Keynote lecture of 60 minutes, including 10-15 minutes of discussion.

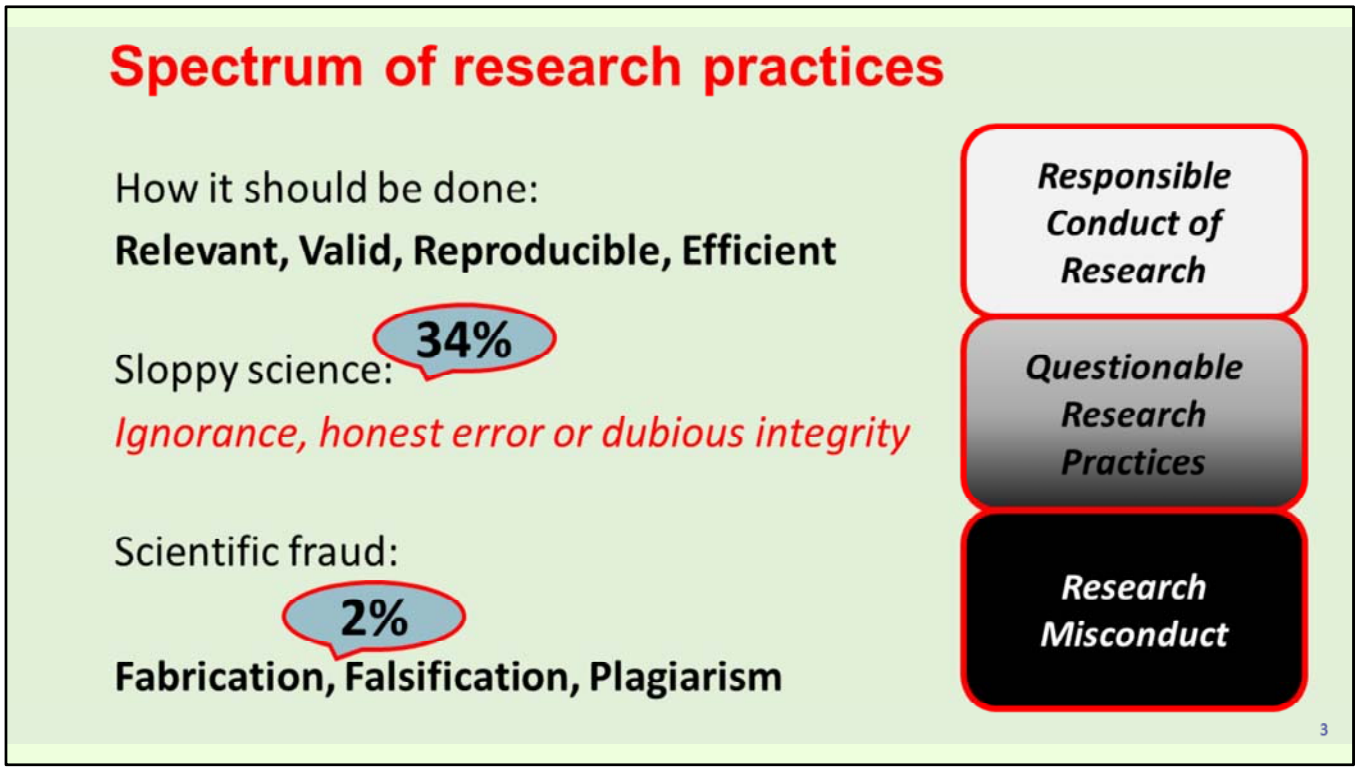
Content

- Research Integrity
- Selective reporting and replication crisis
- Plea for transparency
- What can institutions, journals and funders do?

2

Let me start with a disclaimer. Meta-research is still very much in its infancy. Much of what I have to say is not yet firmly evidence-based.

I have made available a PDF of my presentation to the organizers, containing in the note fields all references to articles and websites I mention. This PDF can be requested from the organizers.



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)

Bouter *et al. Research Integrity and Peer Review* (2016) 1:17
DOI 10.1186/s41073-016-0024-5

Research Integrity and
Peer Review

RESEARCH Open Access

Ranking major and minor research misbehaviors: results from a survey among participants of four World Conferences on Research Integrity

Lex M. Bouter[1,2*], Joeri Tijdink[2,3], Nils Axelsen[4], Brian C. Martinson[5] and Gerben ter Riet[6]

Top 5 – aggregated impact

1. Insufficiently ***supervise*** or mentor junior coworkers
2. Insufficiently ***report study flaws and limitations***
3. Keep ***inadequate notes*** of the research process
4. Turn a ***blind eye to putative breaches*** of research integrity by others
5. Ignore basic principles of ***quality assurance***

5

FFP are nowhere near this top 5 – based on frequency weighted impact on validity (frequency X impact on validity scores).

These are all questionable research practices that research institutes can do something about.

Fabrication and Falsification

Rank numbers

item	freq	truth	freq x truth
Delete data before performing data analysis without disclosure	*45*	*6*	*19*
Selectively delete data, modify data or add fabricated data after performing initial data-analyses	*50*	*2*	*24*
Fabricate data	*59*	*1*	*34*

Please note that F+F are not in the top 5 freq X truth scores – in fact they are in the middle range of the list of 60 items (high impact – low frequency)

Plagiarism

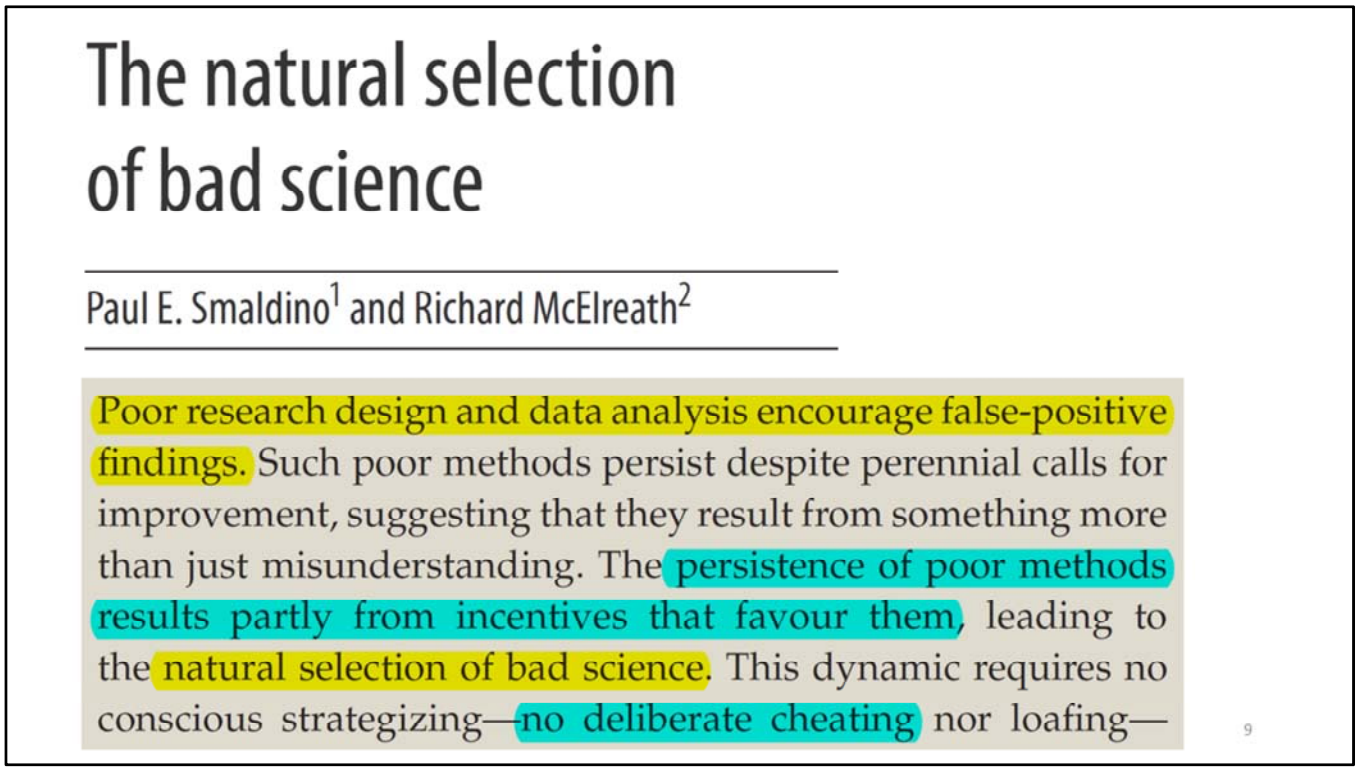
Rank numbers

item	freq	truth	freq x truth
Use published phrases or ideas of others without referencing	12	43	26
Re-use parts of your own publications without referencing	15	55	46
Use unpublished phrases or ideas of others without their permission	21	41	33
Re-use of previously published data without disclosure	29	46	36
Duplicate publication without disclosure	36	49	48

