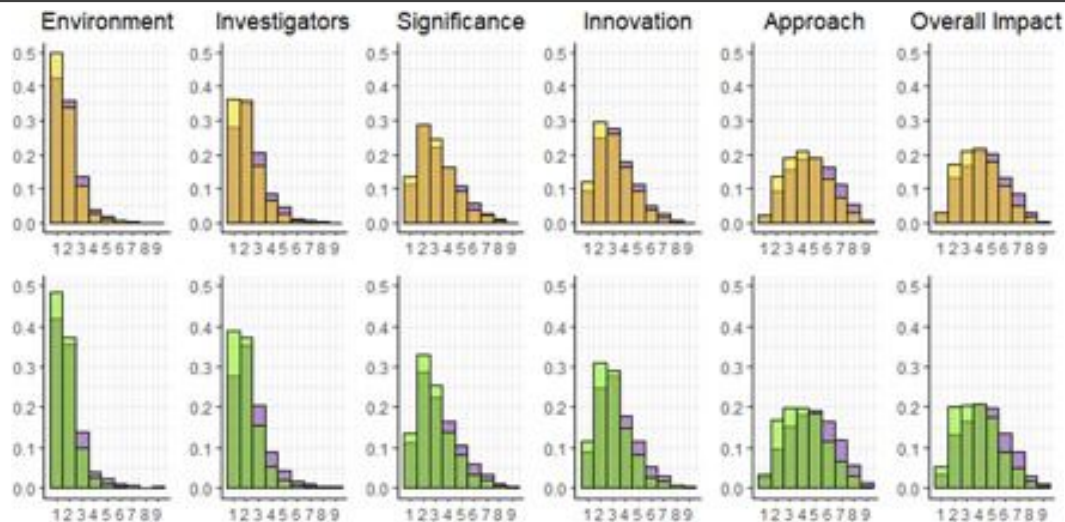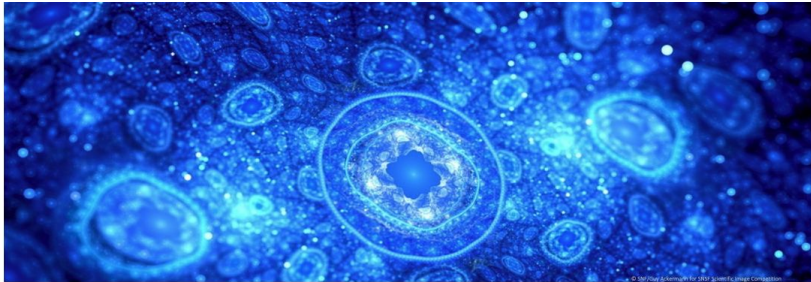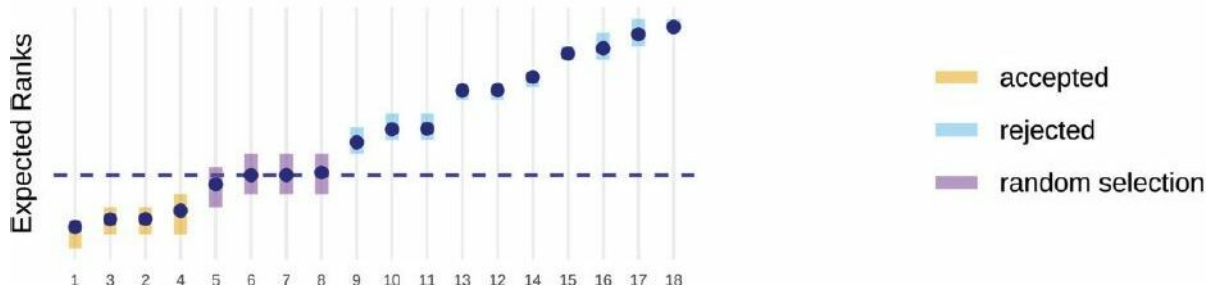
Erosheva et al (2020) 140k NIH reviews

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

Expected Ranks

1 3 2 4 5 6 7 8 9 10 11 13 12 14 15 16 17 18

accepted
rejected
random selection

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.

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

Marco Bieri and Rachel Heyard

Swiss National Science Foundation

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

RoЯI RESEARCH ON RESEARCH INSTITUTE

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

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 detectibly 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

RoЯI RESEARCH ON RESEARCH INSTITUTE

Does name matter?

Percentage of First Time PIs by Observing Cycle

start of dual anonymous peer review
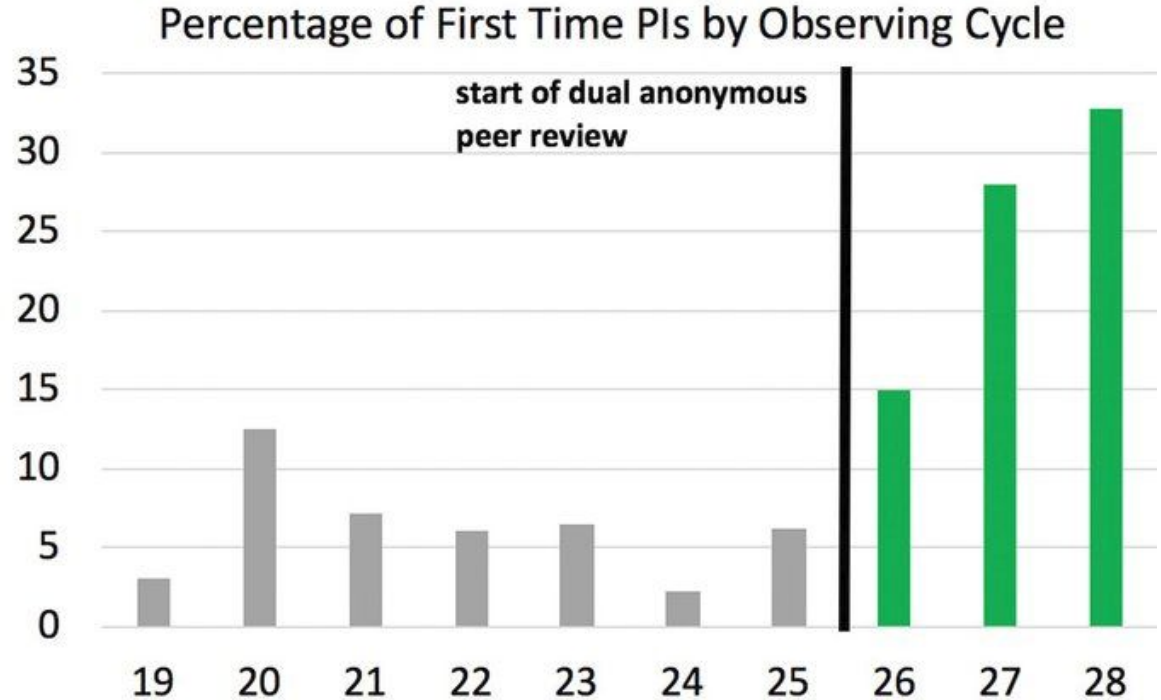
Image: Rachel Osten Space Telescope Science Institute. See also

