Research Integrity, Replicability and Transparency

Lex Bouter

25 minutes, including 5 minutes for Q+A

http://bit.ly/2nVbJRC

http://www.allea.org/wp-content/uploads/2017/03/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf

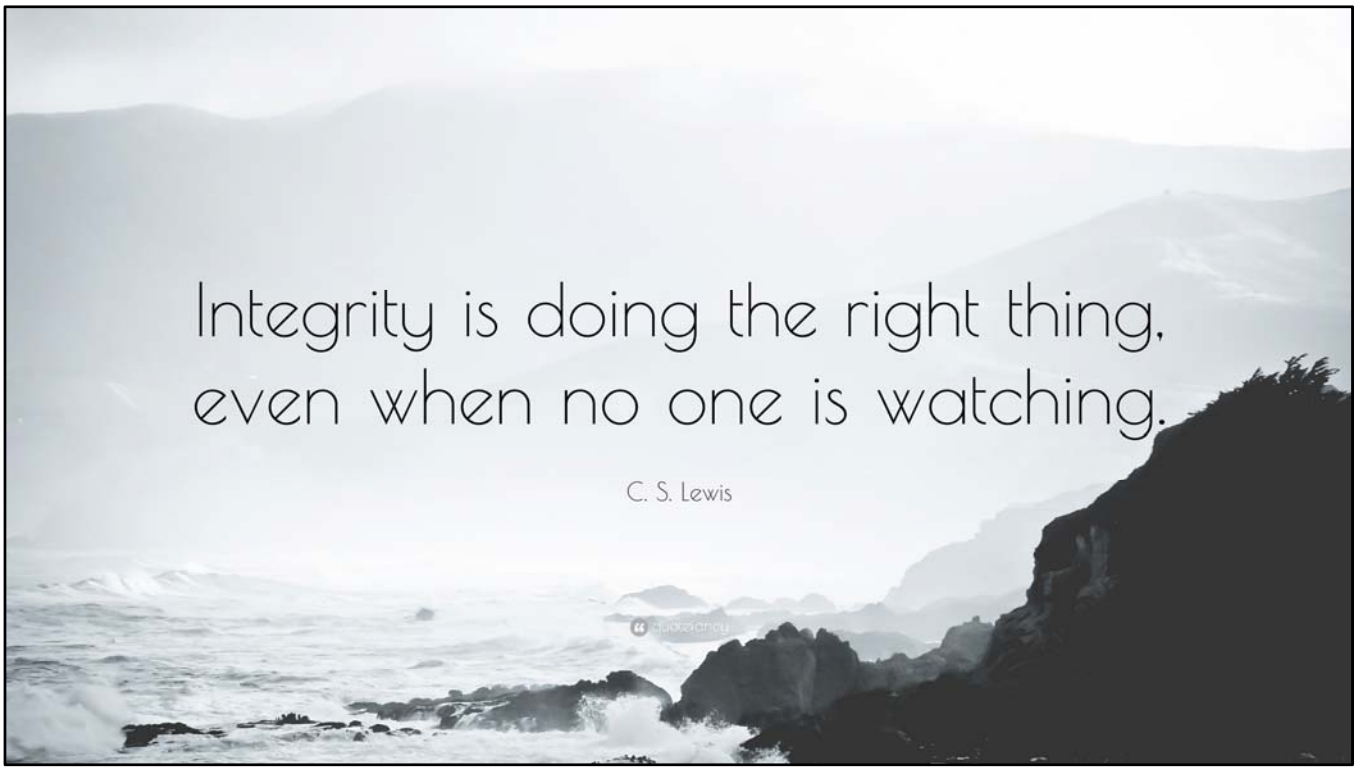
Published in 2017 and made mandatory for research sponsored by the EU (e.g. ERC and Horizon 2020).

Contains 5 principles (values), 61 standards (norms) and 21 duties of care for institutions and prescribes how to deal with allegations af breaches of research integrity.

http://www.vsnu.nl/files/documents/Netherlands%20Code%20of%20Conduct%20for%20Research%20Integrity%202018.pdf

http://www.vsnu.nl/files/documenten/Nederlandse%20gedragscode%20wetenschappelijke%20integriteit%202018.pdf

https://doi.org/10.17026/dans-2cj-nvwu



This seems to be and instance of inadequate referencing by repeating the an incorrect attribution.

So already erly in this presentation we encounter an instance of (unintentional) plagiarism.

This quote is NOT of C.S. Lewis – probably it is a paraphrase of a Charles Marshall quote in Shattering the Glass Slipper.

http://www.cslewis.org/aboutus/faq/quotes-misattributed/
http://www.essentialcslewis.com/2015/11/22/ccslq-13-right-thing/

Functioning of moral compass depends on:

- Individual virtuousness
- Research climate
- Perverse incentives

Codes of conduct are mainly about what researchers should do themselves.

But there are also strong determinants in the research climate and the system of science.

That doesn't decrease the personal responsibility to behave well in research. In fact it makes personal responsilility larger: individual researchers also have to help to improve the research climate and the system of science.

Spectrum of research practices

How it should be done:

Relevant, Valid, Reproducible, Efficient

Sloppy science: **34%**

Ignorance, honest error or dubious integrity

Scientific fraud: **2%**

Fabrication, Falsification, Plagiarism

3

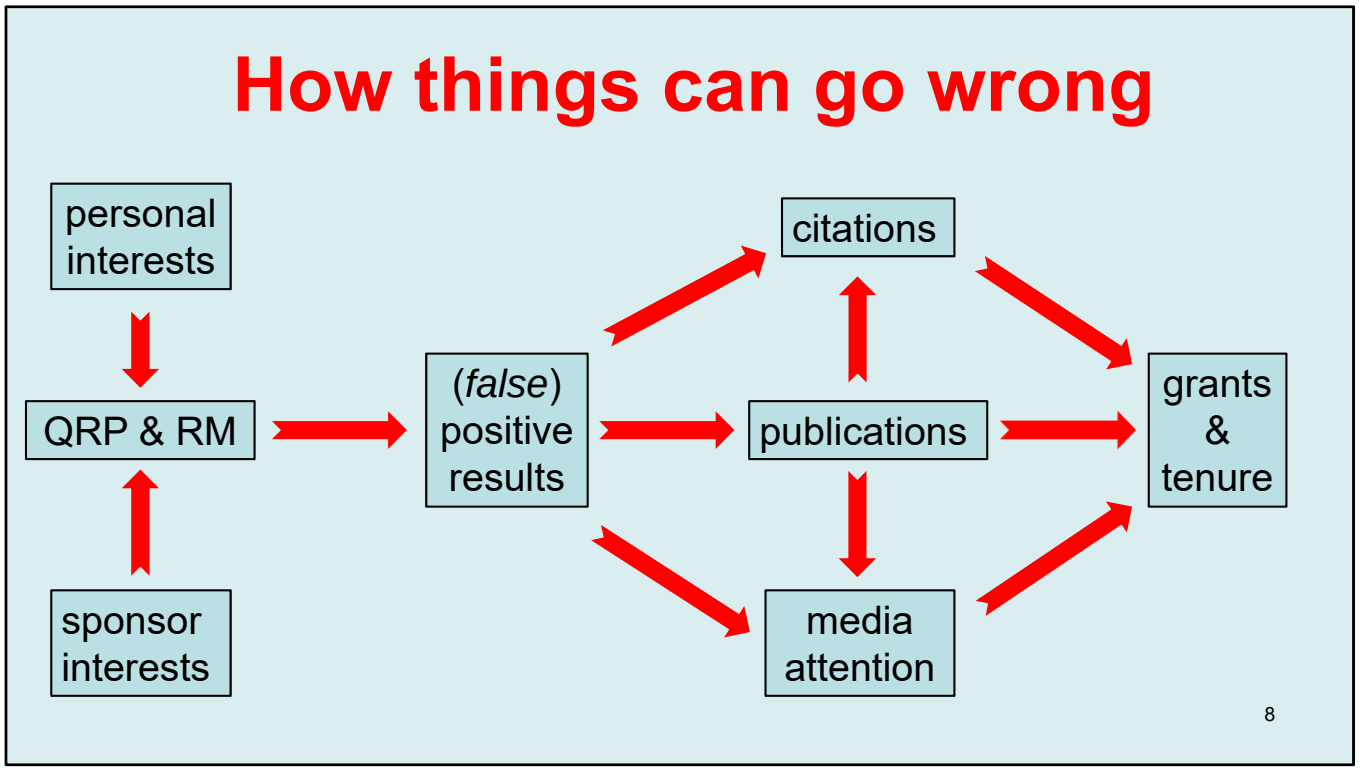
The figures concern the question 'did you at least once in the last 3 years engage in FF / QRP ?' and come from the highly cited meta-analysis:

Fanelli D. How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data. PLoS ONE 2009; 4(5): e5738

We need to talk about the elephants in the room

- Selective reporting
- Poor replicability

7



This slide shows – in a simplified way – how things can go wrong.

In most disciplines the proportion of papers reporting positive results increases over time. Positive results are published and cited more often, and also get more media attention. This will probably increase the likelihood of getting grants and tenure. We have also some evidence that conflicts of interest and sponsor interests may lead to sloppy science or worse. QRP and RM can effectively help to get (false) positive results.

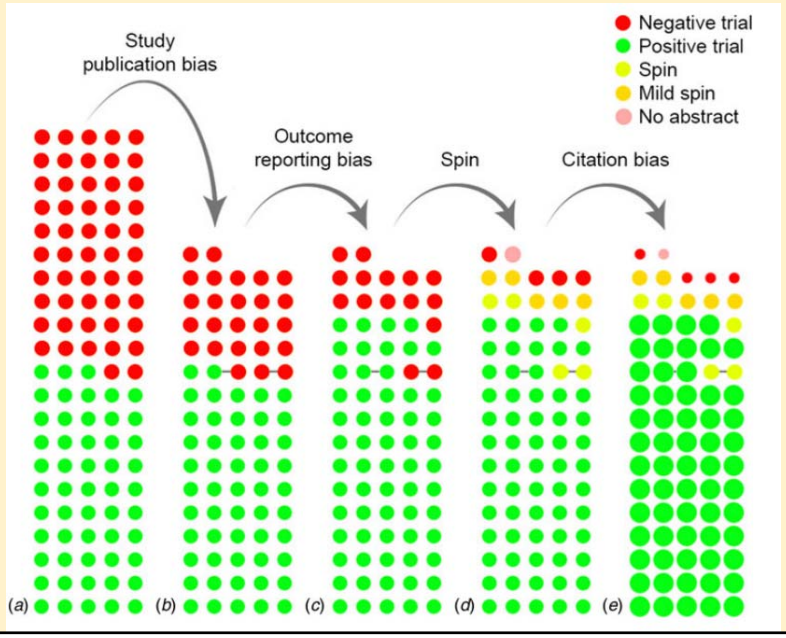
Negative findings are so unpopular that often these are not reported at all. This mechanism will lead to publication bias, selective reporting and selective citation. Especially small studies with positive outcomes will predominantly be chance findings. These phenomena will distort the published record and can explain the large replication difficulties some fields (e.g. preclinical research) experience.

Personal interests and sponsor interests can lead to QRP and RM also if researchers are not aware of it. Many of us want to please our sponsor with a view to motivate them to keep funding our work. That could lead for instance to subtle flaws in the study design, to selective reporting and to spin in the report of the results of the study.

There is evidence for some of the relations suggested in this slide, but no or only little

evidence for most of them. We really need more solid empirical research to clarify how these things work. Gaining this knowledge is important for effectively fostering RCR and preventing QRP and RM.