Please note that Plagiarism not in the top 5 freq X truth scores – in fact they are in the middle range of the list of 60 items (high impact – low frequency)

But of course the picture is different when the focus is on the impact on trust and the freq X trust scores.

Many rewards are linked to having positive and spectacular results as these are published more easily in high impact journals and will be cited more often. The various QRP have in common that they can effectively help to get these positive and spectacular results.



The survival value of cheating in science is probably substantial.

Smaldino et al - The natural selection of bad science - Royal Society Open Science 2016; 3 160384

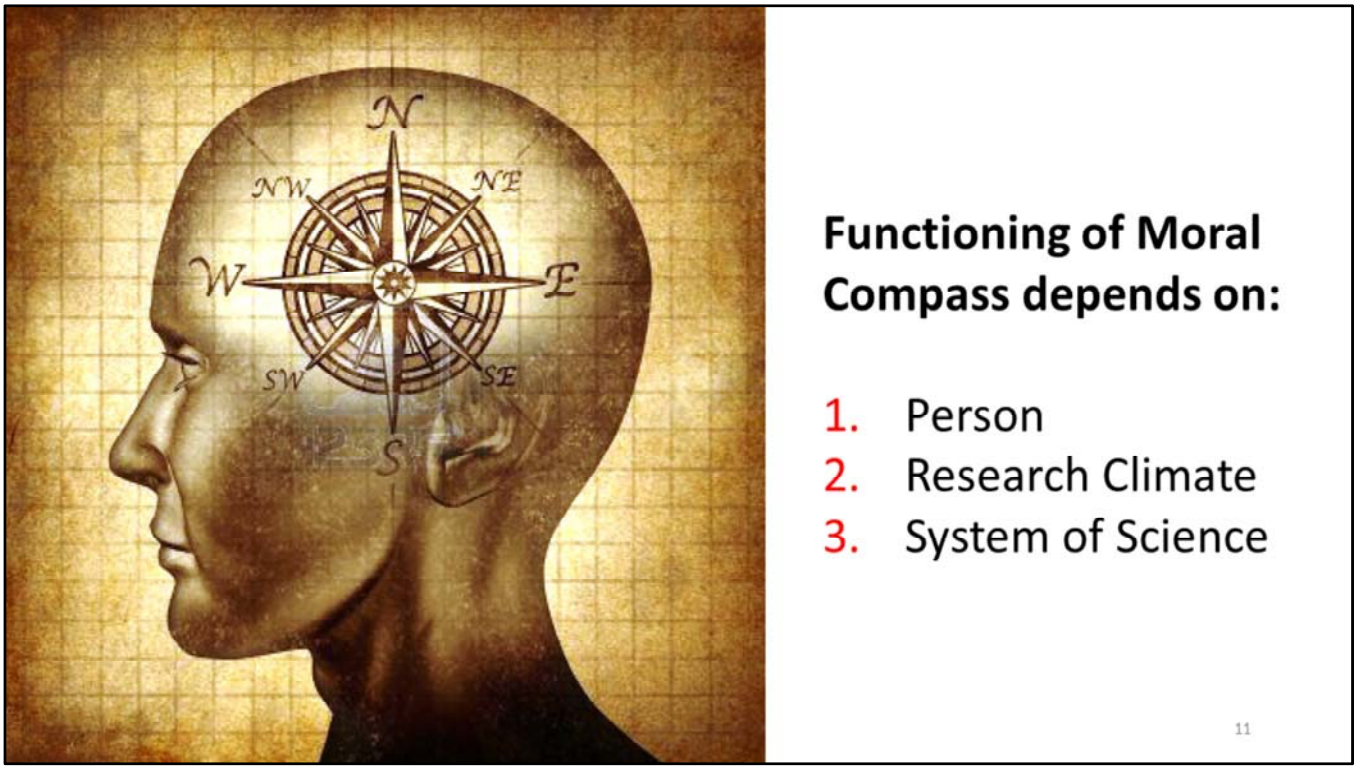
Moral Compass

Functioning depends on:

1. Quality and stability of the compass
2. Compensation for iron and electromagnetic fields nearby
3. Influence of earth magnetic field and its local deviations

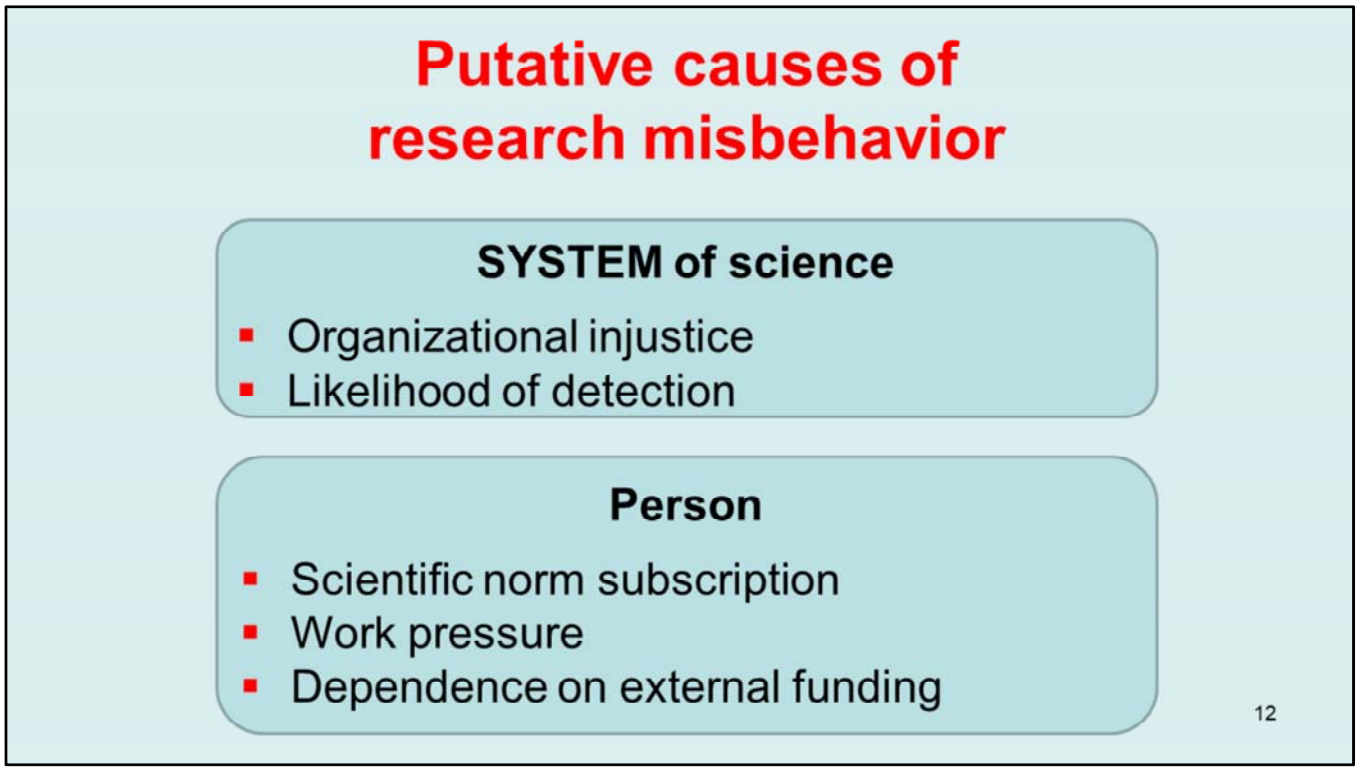
10

Research integrity concerns the behavior of scientists. They are guided by the moral compass in their head.