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

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

Biases in collective decisions

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

RoЯi RESEARCH ON RESEARCH INSTITUTE

Costs and benefits
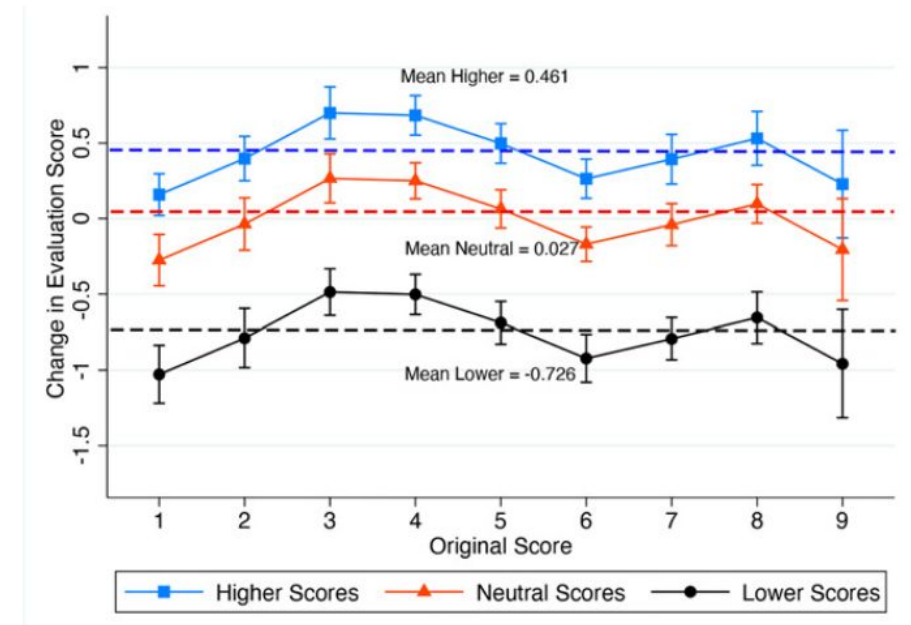
These slides:
http://bit.ly/toms-talks
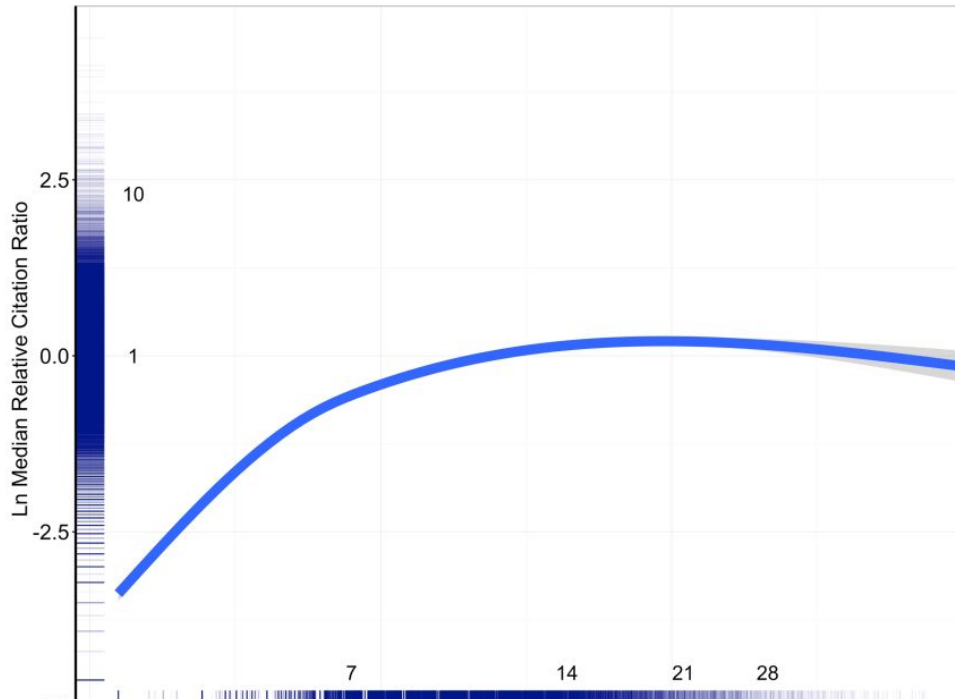
"In a cohort of over 71,000 unique scientists funded by NIH between 1996 and 2014 we analyzed the association of grant support (as measured by annual GSI) with 3 bibliometric outcomes, maximum Relative Citation Ratio (which arguably reflects a scientist's most influential work), median Relative Citation 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.

better research designs -> better inferences
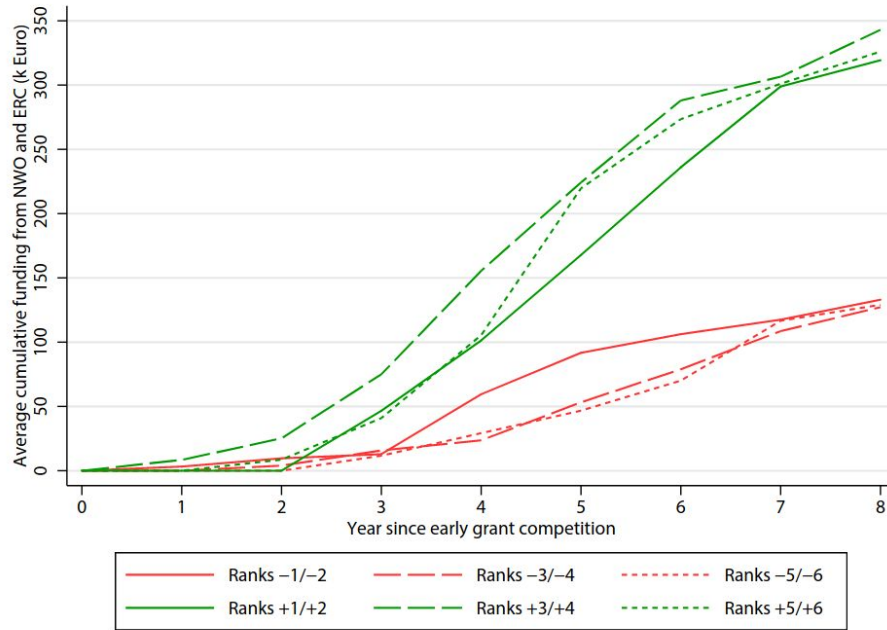
Fig. 3. Accumulation of grant money by early career grant applicants. Shown is the cumulative amount of funding received in NWO and ERC competitions (vertical axis) as a function of the number of years elapsed since the early career grant competition (horizontal axis). This relationship is shown for different ranks above (green, +) and below (red, −) the early career funding threshold.

Bol, T., de Vaan, M., & van de Rijt, A. (2018). The Matthew effect in science funding. *Proceedings of the National Academy of Sciences*, *115*(19), 4887-4890.

These slides:
http://bit.ly/toms-talks
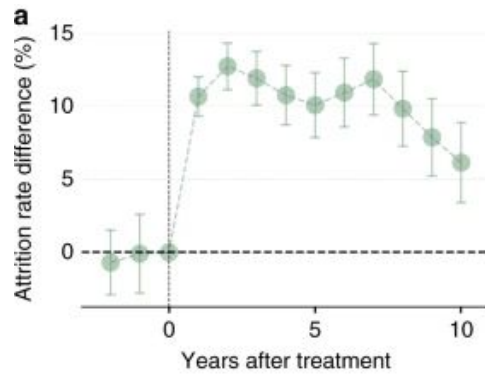
The value of research funding for knowledge creation and dissemination: A study of SNSF Research Grants
https://www.nature.com/articles/s41599-021

RoЯI RESEARCH ON RESEARCH INSTITUTE
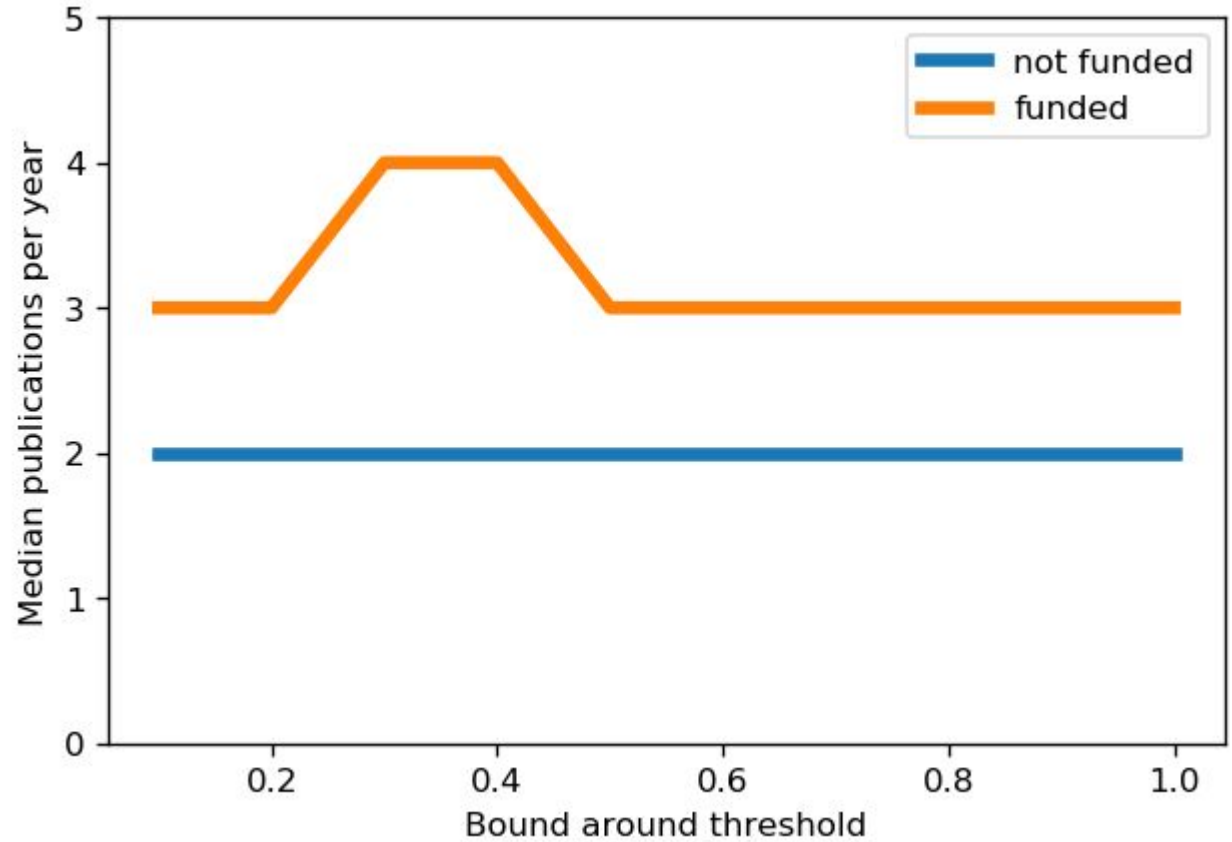
but context still matters

