fMRI in context: Paradigm to practice

Dr Steffen Bollmann

Postdoctoral Researcher and National Imaging Facility Fellow
Centre for Advanced Imaging, University of Queensland, Australia

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

OR: The perish of “Publish or Perish”

A short overview about

- What’s the problem?
  - the problems of p-values
  - the reproducibility crisis
  - questionable research practices

- Possible Solutions
  - Understanding the consequences of claiming $p < 0.05$
  - Incentives on replication
  - Preregistration
  - Blind analysis
  - Let us ignore “impact factors”
  - peer review 2.0 and new forms of publishing

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

OR: The perish of “Publish or Perish”

A short overview about

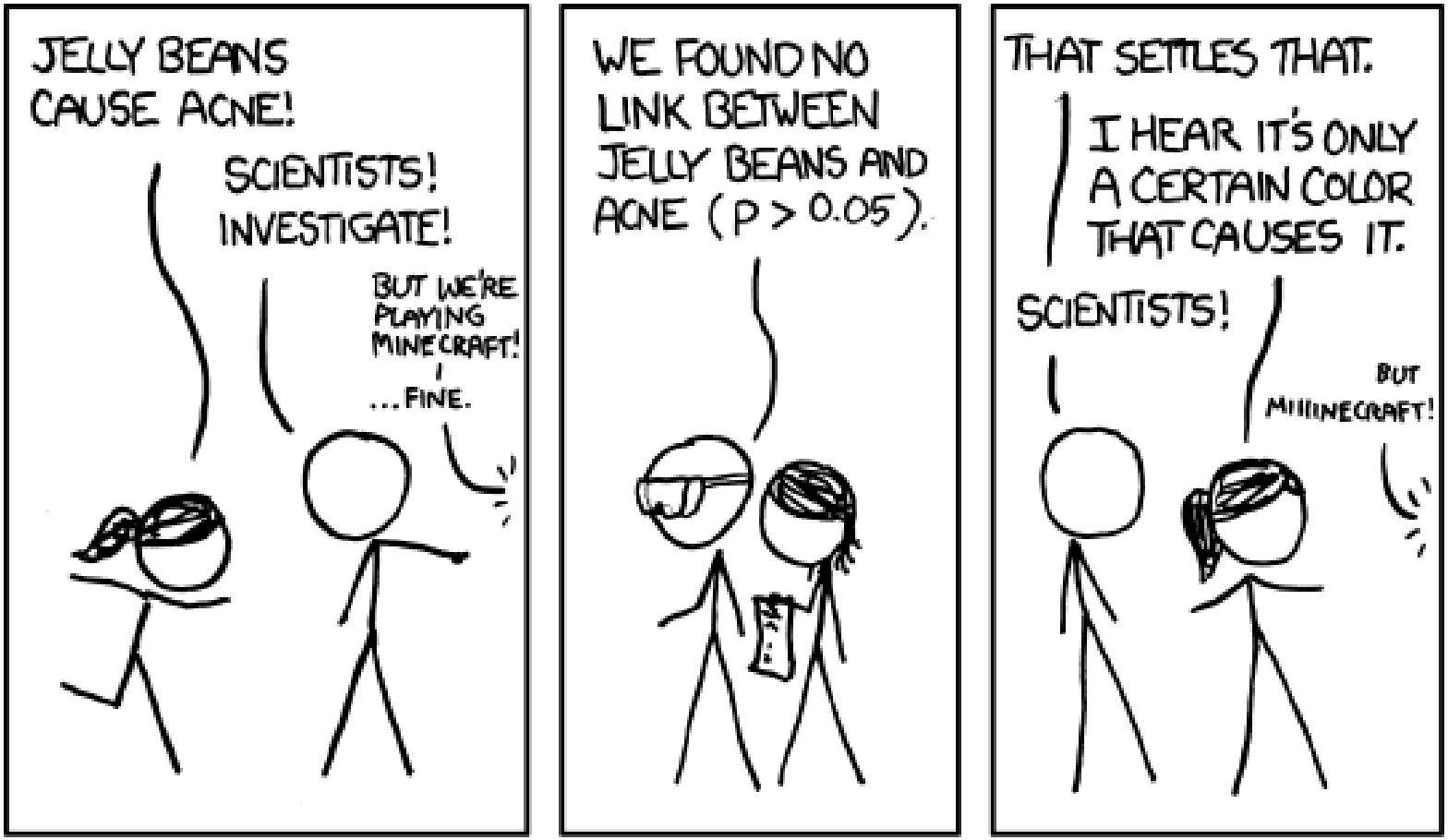
- What’s the problem?
  - **the problems of p-values**
  - the reproducibility crisis
  - questionable research practices