The most important stakeholders are the scientists themselves. Breaches of research integrity and sloppy science result from their professional behaviour. For that behaviour they are responsible as they are for fostering research integrity in their own work and that of their colleagues, first and foremost the PhD students and others they supervise or mentor.

But the behaviour of scientists is of course to a large extent driven by what happens in their environment. Both in the local research climate and in the system of science at large important determinants can be identified. Sadly, some of these can act as a perverse incentive. Empowering the scientist and optimizing the incentives is the responsibility of the other stakeholders. They should together make it more easy to live up to the standards and more difficult to misbehave.



This is all about perceptions, which can deviate substantially from more objective measures.

But perceptions are what matters as these probably drive behaviour to a large extent.

Putative causes of research misbehavior

local research CLIMATE

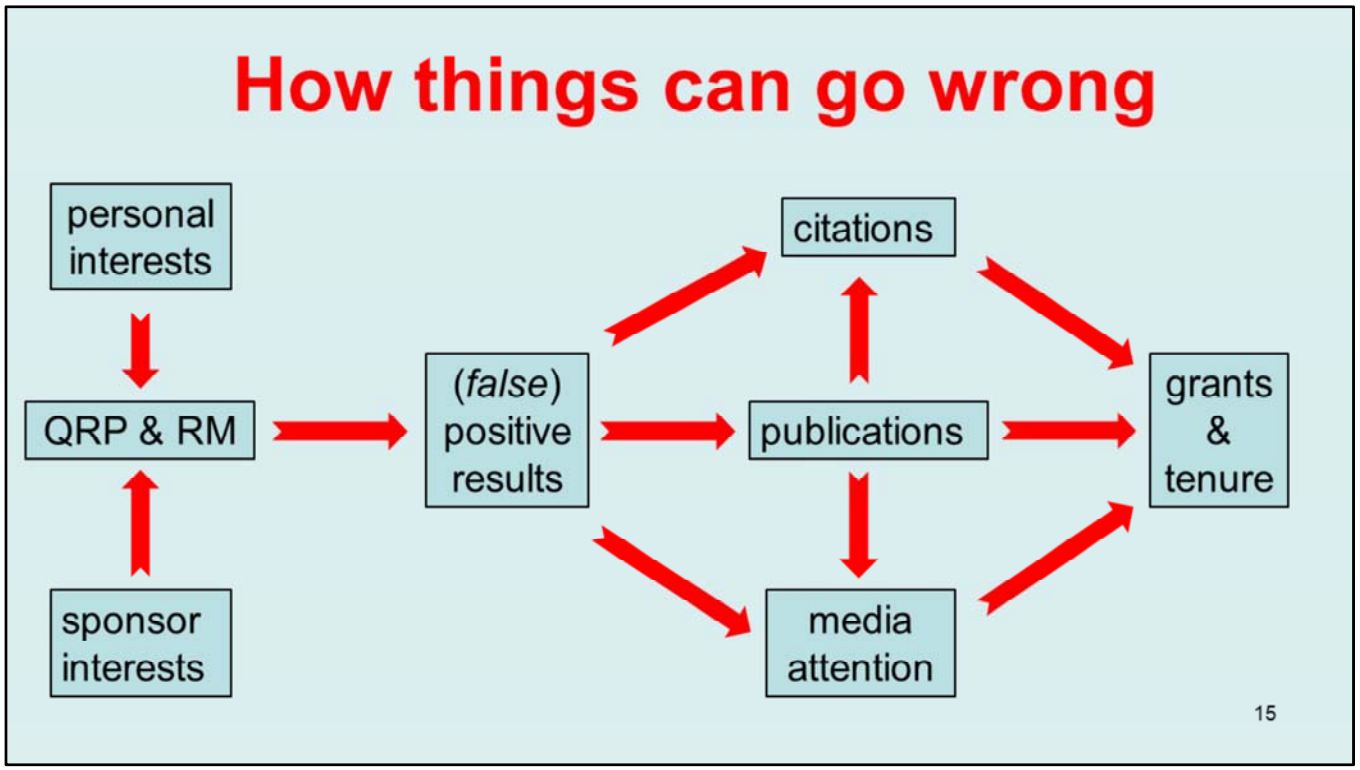
- Local norm adherence
- Local level of competition
- Received mentoring for survival (+) and RCR (-)
- Social support at work
- Local organizational injustice
- Local likelihood of detection

13

Content

- Research Integrity
- **Selective reporting and replication crisis**
- Plea for transparency
- What can institutions, journals and funders do?

14

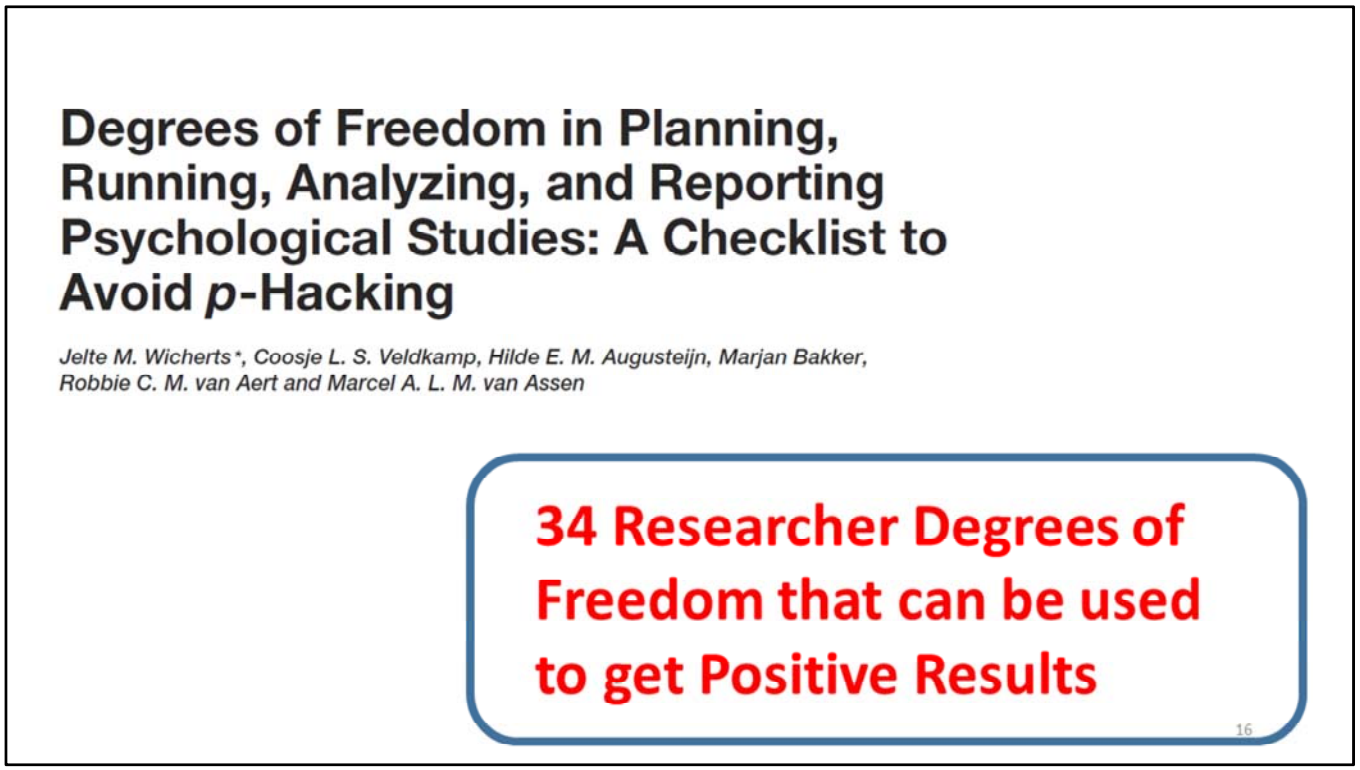


This slide shows – as a simplified summary of what has been explained – how things can go wrong.