How negative results disappear from the published literature



9

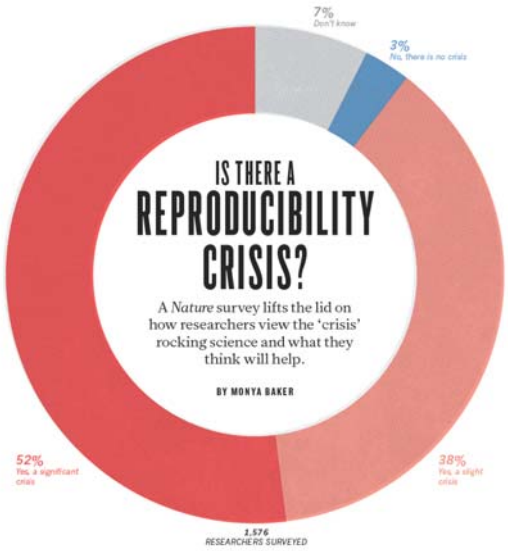
de Vries YA, Roest AM, de Jonge P, Cuijpers P, Munafò MR, Bastiaansen JA (2018). The cumulative effect of reporting and citation biases on the apparent efficacy of treatments: the case of depression. Psychological Medicine 1–3. https://doi.org/10.1017/S0033291718001873

This example concerns the fate of an inception cohort of 105 RCTs of the efficacy of anti-depression drugs from the FDA database. The cohort is complete in the sense that pharmeceutical company must register all trials it intends to use to obtain FDA approval before embarking on data collection. The FDA considered 50% of the trials to be positive after carefully looking at the results.

Replicability of studies is only 10-40 %

FOOLING OURSELVES

HUMANS ARE REMARKABLY GOOD AT SELF-DECEPTION. BUT GROWING CONCERN ABOUT REPRODUCIBILITY IS DRIVING MANY RESEARCHERS TO SEEK WAYS TO FIGHT THEIR OWN WORST INSTINCTS.



Just two Nature headlines. The topic draws attention, and rightly so.

Nuzzo - Fooling ourselves - Nature 2015;526 182-185
Baker - Is there a replicability crisis - Nature 2016; 533 452-4

Let me just give one example, although it's a quite spectacular one.

This is the title of a alarming article in Nature a few years ago.

The authors tried to replicate 53 widely cited high impact preclinical studies on potential new cancer treatments – suprise, surprise, they were all positive.

If needed they even went into the original labs and tried to replicate the study there together with the original PIs.

Begley CG, Ellis LM. Raise standards for preclinical cancer research. Nature 2012; 483: 531-3
Begley - Six red flags for suspect work - Nature 2013; 497 433-4
Begley, Ioannidis - Reproducibility in science - Circulation Research 2015; 116 116-26

Only 6 of 53 preclinical landmark cancer studies could be confirmed by replication

When negative studies are rarely published,
published positive studies are likely to be chance findings

Non-confirmed studies

- sometimes inspire many new studies → **waste of resources**!
- sometimes lead to clinical trials → **unethical situation**!

12

Selective reporting of animal studies is a huge problem, leading to a embarrassing lack of replicability.

The issue is that positive chance findings are difficult to reproduce.

There are many more examples of serious replication problems of preclinical studies.

Begley CG, Ioannidis JPA. Reproducibility in science. Circulation Research 2015; 116 116-26

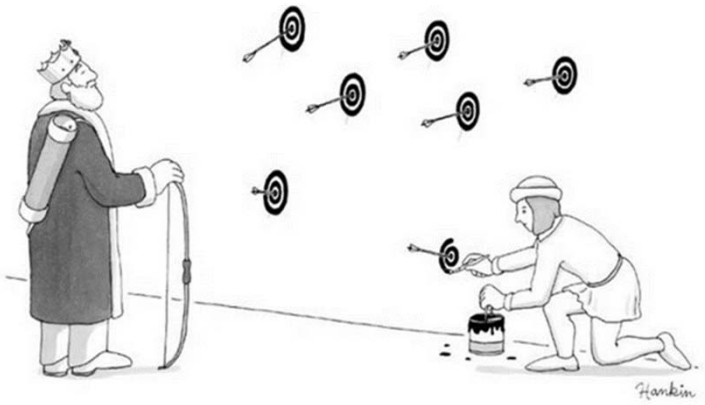
Two videos of John Ioannidis lecturing about reproducibility:

https://www.youtube.com/watch?v=UbQCNOGkc6w
https://www.youtube.com/watch?v=GPYzY9I78CI

Important causes of 'replicability crisis'