- Possible Solutions
  - Understanding the consequences of claiming $p < 0.05$
  - Incentives on replication
  - Preregistration
  - Blind analysis
  - Let us ignore “impact factors”
  - peer review 2.0 and new forms of publishing

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

Significant ($p < 0.05$)

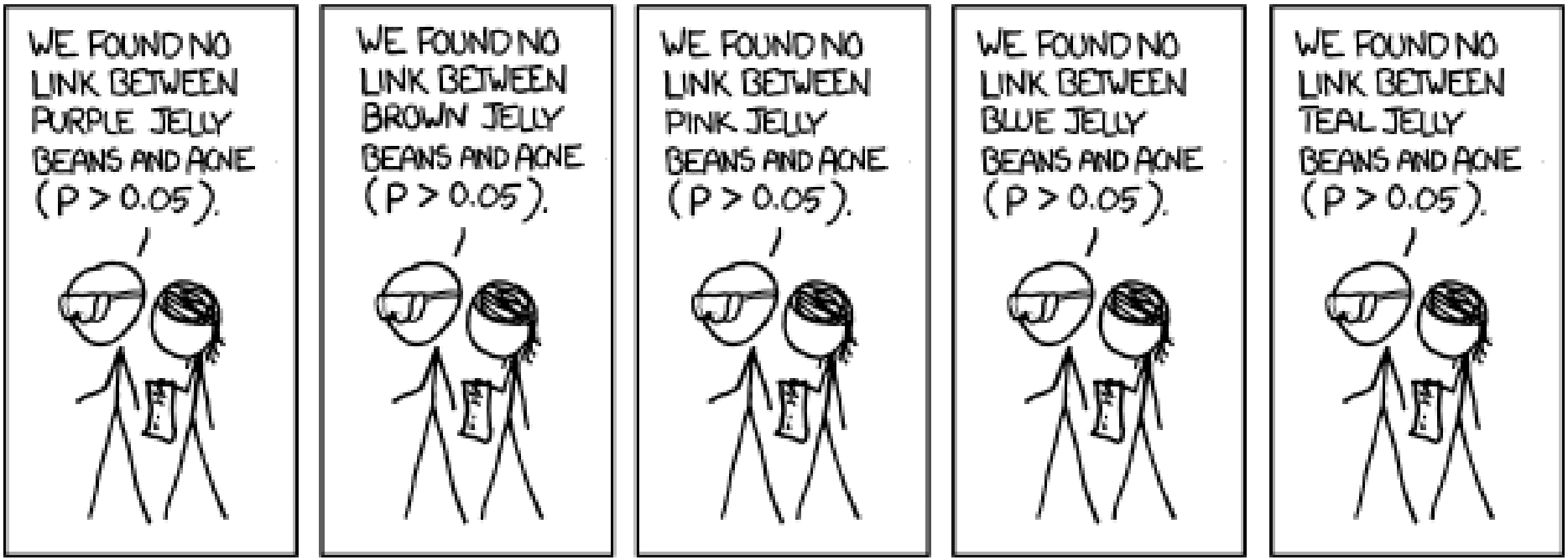


https://xkcd.com/882/

CRICOS Provider No 00025B

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
Create change

Significant ($p < 0.05$)