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 likelyhood 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 truth in the published record and can explain the large replication difficulties some fields (e.g. preclinical research) experience.

There is strong 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.



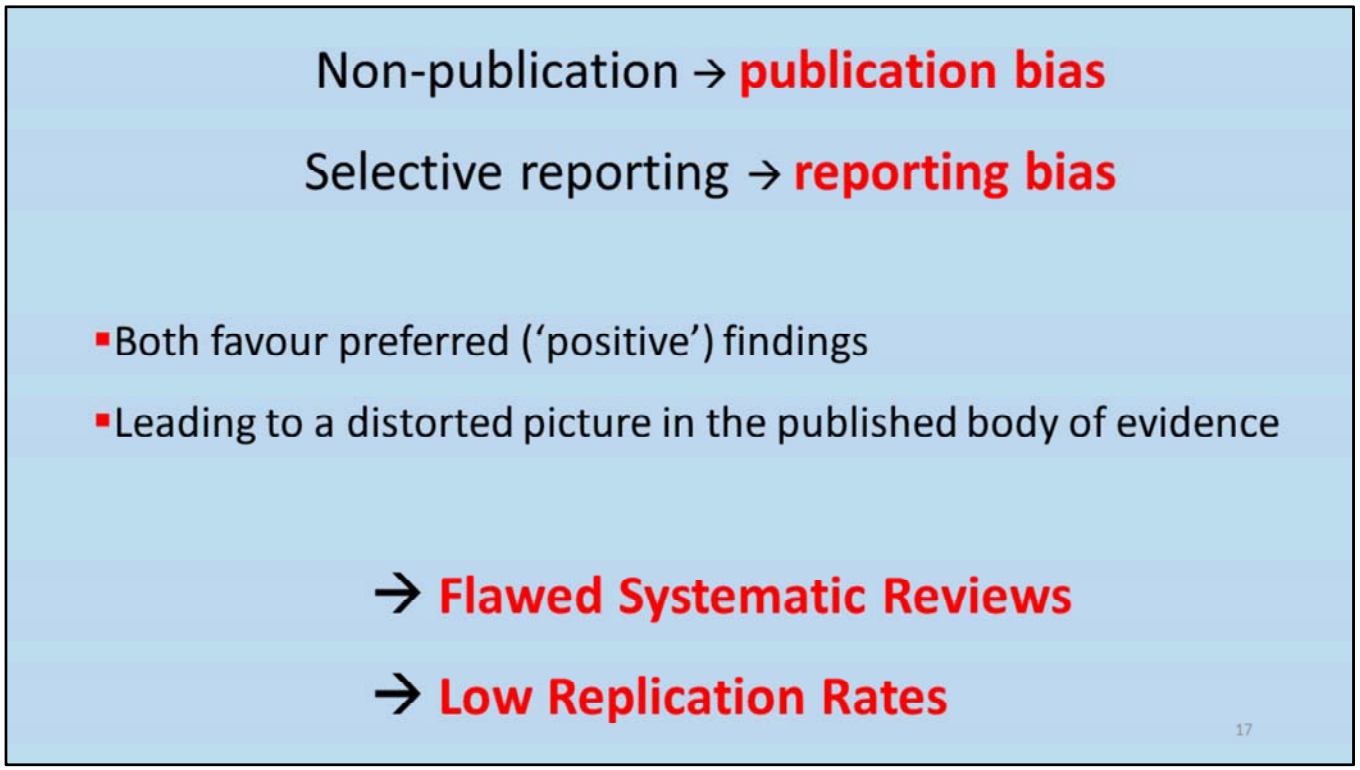
Wicherts et al - Degrees of freedom - checklist to avoid p-hacking - Front Psych 2016; 7 1832

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



There are thus many incentives for selective reporting

Selective publication comes in two forms.

Non-publication and selective reporting of findings may be the single most important source of research waste.

And it is the Achilles' heel of systematic reviews, because these rely on the published reports of research projects.

van der Steen JT, van den Bogert CA, van Soest-Poortvliet MC, Farsani SF, Otten RHJ, ter Riet G, Bouter LM. Determinants of selective reporting: a taxonomy based on content analysis of a random selection of the literature. PLoS ONE 2018: 13: e188247.
van den Bogert CA, Souverein PC, Brekelmans CTM, Janssen SWJ, Koëter GH, Leufkens HGM, Bouter LM. Primary endpoint discrepancies were found in one in ten clinical drug trials: results of an inception cohort study. Journal of Clinical Epidemiology 2017: 89: 199-208.
van den Bogert CA, Souverein PC, Brekelmans CTM, Janssen SWJ, Koëter GH, Leufkens HGM, Bouter LM. Non-publication is common amomg phase 1, single-center, not prospectively registered, or early terminated clinical drug trials: results of a nationwide inception cohort study in the Netherlands. PLoS ONE 2016; 11: e0167709.

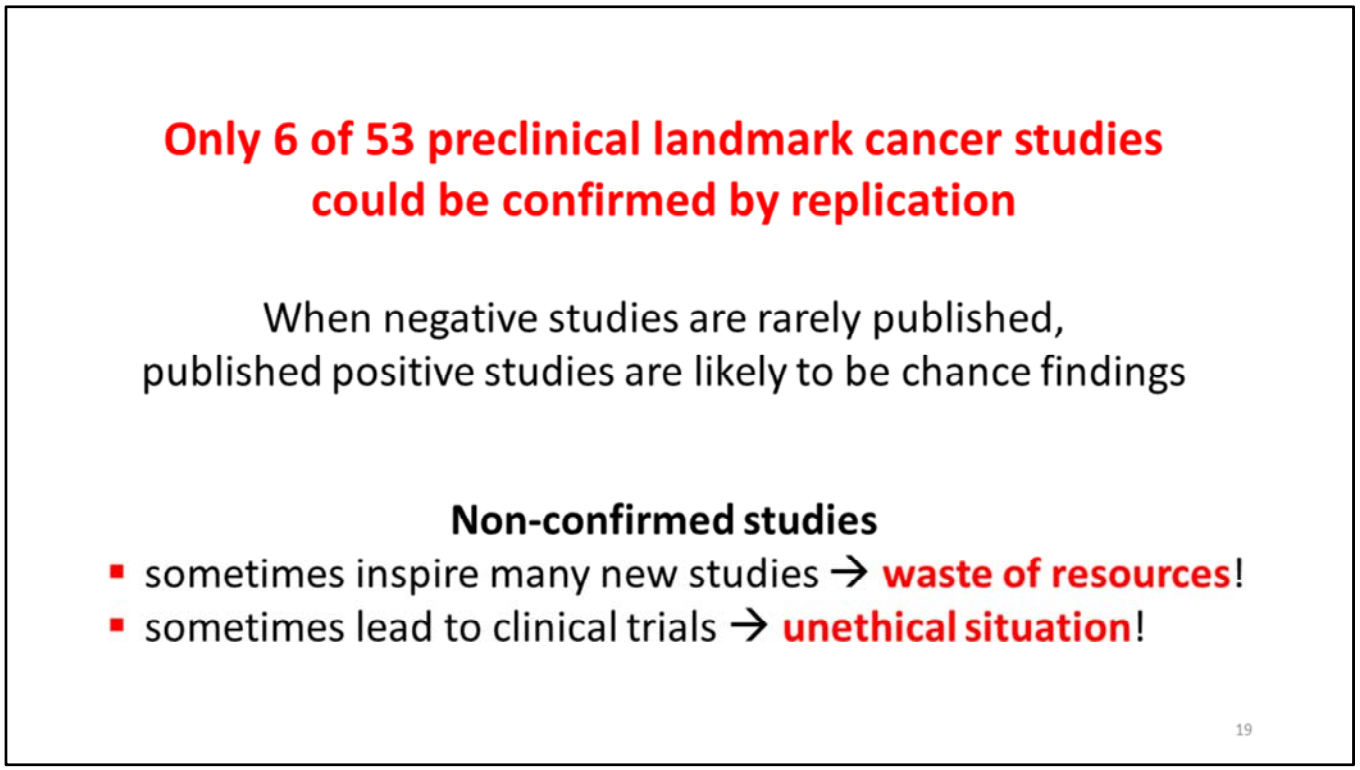
Let me just give one example, although it's a quite spectacular one.