- Selective reporting
- Low power
- P-hacking
- HARKing

Hypothesizing After Results are Known



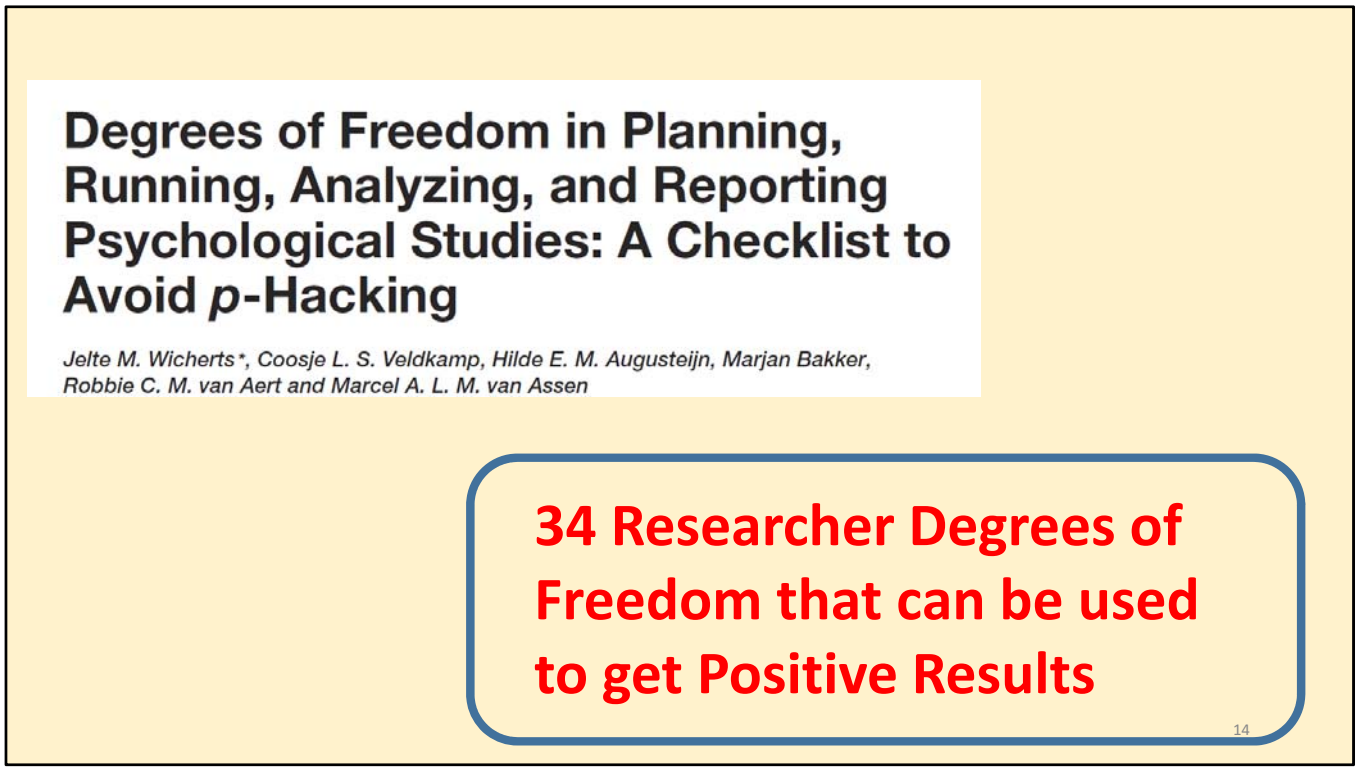
13

Wicherts et al - Degrees of freedom - checklist to avoid p-hacking - Front Psych 2016; 7 1832

Nosek et al - The preregistration revolution - PNAS 2018; 115 2600-6

Bouter - Fostering responsible research practices is a shared responsibility of multiple stakeholders - J Clin Epidemiol 2018; 93 143-6

Ioannidis - Why replication has more scientific value than original discovery - Behavioral and Brain Sciences 2018; 41 e137
http://www.nature.com/articles/s41562-016-0021



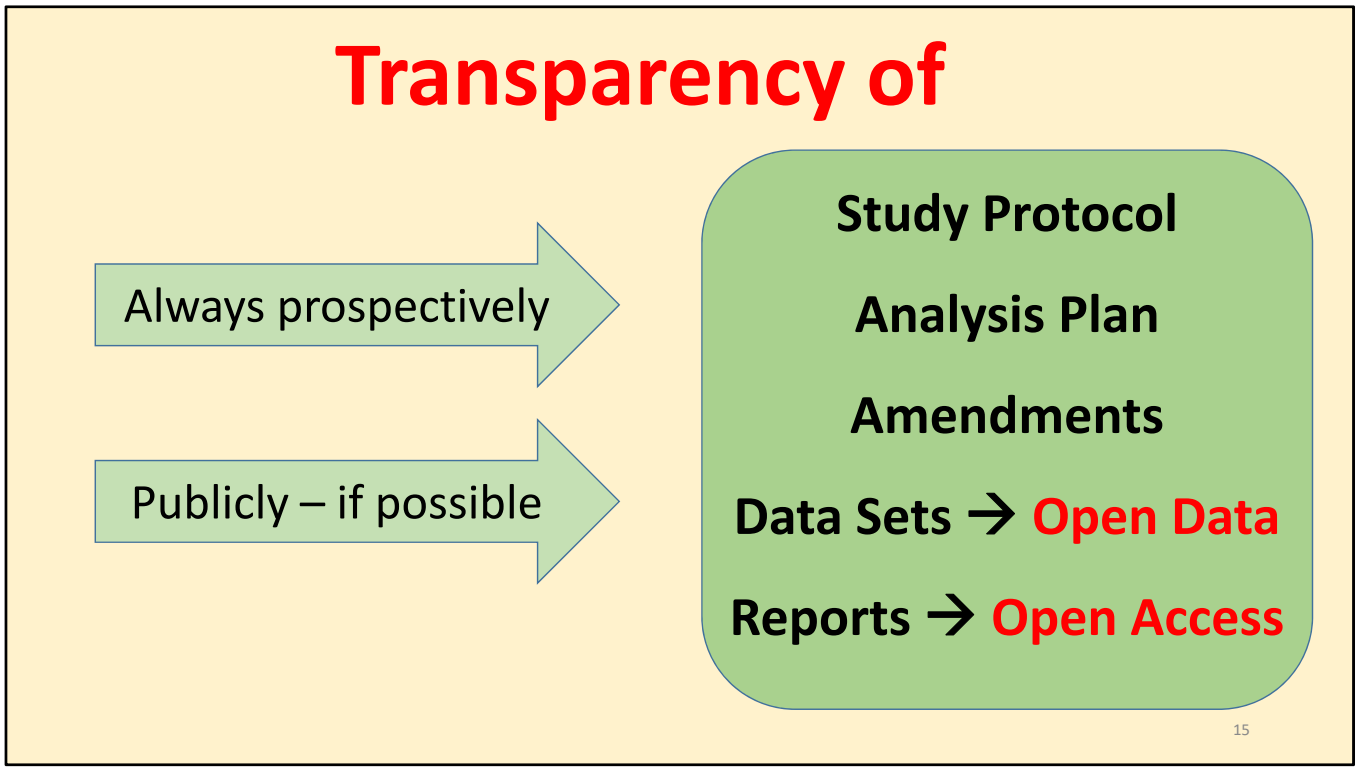
Wicherts et al - Degrees of freedom - checklist to avoid p-hacking - Front Psych 2016; 7 1832

This wonderful article comes from the faculty where Diederik Stapel was dean: never waste a good crisis.