Wang, Y., Jones, B. F., & Wang, D. (2019). Early-career setback and future career impact. Nature communications, 10(1), 1-10.
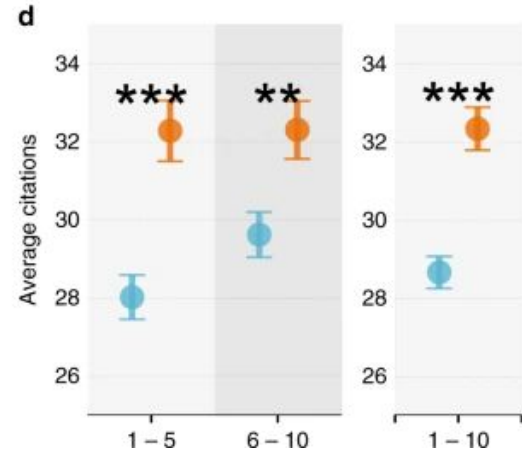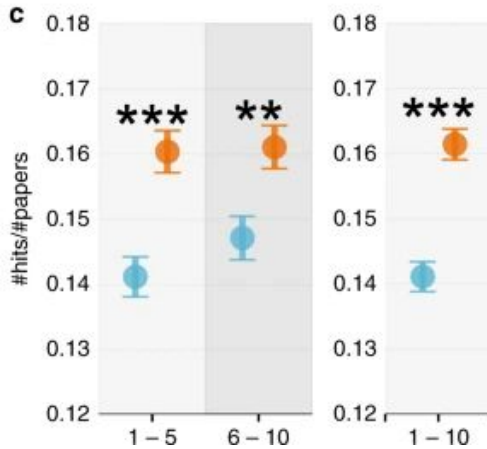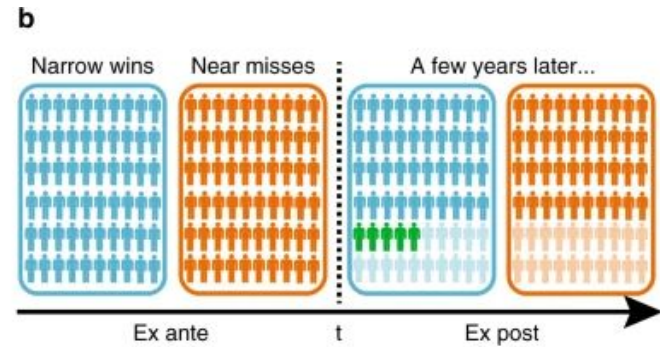
and results can still contradict!

Using funding randomisation to estimate effects of winning funding

Table 4. Estimated delayed mean percentage differences between funded and not funded researchers from the HRC of New Zealand trial.

Outcome	Estimate	Mean	95% CI	P
			Bayesian	
Publications	Rate ratio	1.01	0.91 to 1.12	.82
Citations per paper	Absolute	−0.5	−1.1 to 0.1	.09
Altmetric	Absolute	2.4	−3.8 to 8.5	.46
			Non-Bayesian	
Citations	Rate ratio	1.03	1.02 to 1.04	<.0001

The table includes 95 per cent credible interval for the difference, and estimated probability that the groups differ (P). The citation results use a non-Bayesian confidence interval and P-value. The estimates are a linear change in each year after funding.

RoRI RESEARCH ON RESEARCH INSTITUTE

RoRI's AFIRE project

A Research on Research Institute

https://researchonresearch.org/

We're transforming research systems and cultures

Ensuring that we have the evidence we need to realise the full potential of research.

Projects →   About Us →

Unlocking the potential of the US$2.5 trillion invested globally in research.

RoRI aims to unlock more of the potential of the US$2.5 trillion invested globally in research every year. By turning the tools of research back on itself, RoRI generates data and analysis to improve how we fund, practice, evaluate and communicate research.

RoRI Phase 2 partners
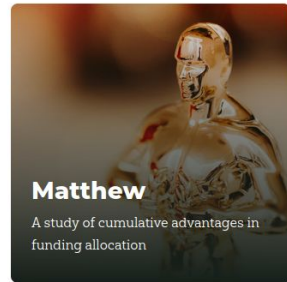
Funder experiments are co-produced

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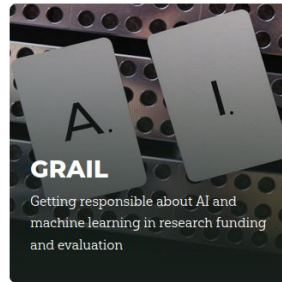http://bit.ly/tom-talks

RoЯI RESEARCH ON RESEARCH INSTITUTE

Matthew
A study of cumulative advantages in funding allocation

Narratives
The uses and evaluation of researchers' narrative CVs

GRAIL
Getting responsible about AI and machine learning in research funding and evaluation

Funder Data Platform
Supporting data-sharing to unlock insights into research funding

AGORRA
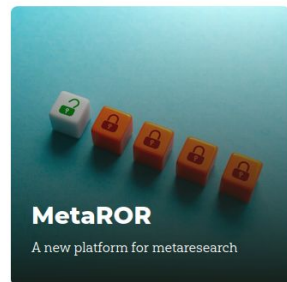A global observatory of responsible research assessment