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



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

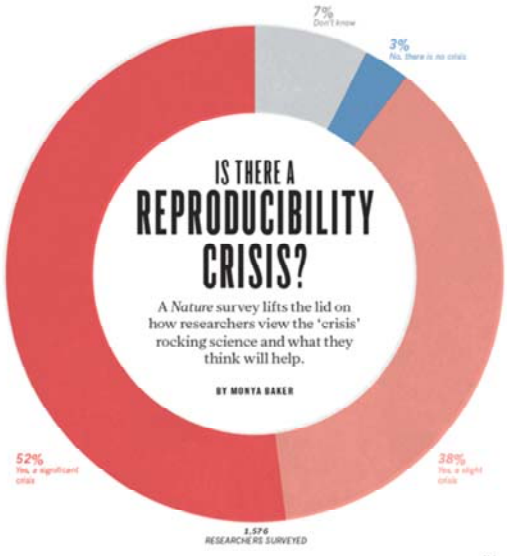
Two videos of John Ioannidis lecturing about reproducibility:

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

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.

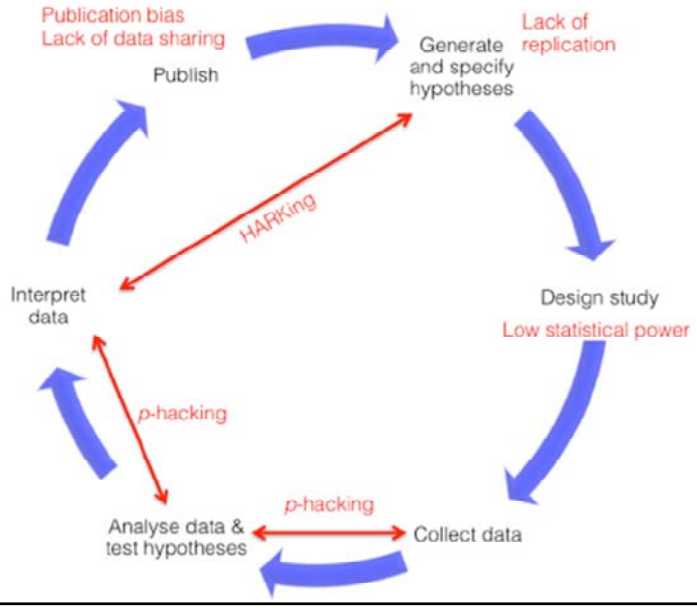


Just two recent Nature headlines. The topic draws attention, and rightly so.

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

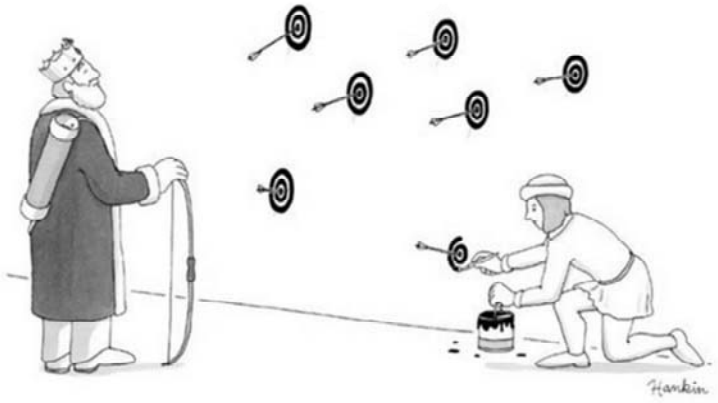
Important causes of 'replicability crisis'

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

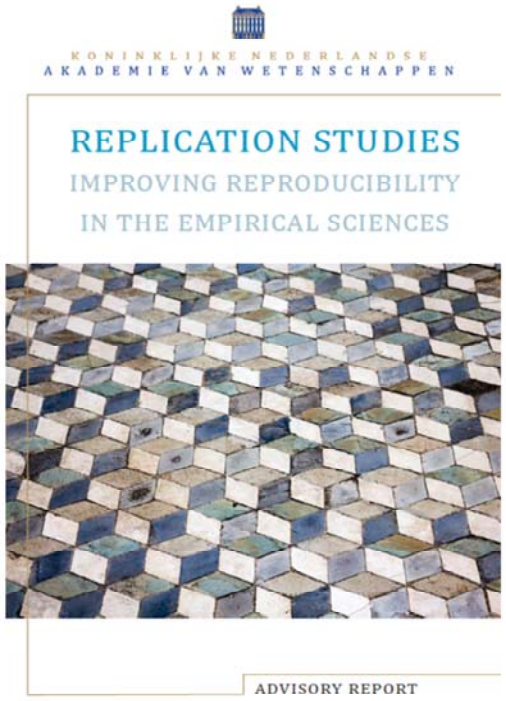


Wicherts et al - Degrees of freedom - checklist to avoid p-hacking - Front Psych 2016; 7 1832
Nosek et al - The preregistration revolution - PNAS 2018; 115 2600-6
Bouter - Fostering responsible research practices is a shared responsibility of multiple stakeholders - J Clin Epidemiol 2018; 93 143-6

Hypothesizing After the Results are Known (HARKing)



22



1. INTRODUCTION 16

2. CONCERNS ABOUT AND STRATEGIES TO IMPROVE REPRODUCIBILITY 18
2.1 Defining reproducibility and replication studies 18
2.2 Determining the reproducibility of results 19
2.3 Occurrence and potential impact of non-reproducibility 21
2.4 Causes of non-reproducibility 23
2.5 Strategies to improve reproducibility 25

3. THE DESIRABILITY OF REPLICATION STUDIES 27
3.1 Goals of replication studies 27
3.2 Kinds of replication studies 29
3.3 Criteria for when to conduct replication studies 30

4. REPLICATION STUDIES IN PRACTICE 32
4.1 Data on replication studies 32
4.2 Reproducibility and replication practices in preclinical animal research 33
4.3 Reproducibility and replication practices in clinical research 34
4.4 Reproducibility and replication practices in empirical psychology 36
4.5 Reproducibility and replication practices in genetic epidemiology 37
4.6 Reproducibility and replication practices in biochemistry 38
4.7 Lessons from replication practices in different disciplines 38

5. BARRIERS TO AND STRATEGIES FOR CONDUCTING MORE REPLICATION STUDIES 40
5.1 Information-sharing about original and replication studies 40
5.2 Knowing when and how to perform replication studies 42
5.3 Incentives for replication studies 43

6. CONCLUSIONS AND RECOMMENDATIONS 47
6.1 Improve reproducibility 47
6.2 Conduct more replication studies 48

See: https://www.nrin.nl/wp-content/uploads/KNAW-Replication-Studies-15-01-2018.pdf


Every baby knows the
scientific method!
1 Make an observation.
2 Form a hypothesis.
3 Perform the experiment.
4 Analyze the data.
5 Report your findings.
6 Invite others to reproduce the results.