The idea of Researcher Degrees of Freedom indicates that sloppy science offers a lot of room to get the findings and conclusions you want.

Please note: we're talking about hypothesis testing research (confirmatory research), NOT about exploratory research. In the latter domain 'anything goes' as long as it's clearly stated that exploration is at issue.

See also: Wicherts – The weak spots of contemporary science (and how to fix them) - Animals 2017, 7, 90; doi:10.3390/ani7120090

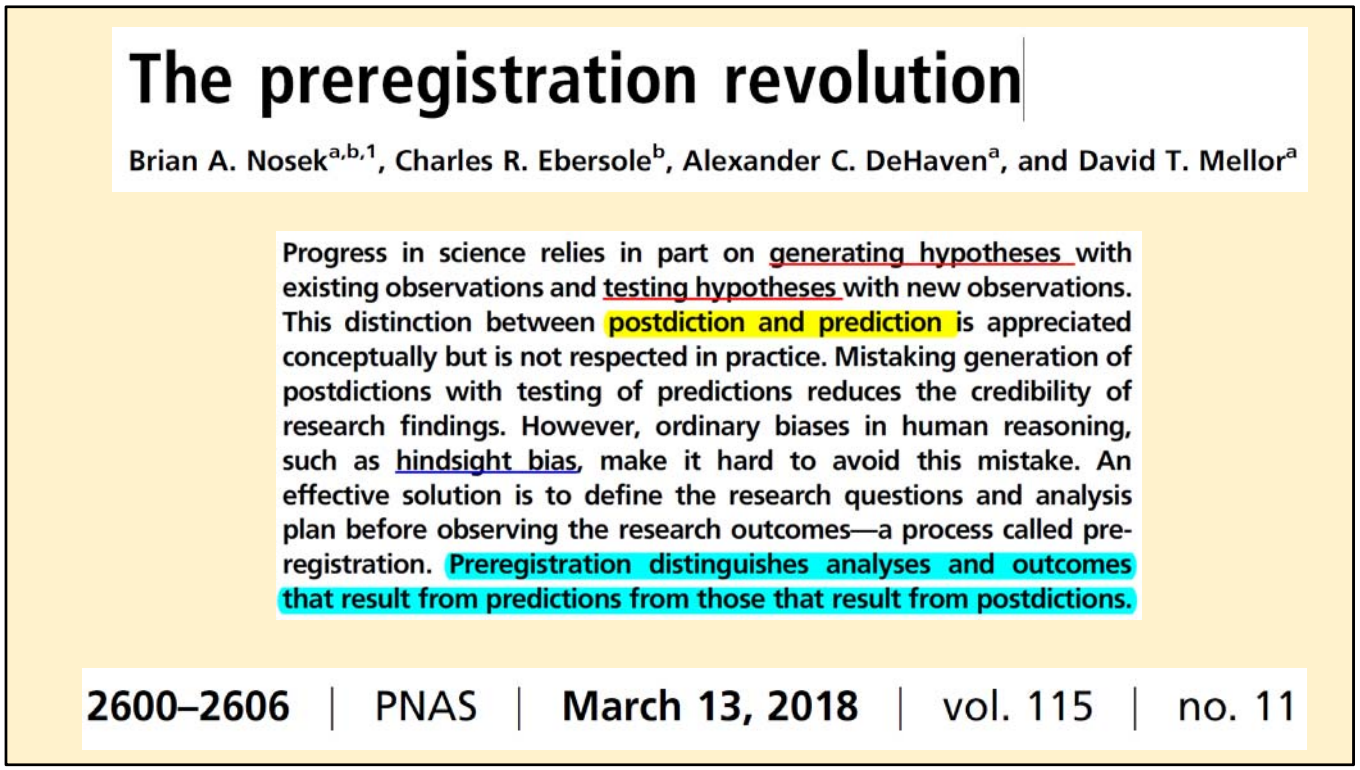


In theory the solution is easy and takes the form of ensuring that all research findings are published and the whole process is transparent, meaning that all steps can be checked and reconstructed. Studies need to be preregistered and a full protocol must be uploaded in a repository before the start of data collection. Similarly a data-analysis plan, syntaxes, data sets and full results need to be uploaded. Amendments and changes are possible but should always leave traces, thus enabling users to identify actions that were potentially data-driven. While ideally these elements of transparency are publicly accessible, there are many situations where delayed, conditional or incomplete access is indicated. But that does not detract from the principle of full transparency: even the process and outcomes of highly classified research for the defence industry should if necessary be made available for a thorough check by an investigation committee that is bound by confidentiality.

Bouter LM. Perverse incentives and rotten apples. Accountability in Research 2015; 22:148-161.
Bouter LM. Open data is not enough to realize full transparency. J Clin Epidemiol 2016; 70: 256-7.
Bouter LM. Fostering responsible research practices is a shared responsibility of multiple stakeholders. Journal of Clinical Epidemiology 2018; 96: 143-6.
ter Riet G, Bouter LM. How to end selective reporting in animal research. In: Martic-Kehl MI, Schubiger PA, eds. Animal models for human cancer: discovery and development of

novel therapeutics. First edition. Weinheim: Wiley, 2016: 61-77.