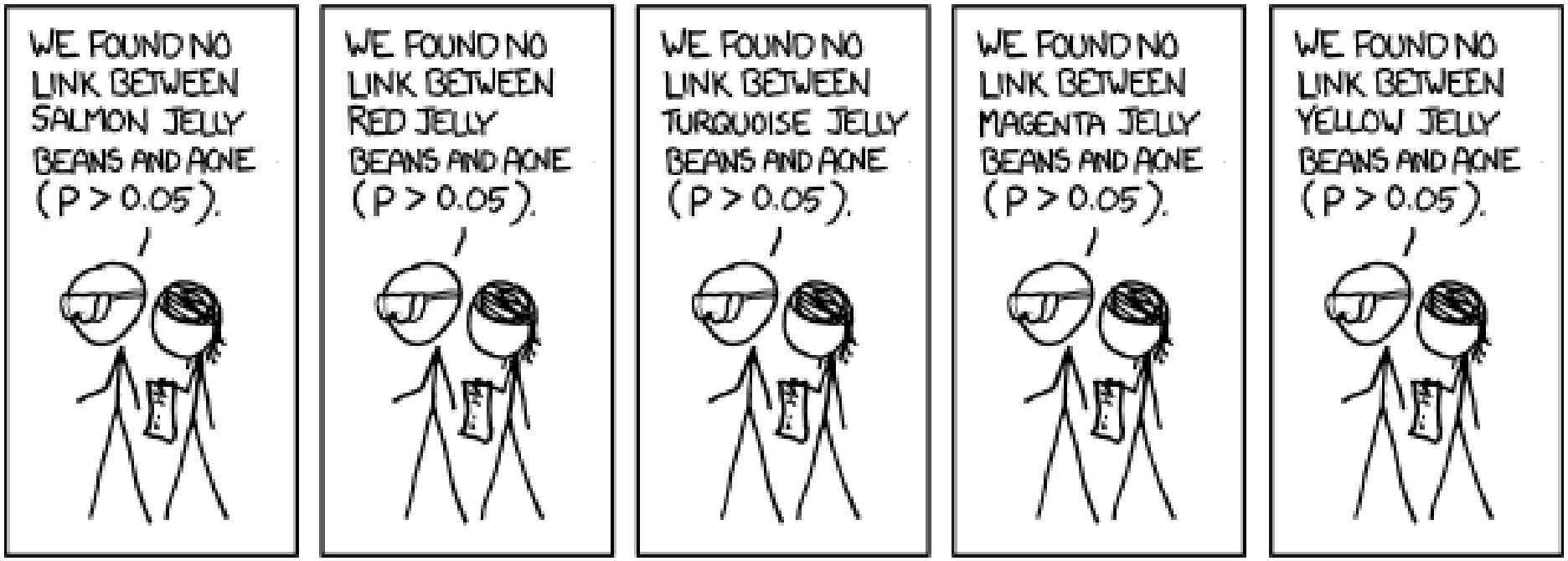


https://xkcd.com/882/

CRICOS Provider No 00025B

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
Create change

Significant ($p < 0.05$)