Undisciplined
Future models of funding and evaluating transdisciplinary research

Peer Review
RoRI Atlas of Peer Review

Portfolios
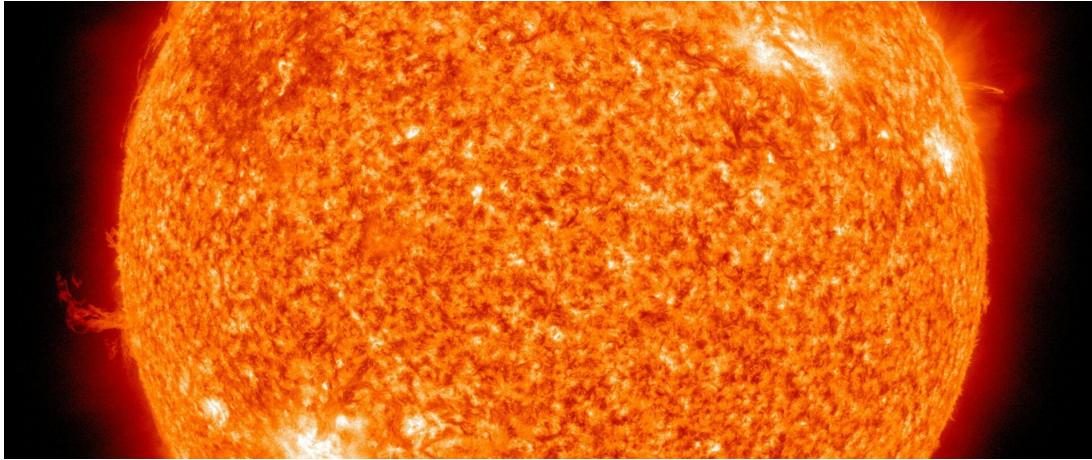Research funding landscape analysis
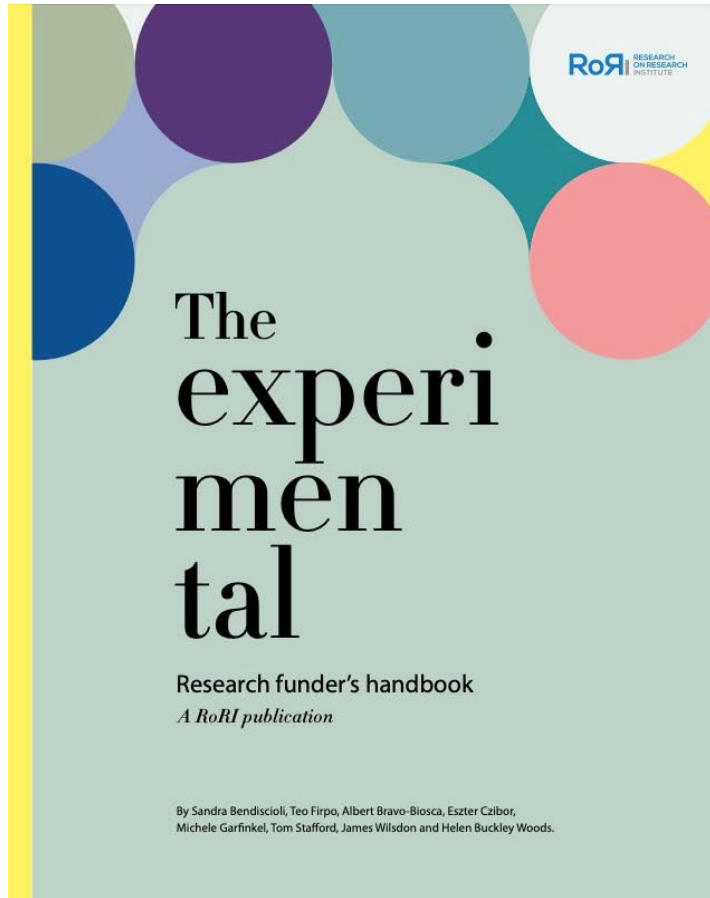
MetaROR
A new platform for metaresearch

AFIRE
Accelerator For Innovation & Research Funding Experimentation

https://researchonresearch.org/projects/

RoЯI RESEARCH ON RESEARCH INSTITUTE

AFIRE

https://researchonresearch.org/project/a-f-i-r-e/

Bendiscioli, Sandra; Firpo, Teo; Bravo-Biosca, Albert; Czibor, Eszter; Garfinkel, Michele; Stafford, Tom; et al. (2022):

The experimental research funder's handbook (Revised edition, June 2022, ISBN 978-1-7397102-0-0).

Research on Research Institute. Report. https://doi.org/10.6084/m9.figshare.19459328.v2

Table 2: Target outcome, unit of analysis and sample availability for one funding call

Target outcome	applicant diversity, beliefs about partial randomisation	proposal novelty, ambition/risk	reviewer burden, review consistency	project productivity, diversity characteristics of awardees, awardee reaction to award by partial randomisation
Unit of analysis	APPLICANTS	APPLICATIONS	REVIEWS	AWARDS
Sample available	number of investigators	number of applications	number of applications x reviews per application	number of applications x proportion funded
Illustrative numbers assuming 100 applications, 3 investigators, 4 reviews per applications, and a 10% success rate	300	100	400	10

Check for updates

REVIEW

Where next for partial randomisation of research funding? The feasibility of RCTs and alternatives [version 1; peer review: 2 approved, 1 approved with reservations]

Tom Stafford [1], Ines Rombach[1], Dan Hind[1], Bilal Mateen [2], Helen Buckley Woods[3], Munya Dimario[1], James Wilsdon [4]

[1]The University of Sheffield, Sheffield, England, UK
[2]Wellcome Trust, London, England, UK
[3]Research on Research Institute, London, England, UK
[4]University College London, London, England, UK

Stafford T, Rombach I, Hind D et al. (2023) Where next for partial randomisation of research funding? The feasibility of RCTs and alternatives Wellcome Open Res 2023, 8:309 https://doi.org/10.12688/wellcomeopenres.19565.1

RoRI RESEARCH ON RESEARCH INSTITUTE

Future funder experiments

Sequential evaluation for review debiasing

Navigating the grey zone: capturing reviewer uncertainty

Matthew: studying cumulative advantages in funding evaluation

A large multi-funder trial of partial randomisation

Experiments with the use of narrative CVs

Designing panel rules for smarter decision making

Responsible uses of AI & machine learning in research evaluation

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

RoЯI RESEARCH ON RESEARCH INSTITUTE

Future funder experiments

Sequential evaluation for review debiasing

BIAS & BIAS MITIGATION

Navigating the grey zone: capturing reviewer uncertainty

JUDGEMENT AND EVALUATION

Matthew: studying cumulative advantages in funding evaluation

EXPERIMENT / TRIAL DESIGN

A large multi-funder trial of partial randomisation

GROUP DECISION MAKING

Experiments with the use of narrative CVs

TRUST & TRANSPARENCY IN DECISION PROCESSES

Designing panel rules for smarter decision making

Responsible uses of AI & machine learning in research evaluation

Volkswagen Foundation introduces experimental Distributed Peer Review

Supported by RoRI researchers, the Foundation will run an experiment in parallel to its standard selection of proposals

RoR news

RoRI RESEARCH ON RESEARCH INSTITUTE

Future AFIRE activities

Funder's Forum

Sprints - capacity building

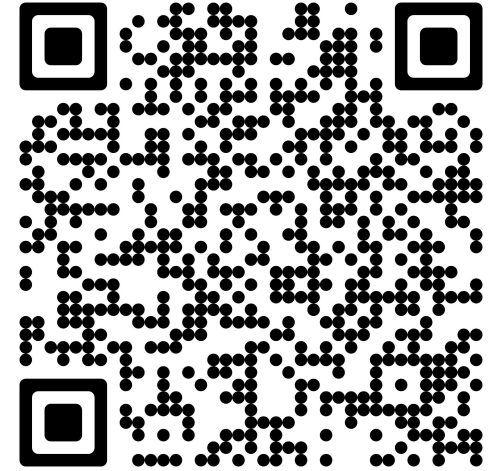Evidence synthesis and sharing

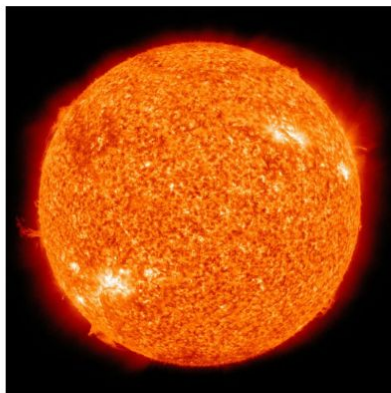Supporting more experiments

metascience.info

tomstafford.github.io

RESERVE SLIDES FOLLOW
(not for show)

Accelerating funder experiments: the launch of AFIRE

Invitation-only online seminar (via Zoom)
**14:00-17:00 BST/UK, Monday 13 May 2024**