"Only when certain events recur in accordance with rules or regularities, as in the case of **repeatable** experiments, can our observations be tested—in principle—by anyone.... Only by such **repetition** can we convince ourselves **that we are not dealing with a mere isolated 'coincidence**,' but with events which, on account of their regularity and **reproducibility**, are in principle inter-subjectively testable."

Karl Popper. The Logic of Scientific Discovery. London: Hutchison. 1959, P. 45

25

Forms of Replication

- **Re-analysis** of the **same data** set (with same or alternative data-analysis plan)
- **Direct replication** (methods reproducibility): collect (and analyse) **new data** with the same study protocol
- **Conceptual replication** (external validity, triangulation): collect (and analyse) new data with an **alternative study protocol** for the same study objective

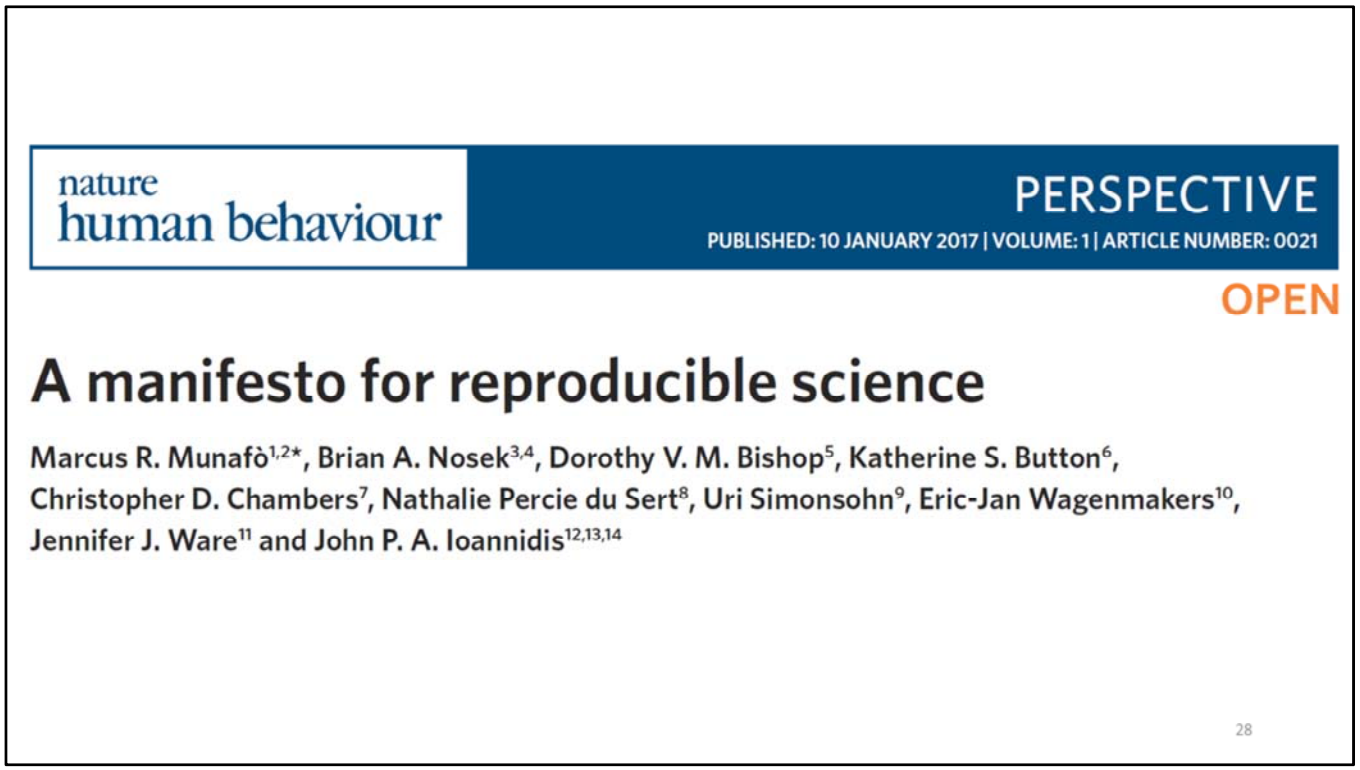
26

Goodman, Fanelli, Ioannidis – What does reproducibility mean? – Science Translational Medicine 2016; 8 341 ps12

Repeating experiments is not enough

27

Munafò and Davey Smith - Repeating experiments is not enough - Nature 2018; 553 399-401



http://www.nature.com/articles/s41562-016-0021

Table 1 | A manifesto for reproducible science.

Theme	Proposal	Examples of initiatives/potential solutions (extent of current adoption)	Stakeholder(s)
Methods	Protecting against cognitive biases	All of the initiatives listed below (* to ****) Blinding (**)	J, F
	Improving methodological training	Rigorous training in statistics and research methods for future researchers (*) Rigorous continuing education in statistics and methods for researchers (*)	I, F
	Independent methodological support	Involvement of methodologists in research (**) Independent oversight (*)	F
	Collaboration and team science	Multi-site studies/distributed data collection (*) Team-science consortia (*)	I, F
Reporting and dissemination	Promoting study pre-registration	Registered Reports (*) Open Science Framework (*)	J, F
	Improving the quality of reporting	Use of reporting checklists (**) Protocol checklists (*)	J
	Protecting against conflicts of interest	Disclosure of conflicts of interest (***) Exclusion/containment of financial and non-financial conflicts of interest (*)	J

29

Table 1 | A manifesto for reproducible science.

Theme	Proposal	Examples of initiatives/potential solutions (extent of current adoption)	Stakeholder(s)
Reproducibility	Encouraging transparency and open science	Open data, materials, software and so on (* to **) Pre-registration (**** for clinical trials, * for other studies)	J, F, R
Evaluation	Diversifying peer review	Preprints (* in biomedical/behavioural sciences, **** in physical sciences) Pre- and post-publication peer review, for example, Publons, PubMed Commons (*)	J
Incentives	Rewarding open and reproducible practices	Badges (*) Registered Reports (*) Transparency and Openness Promotion guidelines (*) Funding replication studies (*) Open science practices in hiring and promotion (*)	J, I, F

Estimated extent of current adoption: *, <5%; **, 5–30%; ***, 30–60%; ****, >60%. Abbreviations for key stakeholders: J, journals/publishers; F, funders; I, institutions; R, regulators.