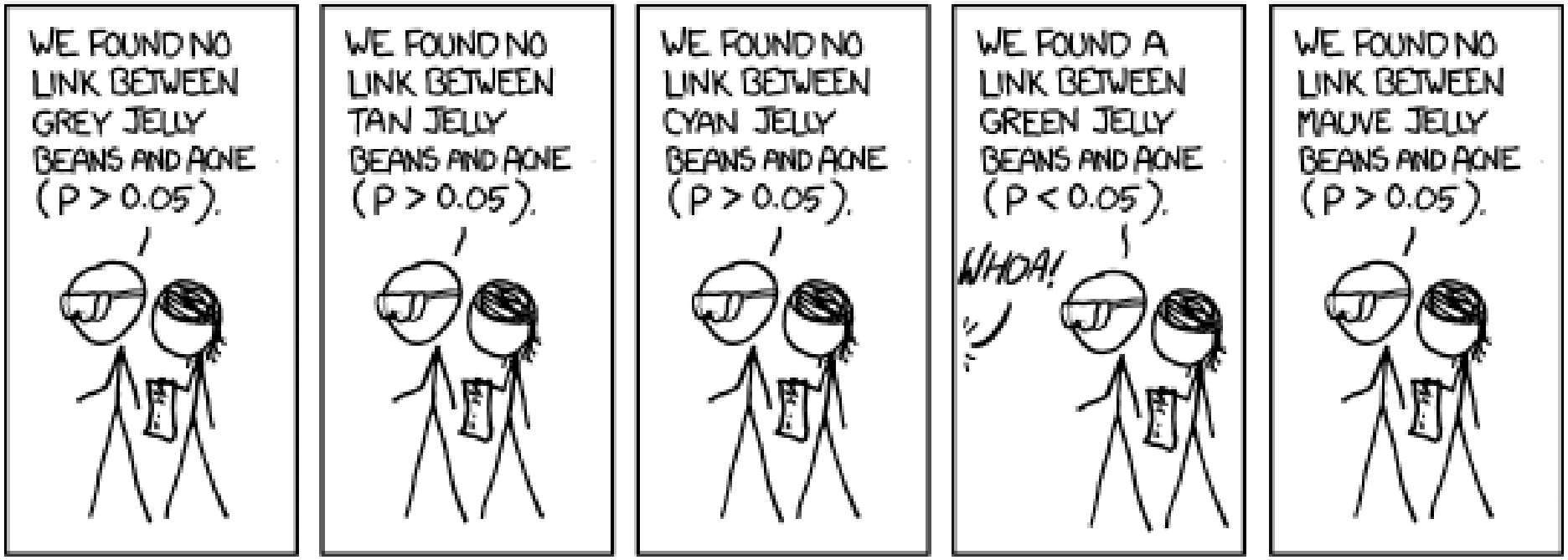


https://xkcd.com/882/

CRICOS Provider No 00025B

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
Create change

Significant ($p < 0.05$)

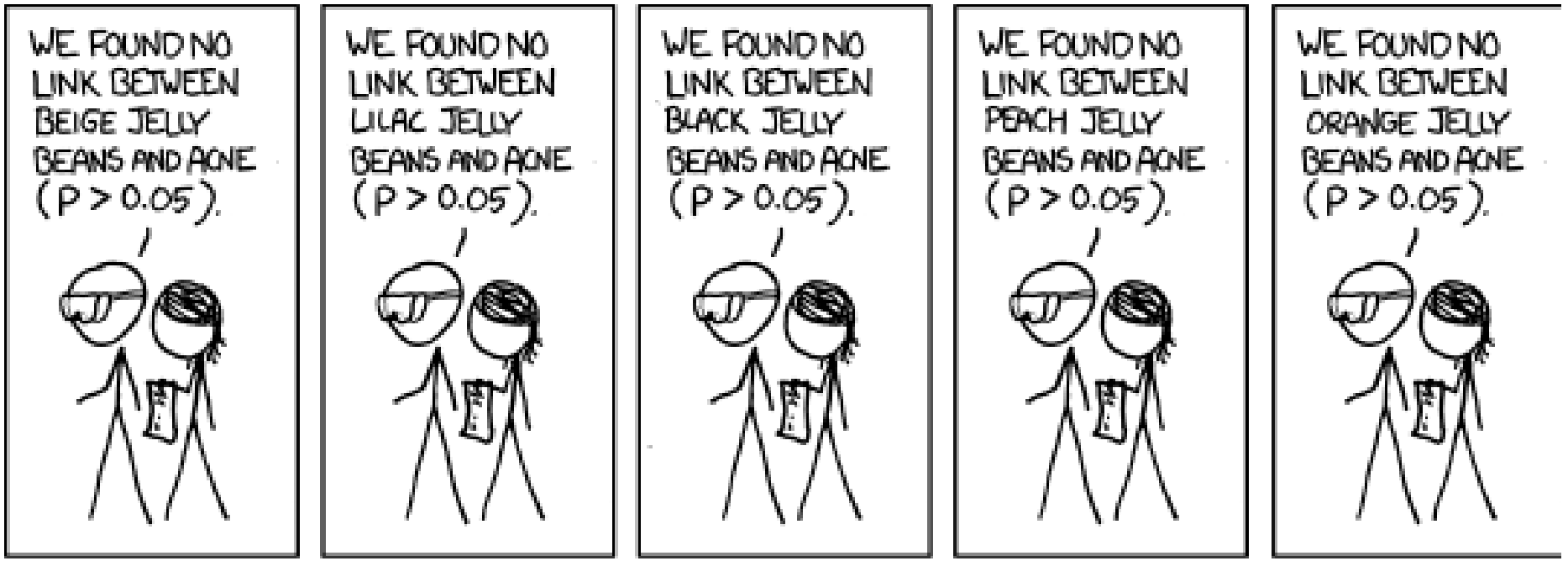


https://xkcd.com/882/

CRICOS Provider No 00025B

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
Create change

Significant ($p < 0.05$)