Please register your interest here:
https://forms.gle/Wzotjz1oCxvLgiHR7

You are warmly invited to participate in this invitation-only event for research and innovation funders to mark the launch of RoRI's new programme — **AFIRE: an Accelerator For Innovation & Research Funding Experimentation.**
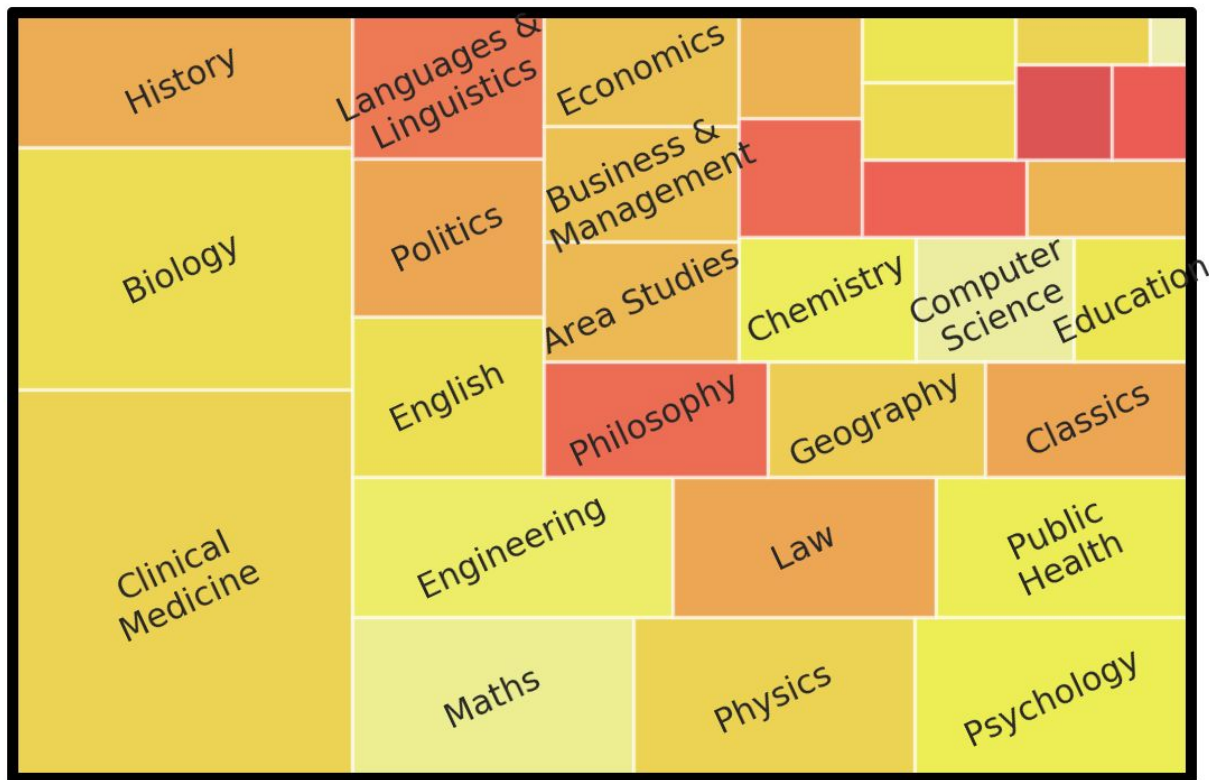
A RDD in Research Policy

Moqi Groen-Xu et al (2022) Short-term
incentives of research evaluations:
Evidence from the UK Research Excellence
Framework
https://doi.org/10.1016/j.respol.2023.104
729
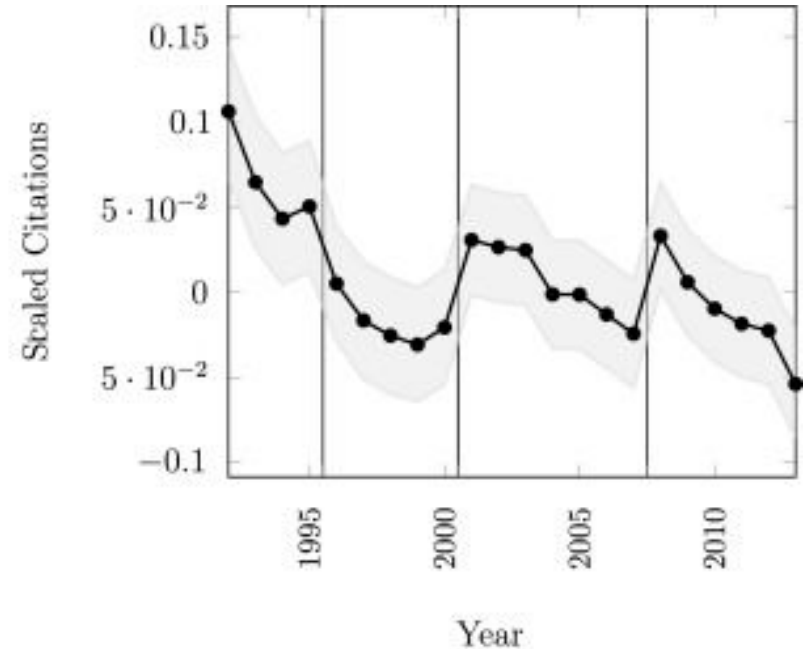
(b) Coefficients of publication year.

REF: funding evaluation and allocation

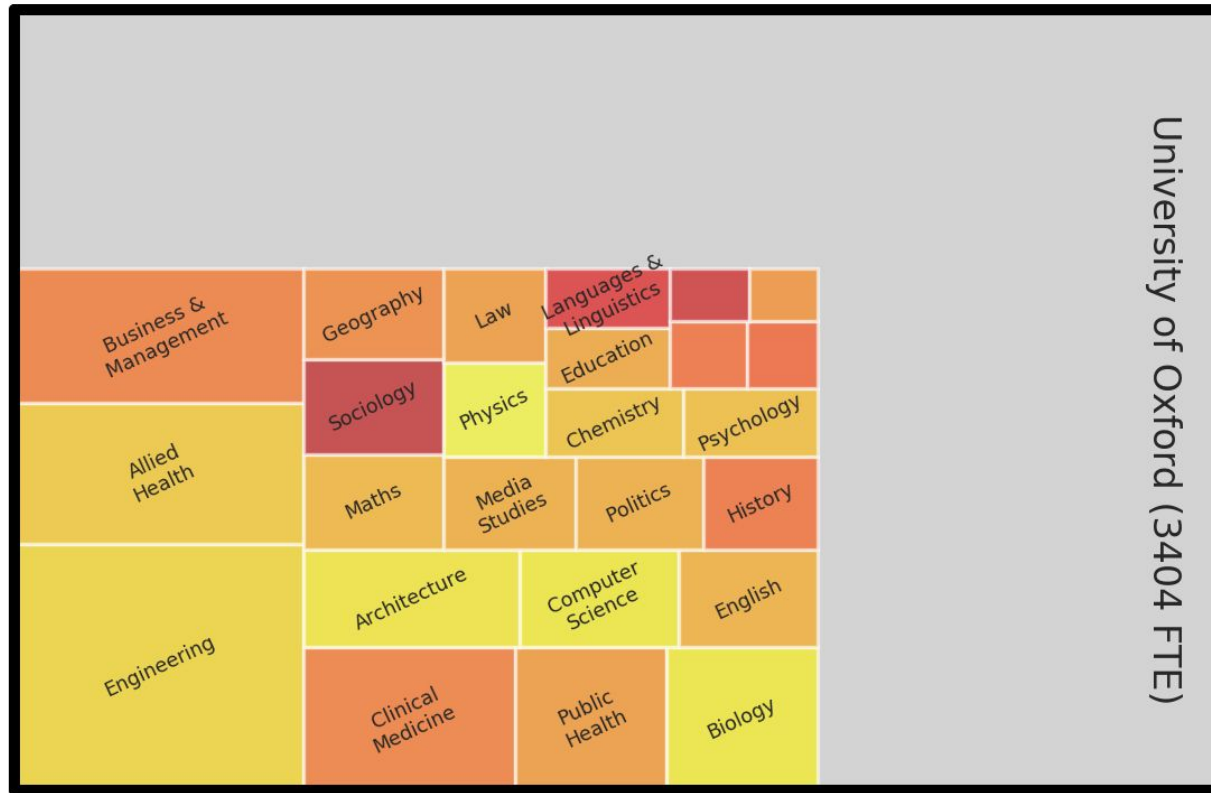These slides:
http://bit.ly/toms-talks

University of Oxford (3404 FTE)