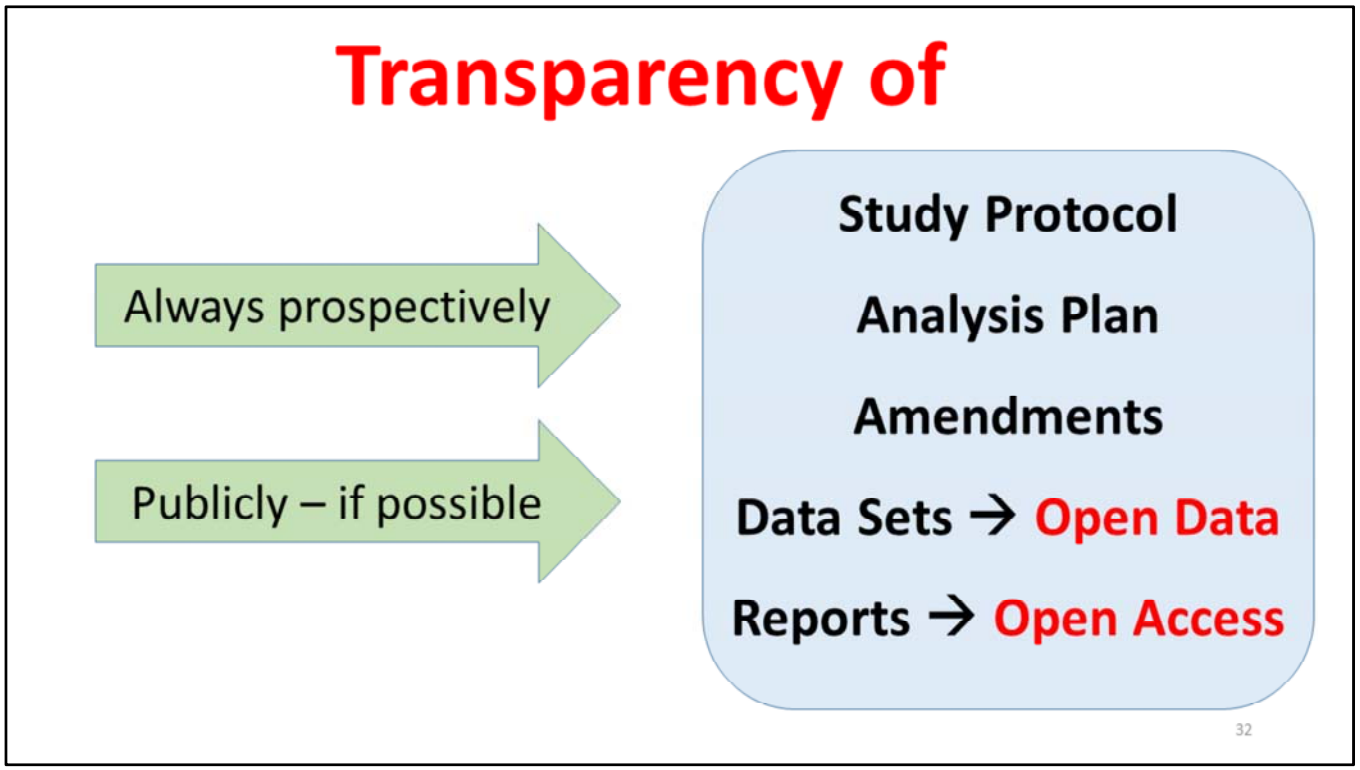
30

Content

- Research Integrity
- Selective reporting and replication crisis
- **Plea for transparency**
- What can institutions, journals and funders do?

31

.



The current low levels of reproducibility are wasteful in the sense that resources were wasted on the production of these false leads in the scientific literature. It's also unethical when animals or humans have been burdened for unpublished studies or for published false positive findings.

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.

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

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

Preregistration

- Essential for hypothesis testing research (**PREDICTION**)
- Alias context of justification, confirmatory research
- Optional for hypothesis-generating research (**POSTDICTION**)
- Alias context of discovery, exploratory research

- p-values only interpretable for **PREDICTION** + preregistration
- **POSTDICTION** p-values likely due to HARKing and p-hacking
- In other words: due to hindsight bias or data-driven

34

https://osf.io/

https://clinicaltrials.gov/

https://dataverse.nl/

https://figshare.com/

https://www.mendeley.com/

https://www.preclinicaltrials.eu/

Content

- Research Integrity
- Selective reporting and replication crisis
- Plea for transparency
- **What can institutions, journals and funders do?**

My wish list - institutions

1. **clear codes, guidelines and SOPs**
   what is expected behaviour in operational terms
2. **fair procedures for handling allegations**
   protect both the whistleblowers and the scientists they accuse
3. **adequate mentoring and training in RCR**
   likely to be important, not only for PhD students
4. **adequate methodological and statistical support**
   many QRPs have to do with poor methods

37

Bouter LM, Hendrix S. Both whistle blowers and the scientists they accuse are vulnerable and deserve protection. Accountability in Research: Policies and Quality Assurance 2017; 24: 359-66.
Bouter LM. Fostering responsible research practices is a shared responsibility of multiple stakeholders. Journal of Clinical Epidemiology 2018; 96: 143-6.

My wish list - institutions

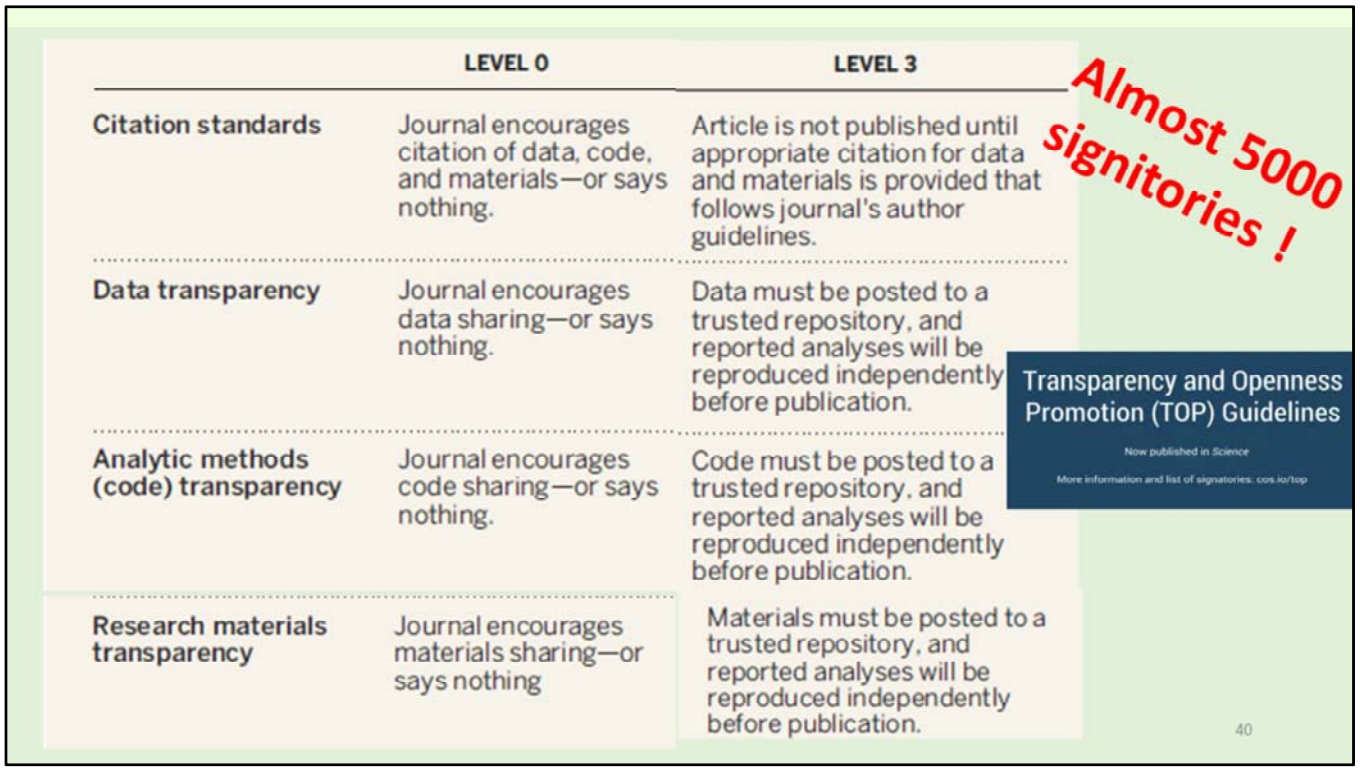
5. **system of internal audits**
   this is so often ignored in academia
6. **good facilities for data-management and –storage**
   web-based solutions for being transparant and accountable
7. **remove the perverse incentives in the reward system**
   not only counting publications and citations
8. **promote an open research climate**
   open discussion of dilemmas and learning from mistakes

38

My wish list - journals

1. **follow TOP guidelines**
   and be as transparant as possible
2. **introduce Registered Reports**
   to avoid selective reporting
3. **use reporting guidelines**
   to make publications more clear and informative
4. **promote preprints and postpublication peer review**
   with a view to enable optimal scholarly discussion

