Better Research?

Tom Stafford (University of Sheffield)
2022-09-02

These slides: http://bit.ly/tom-talks
How The Research On Research Institute (Rori) Is Aiming To Improve How Research Is Funded, Practiced, Communicated, And Evaluated

The Research on Research Institute brings together scholars, research funders and publishers to coordinate research on all aspects of research funding, evaluation, quality and publication. RoRi is now funded for a second phase (2022-2027) and planning projects on all aspects of the research ecosystem. In this talk I will discuss various inefficiencies, biases and problematic features of how we do research, and the opportunities doctoral students have to contribute to the way research is changing - and should change! - to make research better quality, more efficient and more valuable.

https://tomstafford.staff.shef.ac.uk/

Read more about RoRI here https://researchonresearch.org/
Research

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?

These slides:
$2 trillion is invested globally in research every year

What does one TRILLION dollars look like?

These slides: http://bit.ly/tom-talks
$2 trillion is invested globally in research every year

What does one TRILLION dollars look like?

Image: pagetutor.com

These slides: http://bit.ly/tom-talks
The Billion Dollar O Gram

$705bn US defence budget
$175 Could lift 1 billion people out of extreme poverty (per year)
$123 Mark Zuckerberg net worth
$15,800 Iraq & Afghanistan wars - total eventual cost
$67 Ebay goods
$263 NASA
$293 China's defence budget
$78 Alternative medicine market
$327 OPEC revenue
$199 Amazon gross profit
$139 Bill Gates
$195 Jeff Bezos
$277 Elon Musk
$4 Twitter sale price
$33 net profit
$312 Apple revenue
$207 Cocaine
$417 Africa's foreign debt
$100 Owed to China
$167 Profit
$545 Global illegal drugs trade
$640 War on Drugs (total US spend since 1971)
$524 Walmart revenue
$130 To virtually eradicate AIDS by 2030 (per year)
$121 Save the Amazon (per year)
$16 Foreign aid payments by the world's richest nations (per year)

contrast
accumulating earning fighting giving owing spending

updated Jun 2022 // data bit.ly/billions2022

sources United Nations, Guardian, CNBC, Wikipedia and news reports

David McCandless
informationisbeautiful
Research on Research

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

Image: Ludo Waltman / RoRI
<table>
<thead>
<tr>
<th>Issues, problems, incidents &amp; accidents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Researcher training</td>
</tr>
<tr>
<td>Career progression</td>
</tr>
<tr>
<td>Diversity and Inclusion</td>
</tr>
<tr>
<td>(ir)responsible Metrics</td>
</tr>
<tr>
<td>Research culture</td>
</tr>
<tr>
<td>Funding issues</td>
</tr>
<tr>
<td>Bias in peer review</td>
</tr>
<tr>
<td>Ecological (ir)responsibility</td>
</tr>
<tr>
<td>Reliability and reproducibility</td>
</tr>
<tr>
<td>Transparency</td>
</tr>
<tr>
<td>Data Sharing</td>
</tr>
<tr>
<td>Publishing</td>
</tr>
<tr>
<td>Public trust</td>
</tr>
<tr>
<td>Declining research gains</td>
</tr>
</tbody>
</table>

https://researchonresearch.org/projects
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

https://researchonresearch.org/projects
Research evaluation is costly

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

“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)”
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

Research evaluation is unreliable


**ABSTRACT**

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).

**KEYWORDS:** peer-review, consistency, grant funding.
Funding allocation is biased

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


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

Image: By Unknown author - A Woodcut from Historiae celebriores Veteris Testamenti Iconibus representatae. , Public Domain, Wikimedia Commons
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.

“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.”

Improvements are possible

These slides: http://bit.ly/tom-talks
Percentage of First Time PIs by Observing Cycle

start of dual anonymous peer review


These slides: http://bit.ly/tom-talks
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/
A Research on Research Institute

These slides: http://bit.ly/tom-talks
### RoRI Phase 2: current partners update (22 June 2022)

<table>
<thead>
<tr>
<th>Confirmed Core Partner &amp; Member of the CIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Sheffield</td>
</tr>
<tr>
<td>Novo Nordisk Foundation</td>
</tr>
<tr>
<td>SNSF</td>
</tr>
<tr>
<td>Digital Science</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Confirmed Core Partner</th>
</tr>
</thead>
<tbody>
<tr>
<td>CWTS-Leiden</td>
</tr>
<tr>
<td>Volkswagen Foundation</td>
</tr>
<tr>
<td>CIHR</td>
</tr>
<tr>
<td>Michael Smith Health Research BC</td>
</tr>
<tr>
<td>FNR-Luxembourg</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Core Partner likely</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wellcome</td>
</tr>
<tr>
<td>Chan Zuckerberg Initiative</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Core or Project – live discussions</th>
</tr>
</thead>
<tbody>
<tr>
<td>NRF-South Africa</td>
</tr>
<tr>
<td>NWO &amp; Zon MW (Netherlands)</td>
</tr>
<tr>
<td>UK Research &amp; Innovation (UKRI)</td>
</tr>
<tr>
<td>Cancer Research UK</td>
</tr>
<tr>
<td>SSHRC-Canada</td>
</tr>
<tr>
<td>British Academy</td>
</tr>
<tr>
<td>Sloan Foundation</td>
</tr>
<tr>
<td>Moore Foundation (US)</td>
</tr>
<tr>
<td>India Alliance</td>
</tr>
<tr>
<td>Israel Science Foundation</td>
</tr>
<tr>
<td>African Academy of Sciences</td>
</tr>
</tbody>
</table>
Gert V Balling, Katrin Milzow and Sarah de Rijcke (Co-Chairs of RoRI)
Funder experiments are co-produced

These slides: http://bit.ly/tom-talks
Experiment with evaluation

Inventory of challenges with the implementation of narrative CVs

Key ideas: Create an inventory of challenges related to the implementation of narrative CVs across different funders, and develop strategies to tackle them – e.g. inertia of established indicator-centric review practices, resistance to the narrative format by (some) reviewers, inertia of organizational culture, technical difficulties, effects of disciplinary culture on use etc.