See also: Nosek BA, Ebersole CR, DeHaven AC, Mellor D. The preregistration revolution. PNAS 2018;115:2600-6 - http://www.pnas.org/content/115/11/2600



The preregistration revolution

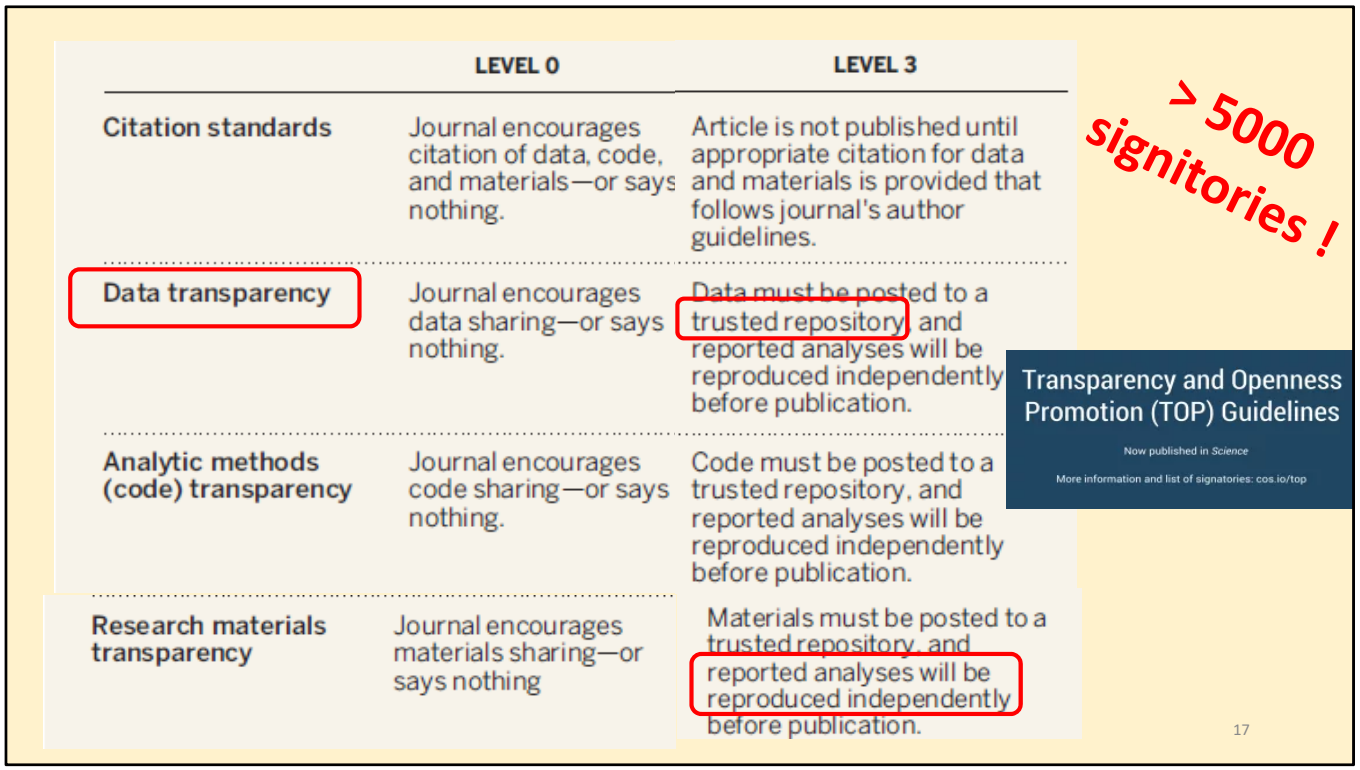
Brian A. Nosek[a,b,1], Charles R. Ebersole[b], Alexander C. DeHaven[a], and David T. Mellor[a]

Progress in science relies in part on generating hypotheses with existing observations and testing hypotheses with new observations. This distinction between postdiction and prediction is appreciated conceptually but is not respected in practice. Mistaking generation of postdictions with testing of predictions reduces the credibility of research findings. However, ordinary biases in human reasoning, such as hindsight bias, make it hard to avoid this mistake. An effective solution is to define the research questions and analysis plan before observing the research outcomes—a process called preregistration. Preregistration distinguishes analyses and outcomes that result from predictions from those that result from postdictions.

2600–2606 | PNAS | March 13, 2018 | vol. 115 | no. 11

Nosek BA, Ebersole CR, DeHaven AC, Mellor D. The preregistration revolution. PNAS 2018;115:2600-6.

http://www.pnas.org/content/115/11/2600



Nosek et al - Promoting an open research culture - Science 2015; 348 1422-5
https://cos.io/our-services/top-guidelines/

Registered Reports

Adopted by 204 journals

18

Chambers et al - Instead of playing the game its time to change the rules - registered reports - AIMS Neuroscience 2014; 1 4-17

Chambers - Ten reasons why journals must review manuscripts before results are known - Addiction 2015; 110 10-11