Area scaled to FTE of that UoA, colour scaled to Overall GPA
Grey box is the FTE area of the largest institution in REF2021
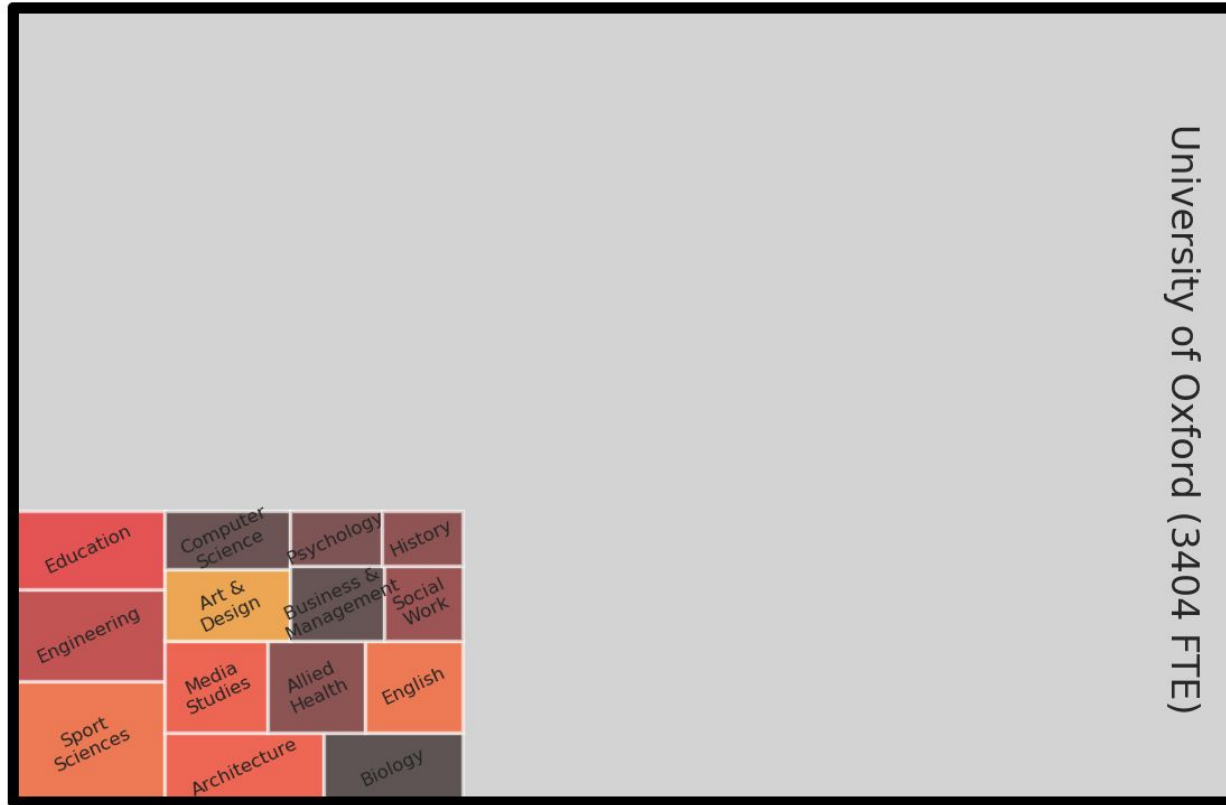
These slides:
http://bit.ly/toms-talks

RoRI RESEARCH ON RESEARCH INSTITUTE

The University of Sheffield (1518 FTE)

University of Oxford (3404 FTE)

Business & Management, Geography, Law, Languages & Linguistics, Sociology, Physics, Education, Allied Health, Chemistry, Psychology, Maths, Media Studies, Politics, History, Engineering, Architecture, Computer Science, English, Clinical Medicine, Public Health, Biology

Area scaled to FTE of that UoA, colour scaled to Overall GPA
Grey box is the FTE area of the largest institution in REF2021

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

https://github.com/tomstafford/ref2021

RoЯI RESEARCH ON RESEARCH INSTITUTE

Sheffield Hallam University (458 FTE)

University of Oxford (3404 FTE)

Education, Engineering, Sport Sciences, Computer Science, Art & Design, Media Studies, Architecture, Psychology, Business & Management, Allied Health, Biology, History, Social Work, English

Area scaled to FTE of that UoA, colour scaled to Overall GPA
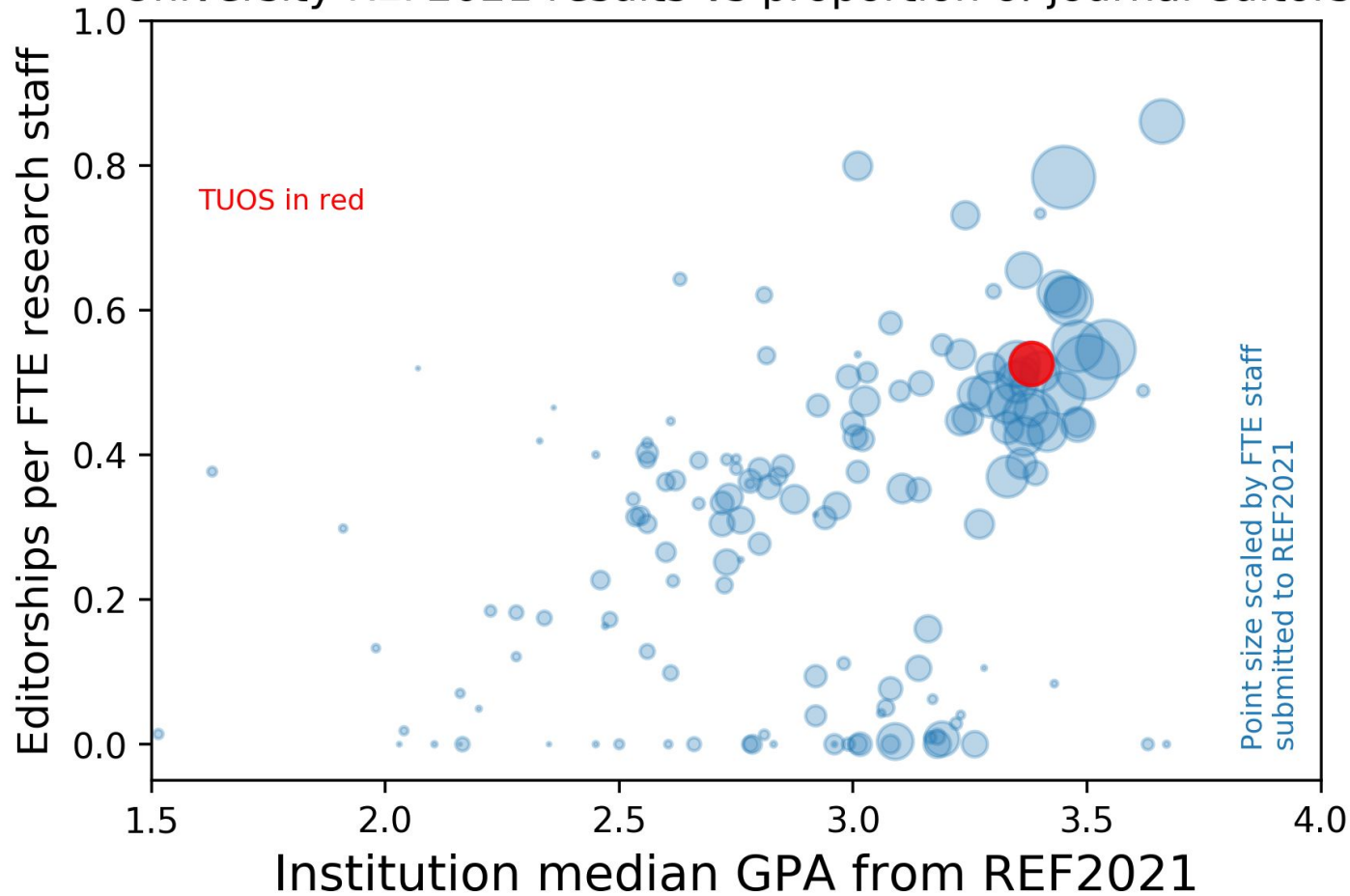Grey box is the FTE area of the largest institution in REF2021

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

https://github.com/tomstafford/ref2021

RoЯI RESEARCH ON RESEARCH INSTITUTE

University REF2021 results vs proportion of journal editors

TUOS in red

Point size scaled by FTE staff submitted to REF2021

Editorships per FTE research staff

Institution median GPA from REF2021

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

https://tomstafford.github.io/editors/

RoЯI RESEARCH ON RESEARCH INSTITUTE

Why experiment?

"*If I look back on many years of involvement in political decision-making and policy-making around science, innovation and R&D,* **I am struck by how much of it tends to turn on gut feel of the individuals involved, than on hard evidence and analysis***. This is of course ironic, since good science is all about testing hypotheses against data, empirical results and facts.*"

Sir John Kingman,

Reflections on his time as Chair of UK Research and Innovation, 14 July 2021.