Infrastructures & Data Sharing

Research on Research on Research: a meta RoRI case study

Key idea(s): What can we learn from our uncommon modes of working as a large scale multi-stakeholder research consortium that will add to knowledge on practice/research co-produced projects? Considering different vectors of engagement, what lessons have we learned about bringing different knowledge types together, conducting co-produced

Research Priorities & Portfolios

Engaging the citizen. The impact of crowd knowledge on science

Key idea(s): A research study funded by a Marie Curie grant, led by Cindy Lopez-Bento (Head of Science of Science at FNR and Professor at the University of Leuven). The project...aims to increase the understanding we have of crowd science. ...
an approach that allows a wide base of (non-scientific) volunteers to participate in research projects. The project will, for example, use

Impacts, Indicators & Culture Change

The productivity and public value of research

Key idea(s): Over the last decade we have seen a tendency to connect more directly the setting of research agendas with the attainment of socio-economic goals through devices such as "grand challenges" and "mission driven research". At the same time, there are some indications that

About this board

Please comment and vote on the cards, as well as adding your own ideas.

Colour coding:
Orange: Continuation and development of phase 1 projects
Purple: Ideas from pilot phase long-list
White: New ideas
Blue: Infrastructure, networks and capacity-building
Experimentation – RoRI Funder’s Lab

The experimental
Research funder’s handbook
A RoRI publication
Bendiscioli, Sandra; Firpo, Teo; Bravo-Biosca, Albert; Czibor, Eszter; Garfinkel, Michele; Stafford, Tom; et al. (2022):


Part 1

The case for experimental research funding

1. Summary 13
2. Why experiment? 14
3. Tools for experimenting with research funding 19
   3.1 Tools to diagnose 20
   3.2 Tools to design 26
   3.3 Tools to evaluate 30
4. Becoming more experimental 37
   4.1 Attracting applicants 39
   4.2 Selecting reviewers 46
   4.3 Assessing proposals 52
   4.4 Making funding decisions 61
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
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
We’re hiring! (soon)

- From October, at least 1 PDRA, contract until 2027, full time
- would suit researcher with a background in experiment design, quantitative analysis, and/or metascience (research on research)
- based at RoRI and The Department of Psychology, University of Sheffield, UK
- opportunities to design & lead projects, journal publications; grow collaborations; mentoring
- details [https://tomstafford.staff.shef.ac.uk/?p=954](https://tomstafford.staff.shef.ac.uk/?p=954)
- informal enquiries welcome t.stafford@sheffield.ac.uk


Image: Google Maps
Join the conversation, get involved...

http://researchonresearch.org
RESERVE SLIDES
FOLLOW
(not for show)
RoRI work on partial randomisation
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.
RoRI Working Paper No. 7
Why draw lots? Funder motivations for using partial randomisation to allocate research grants

Helen Buckley Woods and James Wilidon
December 2021

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

Stanford survey 2018 of biosciences PhD

low efficacy in identify biases, study design, data management, statistics
“we investigated the precision and reproducibility of a prototypical perturbational experiment performed in cell lines: drug dose-response as measured by cell viability.”

“most examples of irreproducibility are themselves irreproducible and that technical factors responsible for any specific outlier measurement are difficult to pin down”

"we found that irreproducibility most commonly arose from unexpected interplay between experimental protocol and true biological variability"

across-lab standardization of electrophysiological procedures can lead to reproducible results across labs

RoRI
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
Phase 2 Structure

Core partners
12-15 research funders, universities and technology providers. Together these partners govern and sustain the core functions of RoRI.

Partnership board
Core partners (including those who are CIC members) Strategic direction, priorities, management of funds

Sub-committees of Partnership Board
Including Audit & Finance

Executive team
Day-to-day operations and oversight of projects

RoRI research fellows
Lead individuals from within Core Partners Dedicated time on RoRI projects

RoRI nodes / labs
Smaller nodes or labs hosted by core partners Operating at national or regional system level

RoRI Ltd
A Community Interest Company (CIC) under UK law. Brings legal and bureaucratic simplification Holds funds. Owns, manages and preserves infrastructure and other IP. Devolves all other responsibilities to the Partnership Board.

Board of directors
Nominated from CIC members. Light touch governance: audit, accounts

Infrastructure
Funder Data Platform, FAIRware etc

Project partners
Funders, scholarly communication platforms & publishers; researchers; tech providers. Ad hoc project participation & support.

International advisory group
Academic, policy, funder & scholarly communication leaders. Source of informal cross-level guidance and advice

Associate faculty
Leading researchers from universities and institutions worldwide who collaborate on specific RoRI projects
Figure S9: Accumulation of grant money by applicants