https://cos.io/our-services/registered-reports

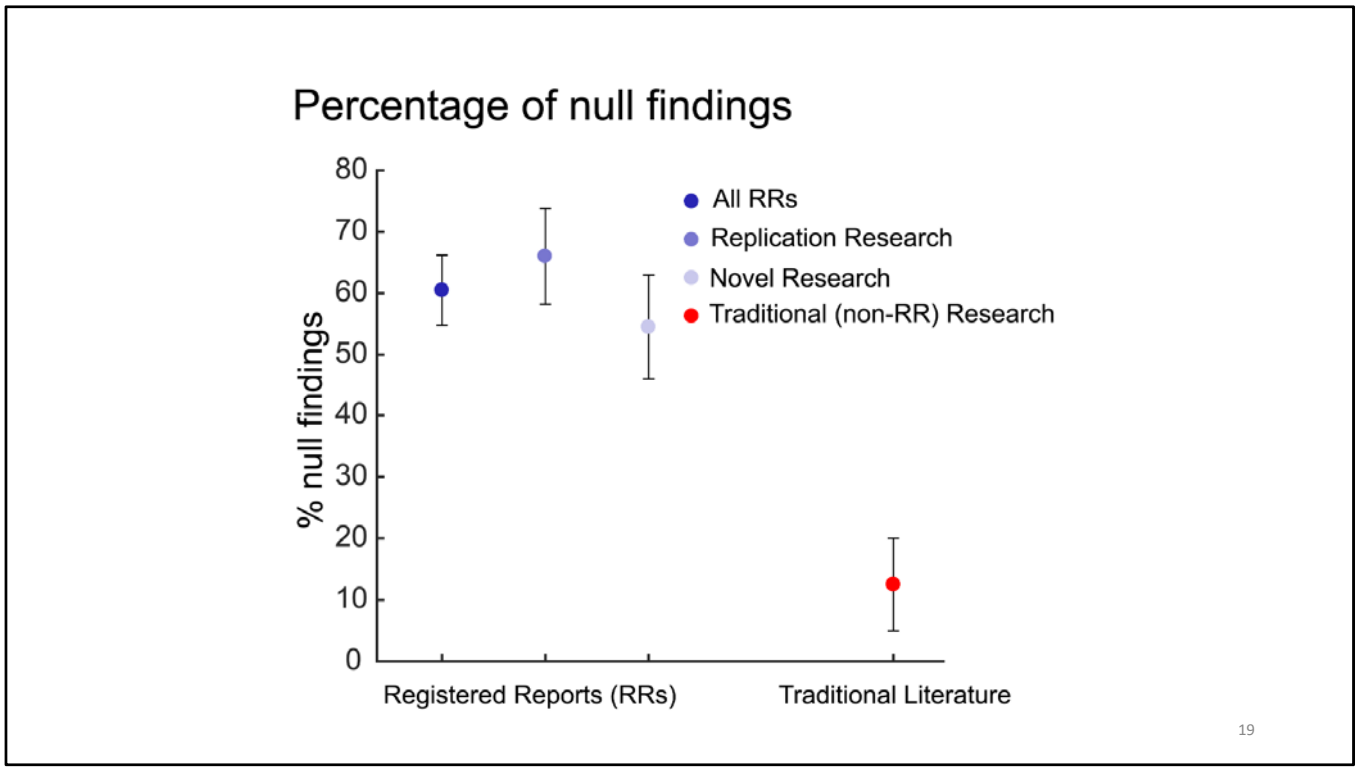
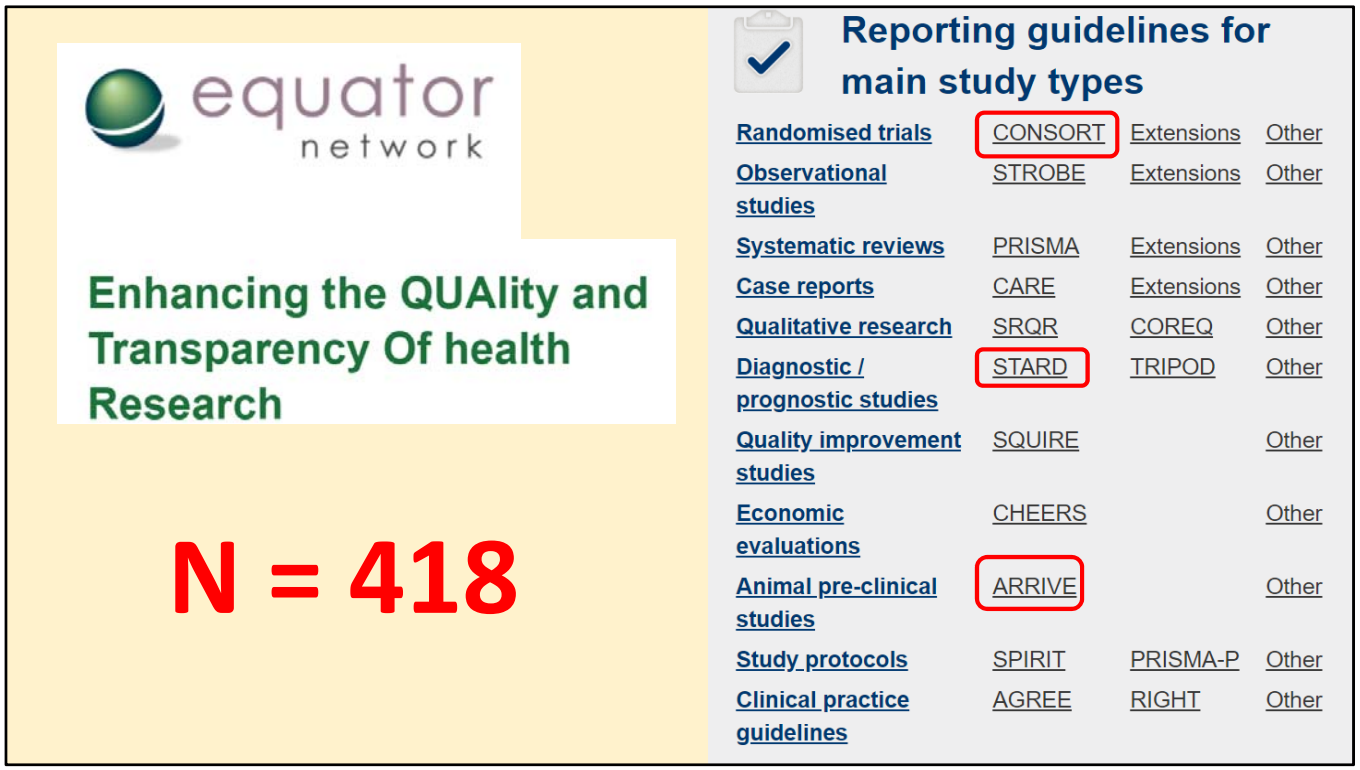