39

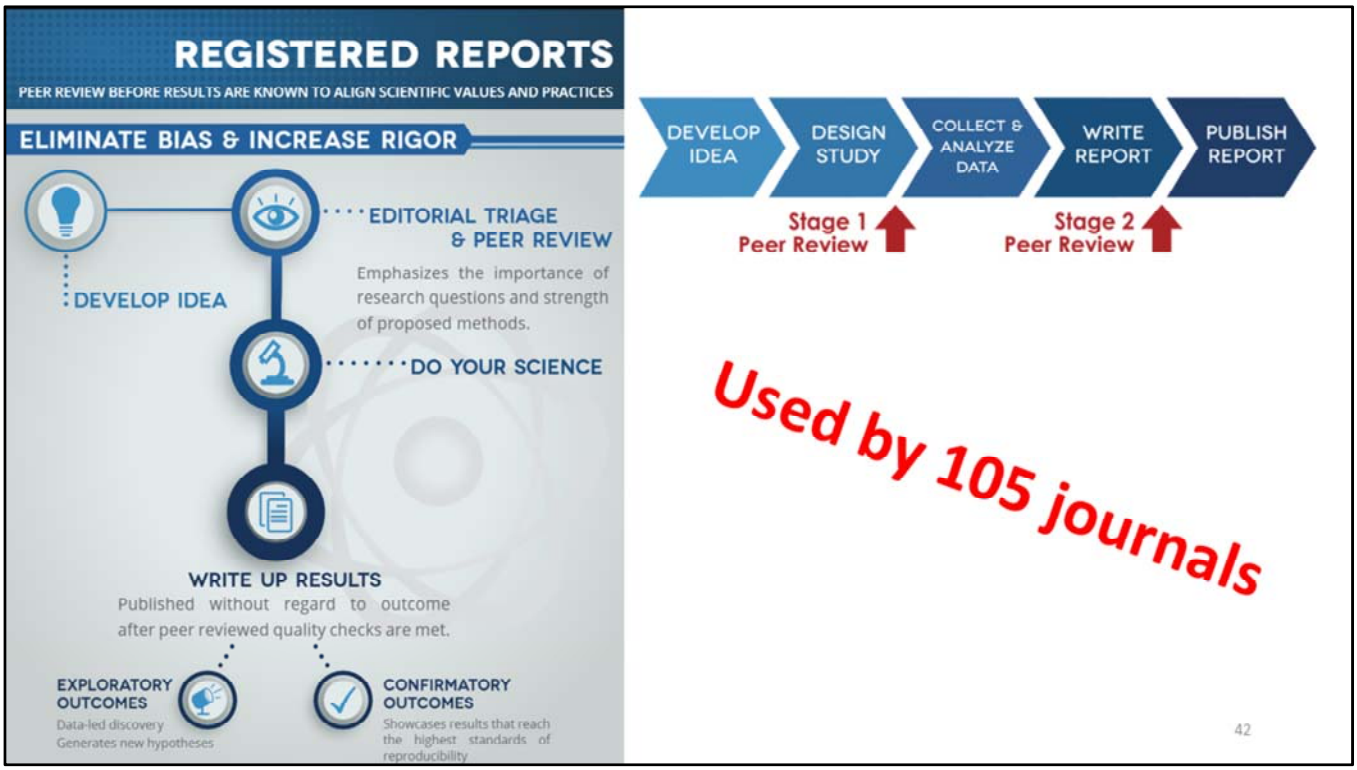


Nosek et al - Promoting an open research culture - Science 2015; 348 1422-5
https://cos.io/our-services/top-guidelines/

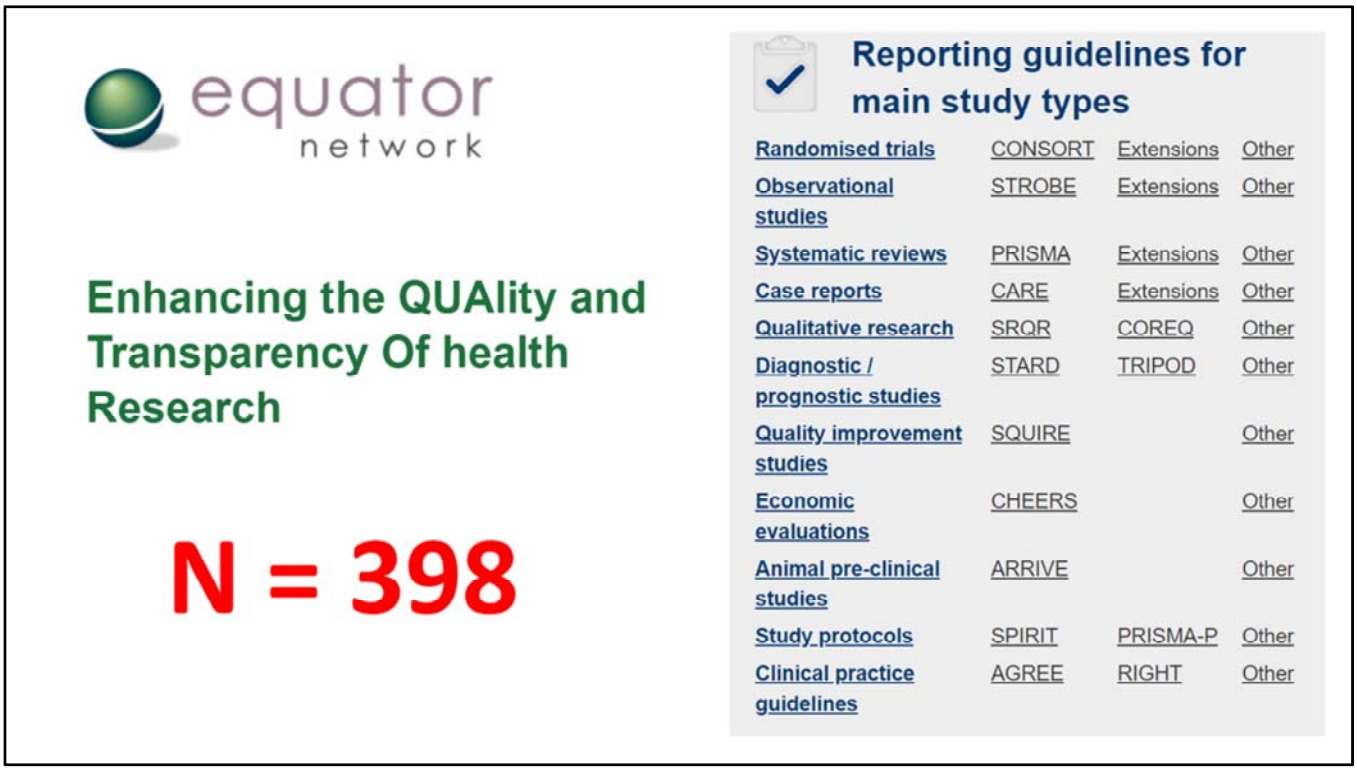
Design and analysis transparency	Journal encourages design and analysis transparency or says nothing.	Journal requires and enforces adherence to design transparency standards for review and publication.
Preregistration of studies	Journal says nothing.	Journal requires preregistration of studies and provides link and badge in article to meeting requirements.
Preregistration of analysis plans	Journal says nothing.	Journal requires preregistration of studies with analysis plans and provides link and badge in article to meeting requirements.
Replication	Journal discourages submission of replication studies—or says nothing.	Journal uses Registered Reports as a submission option for replication studies with peer review before observing the study outcomes.

41

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



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



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

My wish list - funders

1. **check eligibility of institutions**
   fulfilling the essential 'duties of care'
2. **check scientific or societal relevance**
   and not yet sufficiently investigated (systematic review)
3. **demand transparency**
   preregistration, protocol, data-analysis plan, data sets
4. **prescribe Open Data and Open Access**
   unless there are good reasons against this

44


6th WCRI 2019
6th WORLD CONFERENCE ON RESEARCH INTEGRITY
HONG KONG
2 - 5 JUNE 2019
www.wcri2019.org


WCRI
WORLD CONFERENCES
ON RESEARCH INTEGRITY
www.wcrif.org