For unto every one that hath shall be given, and he shall have abundance: but from him, that hath not shall be taken away even that which he hath. And cast ye the unprofitable servant into outer darkness: there shall be weeping and gnashing of teeth.
— Matthew 25:24–30

Topic

RANDOMISATION
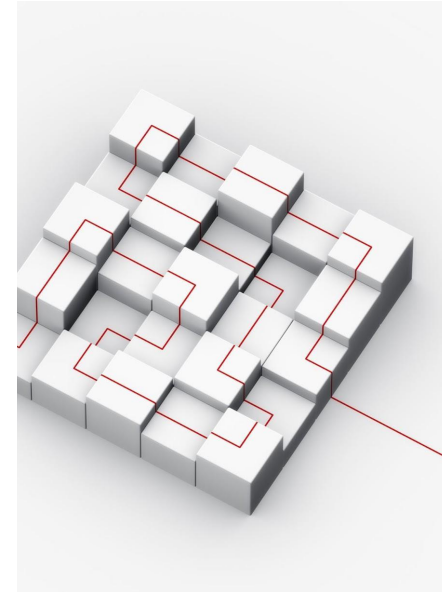
Partial randomisation

Targeted randomisation

Focal randomisation

Random selection

Lottery

Modified lottery

Partial randomisation. What is it?

- A mechanism *complementing* **peer review** for allocating research funding.

- Also called **focal** or **targeted randomisation**, or a **modified lottery**.

- Only applied to a **subset** of peer reviewed applications

- **It relies on peer reviewers' expertise** to first recommend applications for funding - those that meet the quality and criteria. Randomisation is applied to select among the recommended applications.

- **Variations:** One or more peer review rounds

- **Different tools:** Manual lottery drum, plastic capsules in a bowl, software

Questions in the mix:

Bias – whether randomisation results in different patterns of allocation (by discipline, institution, gender, career stage or other variables);

Burden – whether randomisation reduces burden and bureaucracy and burden, both on applicants and on the funding agency;

Risk — whether randomisation is particularly useful for funding processes intended to support highly innovative research in areas which might be considered "too risky" under more conventional modes of assessment;

Legitimacy — the extent to which attitudes and perceptions of focal randomisation vary within the wider research community and its stakeholders (as explored in a recent Health Research Council of New Zealand study);

Outcomes – whether random allocation ultimately results in projects with different impacts and outcomes, relative to other allocation modes. This is the most important question but also the hardest and slowest to study and measure.

Research Integrity and Peer Review

Home   About   Articles   Submission Guidelines

Research | Open Access | Published: 03 February 2020

The acceptability of using a lottery to allocate research funding: a survey of applicants

Mengyao Liu, Vernon Choy, Philip Clarke, Adrian Barnett, Tony Blakely & Lucy Pomeroy ✉

Research Integrity and Peer Review **5**, Article number: 3 (2020) | Cite this article

5930 Accesses | 1 Citations | 186 Altmetric | Metrics

Abstract

Background

The Health Research Council of New Zealand is the first major government funding agency to use a lottery to allocate research funding for their Explorer Grant scheme. This is a somewhat controversial approach because, despite the documented problems of peer review, many researchers believe that funding should be allocated solely using peer review, and peer review is used almost ubiquitously by funding agencies around the world. Given the rarity of alternative funding schemes, there is interest in hearing from the first cohort of researchers to ever experience a lottery. Additionally, the Health Research Council of New Zealand wanted to hear from applicants about the acceptability of the randomisation process and anonymity of applicants.
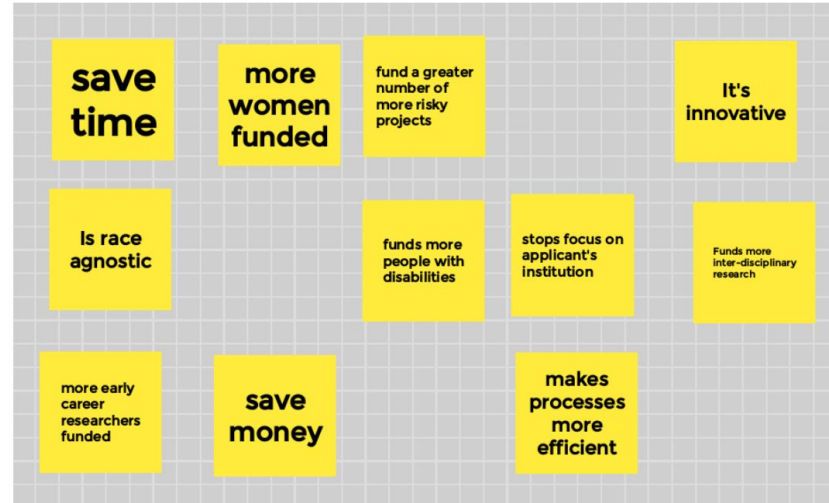
Why draw lots? Funder motivations for using partial randomisation to allocate research grants

Helen Buckley Woods and James Wilsdon

December 2021

Figure 3: Example Jamboard prepared for elicitation exercise

Summary: organisational motivations

Fairness: decision making, diversity, perceived fairness, the law

The Grey Zone: eliminating deadlock and overcoming unhelpful group dynamics

Disciplinary spread: overcoming bias to creative research, overlooked fields and 'cold' topics

Innovation: allied to values, a 'nice to have' by-product, is it really innovative?

Efficiency: money saving or more costly? Time saving: desirable, but gains may be negligible

Funder experiments with partial randomisation: conclusions (1)

- ✔ Well accepted by applicants, reviewers, scientific community and media
- ✔ Acceptance is conditional to an initial peer reviewed selection
- ✔ No negative effects
- ✔ PR extended to other schemes
- ✔ More data is needed to draw meaningful conclusions
- ✔ To be able to make comparisons, it is important to evaluate the same aspects or effects

VolkswagenStiftung

Experiment! In search of bold research ideas

Swiss National Science Foundation

Postdoc.Mobility Fellowships

FWF — Der Wissenschaftsfonds.

1000 Ideas Programme

hrc nz

Explorer Grants

Who made it possible?

Collaboration between 15 strategic partners, RoRI core team, EMBO, SNSF and Nesta's Innovation & Growth Lab

Steering Group: Gert Balling, Marco Bieri, Amanda Blatch-Jones, Michele Garfinkel, Jon Holm, Vincent Traag; Helen Buckley Woods, James Wilsdon

Reporting (motivations, handbook, earlier scoping paper): Sandra Bendiscioli, Albert Bravo-Biosca, Ester Czibor, Teo Firpo, Michele Garfinkel, Tom Stafford, James Wilsdon, Helen Buckley Woods

Australian Research Council
Alfred P. Sloan Foundation
Austrian Science Fund
Chan Zuckerberg Initiative
European Molecular Biology Organization
Michael Smith Health Research BC
National Institute for Health Research
Innovation Growth Lab at Nesta (non-RoRI partner)
Netherlands Organisation for Scientific Research
Novo Nordisk Fonden
Research Council Norway
Swiss National Science Foundation
UK Research and Innovation
Volkswagen Foundation
Wellcome Trust

RoRI RESEARCH ON RESEARCH INSTITUTE

Image: Figure 1, Dror, I. E., & Kukucka, J. (2021). Linear Sequential Unmasking–Expanded (LSU-E): A general approach for improving decision making as well as minimizing noise and bias. Forensic Science International: Synergy, 3.

Dr Charlotte Brand
lottybrand.wordpress.com

https://ces-transformationfund.org/

Cultural Evolution Society Transformation Fund

A kinder research culture is possible

Wellcome is right to call out hyper-competitiveness in research and question the focus on excellence. But other funders must follow its move.