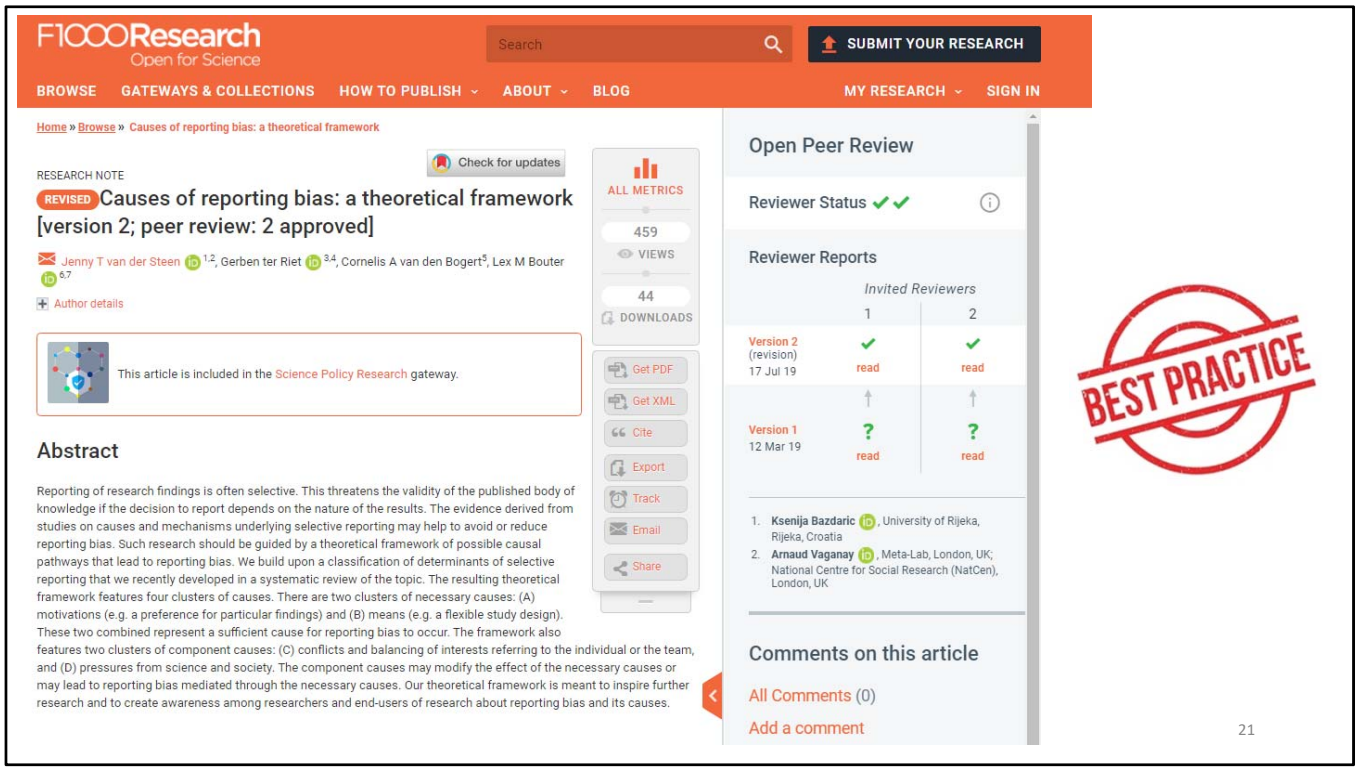
Fig 1
Percentages of null findings among RRs and traditional (non-RR) literature [46,47], with their respective 95% confidence intervals.

In total, we extracted n = 153 hypotheses from RRs that were declared as replication attempts and n = 143 hypotheses that were declared as original research. The bounds of the confidence intervals shown for traditional literature were based on estimates (5% and 20%, respectively) of null findings that have been previously reported for traditional literature [46,47]. Data is available on the Open Science Framework (https://osf.io/wy2ek/) and in S1 Data. RR, registered report.

Allen C, Mehler DMA (2019) Open science challenges, benefits and tips in early career and beyond. PLoS Biol 17(5): e3000246. https://doi.org/10.1371/journal.pbio.3000246



http://www.equator-network.org/


F1000Research
Open for Science
Search
SUBMIT YOUR RESEARCH
BROWSE
GATEWAYS & COLLECTIONS
HOW TO PUBLISH
ABOUT
BLOG
MY RESEARCH
SIGN IN
Home » Browse » Causes of reporting bias: a theoretical framework
Check for updates
RESEARCH NOTE
REVISED
Causes of reporting bias: a theoretical framework [version 2; peer review: 2 approved]
Jenny T van der Steen [1,2], Gerben ter Riet [3,4], Cornelis A van den Bogert[5], Lex M Bouter [6,7]
Author details
This article is included in the Science Policy Research gateway.
ALL METRICS
459
VIEWS
44
DOWNLOADS
Get PDF
Get XML
Cite
Export
Track
Email
Share
Abstract
Reporting of research findings is often selective. This threatens the validity of the published body of knowledge if the decision to report depends on the nature of the results. The evidence derived from studies on causes and mechanisms underlying selective reporting may help to avoid or reduce reporting bias. Such research should be guided by a theoretical framework of possible causal pathways that lead to reporting bias. We build upon a classification of determinants of selective reporting that we recently developed in a systematic review of the topic. The resulting theoretical framework features four clusters of causes. There are two clusters of necessary causes: (A) motivations (e.g. a preference for particular findings) and (B) means (e.g. a flexible study design). These two combined represent a sufficient cause for reporting bias to occur. The framework also features two clusters of component causes: (C) conflicts and balancing of interests referring to the individual or the team, and (D) pressures from science and society. The component causes may modify the effect of the necessary causes or may lead to reporting bias mediated through the necessary causes. Our theoretical framework is meant to inspire further research and to create awareness among researchers and end-users of research about reporting bias and its causes.
Open Peer Review
Reviewer Status
Reviewer Reports
Invited Reviewers
1
2
Version 2 (revision) 17 Jul 19
read
read
Version 1 12 Mar 19
read
read
1. Ksenija Bazdaric, University of Rijeka, Rijeka, Croatia
2. Arnaud Vaganay, Meta-Lab, London, UK; National Centre for Social Research (NatCen), London, UK
Comments on this article
All Comments (0)
Add a comment
BEST PRACTICE
21

https://www.reward-equator-conference-2020.com/

www.nrin.nl

@NRIN_Integrity

23