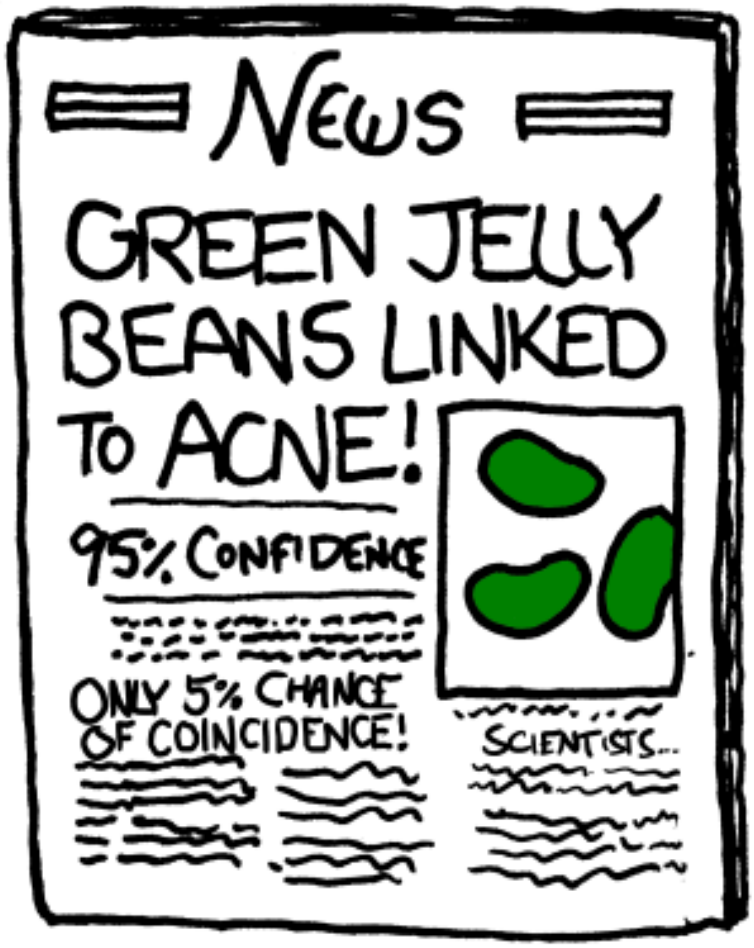


https://xkcd.com/882/

CRICOS Provider No 00025B

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
Create change

Significant ($p < 0.05$)



https://xkcd.com/882/

CRICOS Provider No 00025B

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
Create change

Significant ($p < 0.05$)

P-value-hacking, data dredging, torturing the data until it confesses, data massaging, not conducting proper multiple comparison correction, researcher degrees of freedom …. It has many names … and leads to real world problems

- Are we really p-hacking?
- How are we p-hacking (most of us unintentionally)?
- Why is it a problem?
- How can we solve the problem?