"Wellcome and its partners in RoRI should be commended for taking an important first step. They have recognized that there are problems in research culture and that these need to be fixed. RoRI will help to probe some of the causes of distress, and suggest solutions. Now, other funders and research-management societies must join the mission…"

Nature editorial, 1 October 2019

Research on research

Metaresearch

Metascience

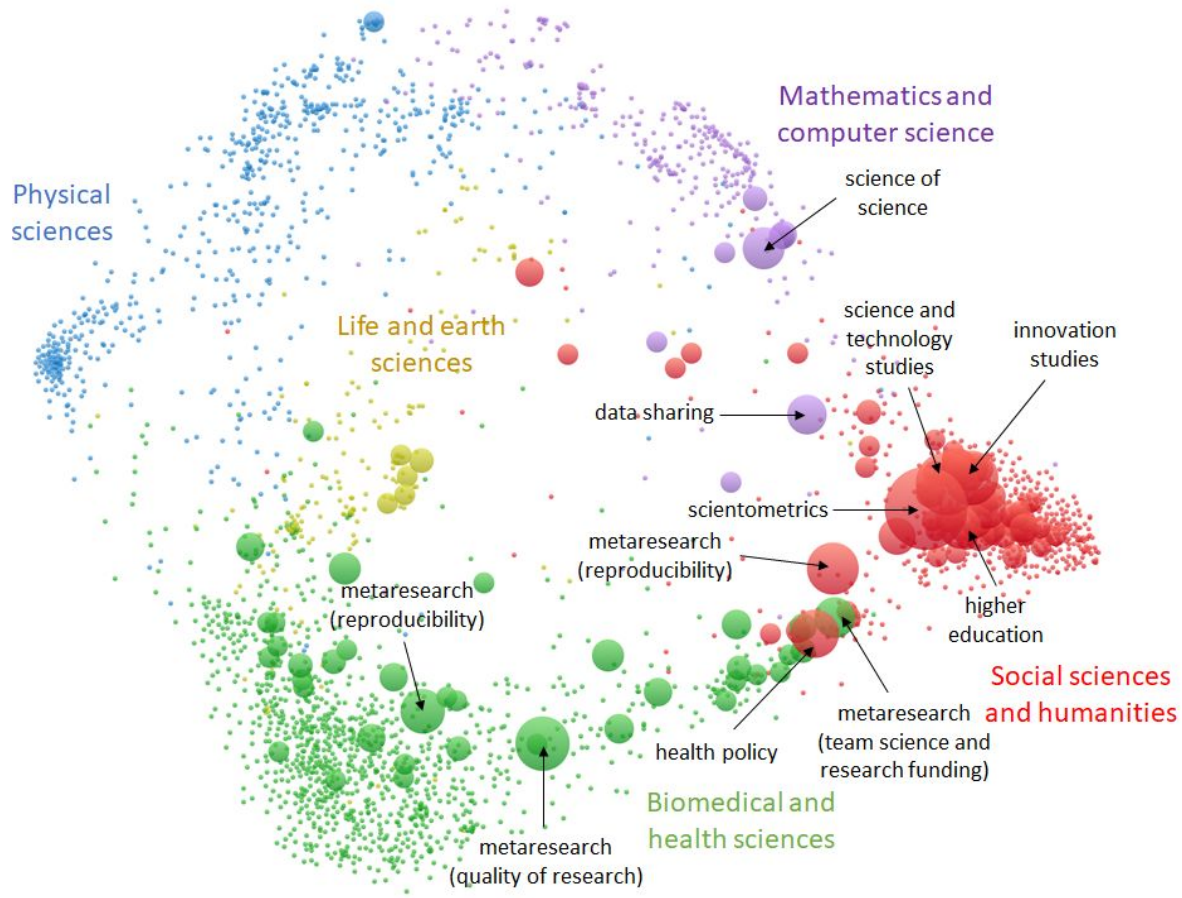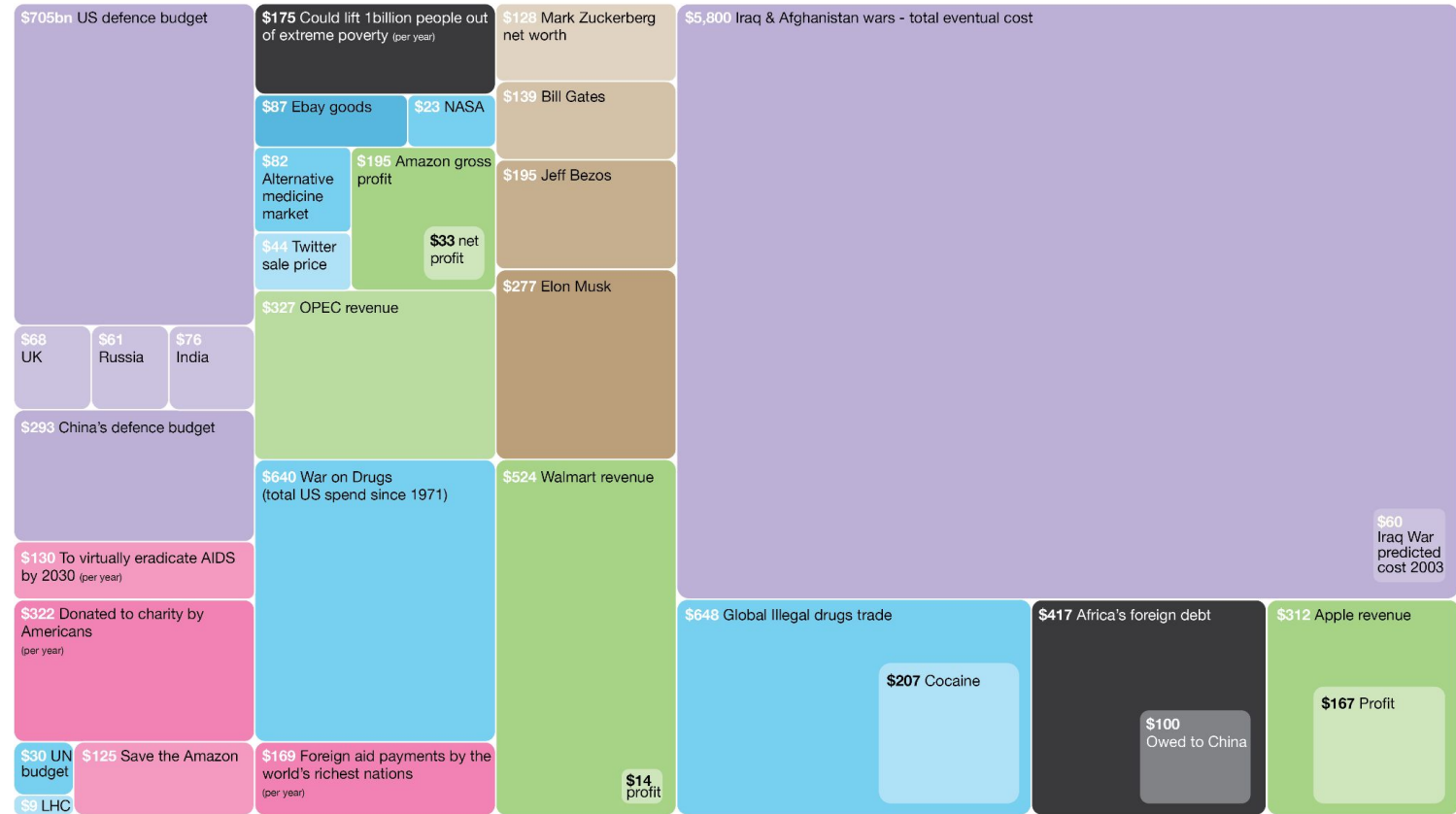Science of science

Scientometrics

Science and Technology Studies

Image: Ludo Waltman / RoRI

The Billion Dollar O Gram

accumulating · earning · fighting · giving · owing · spending

$705bn US defence budget

$68 UK

$61 Russia

$76 India

$293 China's defence budget

$130 To virtually eradicate AIDS by 2030 (per year)

$322 Donated to charity by Americans (per year)

$30 UN budget

$125 Save the Amazon

$9 LHC

$175 Could lift 1billion people out of extreme poverty (per year)

$87 Ebay goods

$23 NASA

$82 Alternative medicine market

$195 Amazon gross profit

$33 net profit

$44 Twitter sale price

$327 OPEC revenue

$640 War on Drugs (total US spend since 1971)

$169 Foreign aid payments by the world's richest nations (per year)

$128 Mark Zuckerberg net worth

$139 Bill Gates

$195 Jeff Bezos

$277 Elon Musk

$524 Walmart revenue

$14 profit

$5,800 Iraq & Afghanistan wars - total eventual cost

$60 Iraq War predicted cost 2003

$648 Global Illegal drugs trade

$207 Cocaine

$417 Africa's foreign debt

$100 Owed to China

$312 Apple revenue

$167 Profit

david mccandless
informationisbeautiful

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

RoЯi RESEARCH ON RESEARCH INSTITUTE

"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"

ukrn.org

Issues, problems, incidents & accidents

Researcher training

Career progression

Diversity and Inclusion

(ir)responsible Metrics

Research culture

Funding issues

Bias in peer review

Ecological (ir)responsibility

Reliability and reproducibility

Transparency

Data Sharing

Publishing

Public trust

Declining research gains