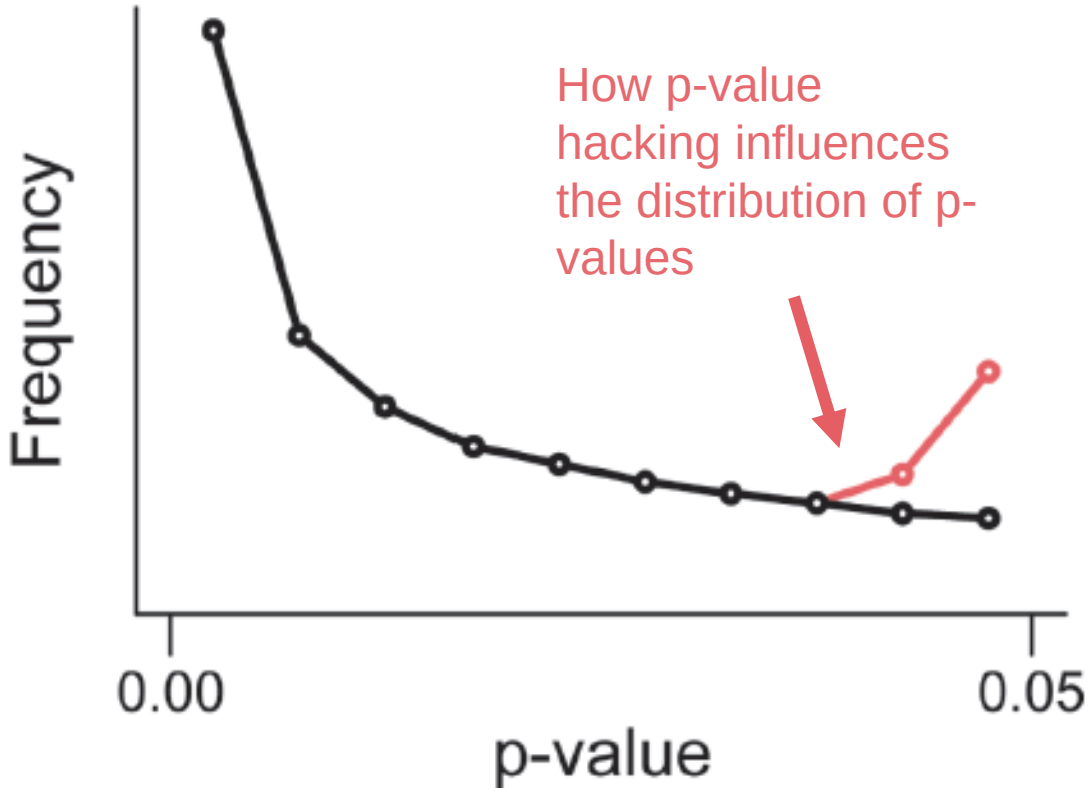
CRICOS Provider No 00025B

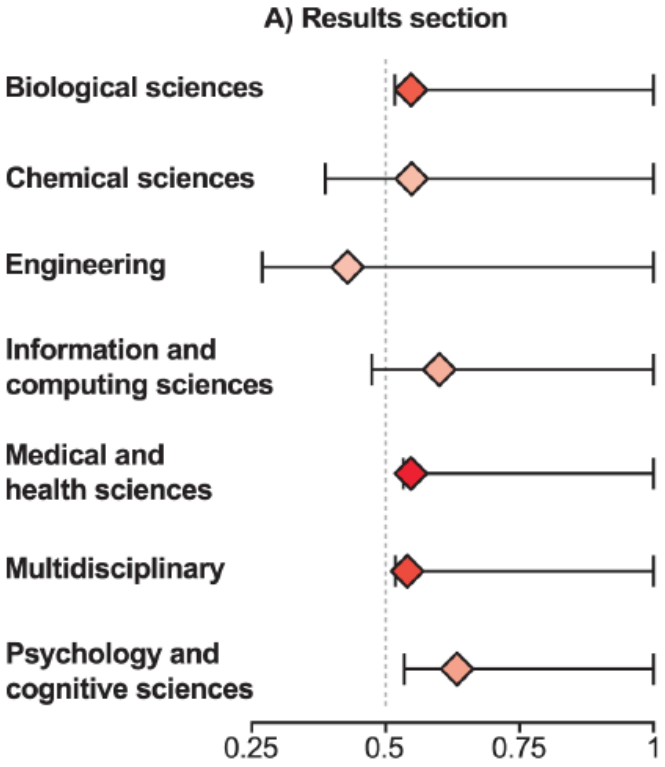
The University Of Queensland
AUSTRALIA
Create change

Are we really p-hacking?

Head et al. 'The Extent and Consequences of P-Hacking in Science'. *PLoS Biol* 2015.

- strong evidence for p-hacking in the literature





CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

Are we really p-hacking?

Head et al. 'The Extent and Consequences of P-Hacking in Science'. *PLoS Biol* 2015.

- Conclusion: evidence of p-hacking, but effect weak in relation to real effect sizes, not likely to cause distortions in the literature

Head et al. study criticized by Dorothy V Bishop et al: https://peerj.com/preprints/1266/

- no control over the type of p-values entered into the analysis
- lack of 'bump' is not indicative of lack of p-hacking
- study by Head et al. provides evidence of p-hacking, but cannot be used to estimate extent and consequences of p-hacking!
- **-> Bishop's simulation shows instead: potential for systematic bias is substantial!**

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

How are we p-hacking?

Analytic flexibility

- We can analyse our data in multiple justifiable ways, but only report the way “that worked” (i.e.: $p < 0.05$)
- How to p-hack by neuroskeptic: https://youtu.be/A0vEGuOMTyA

Selection/publication bias

- we (and journals) prefer the lowest p-value given a choice and often do not report higher p-values (or do not publish non-significant results)

Selective debugging

- We are more likely to look for mistakes in an analysis which created non-significant results -> we are selecting in favour of mistakes that produce false positives

[http://blogs.discovermagazine.com/neuroskeptic/2015/05/18/p-hacking-a-talk-and-further-thoughts/]

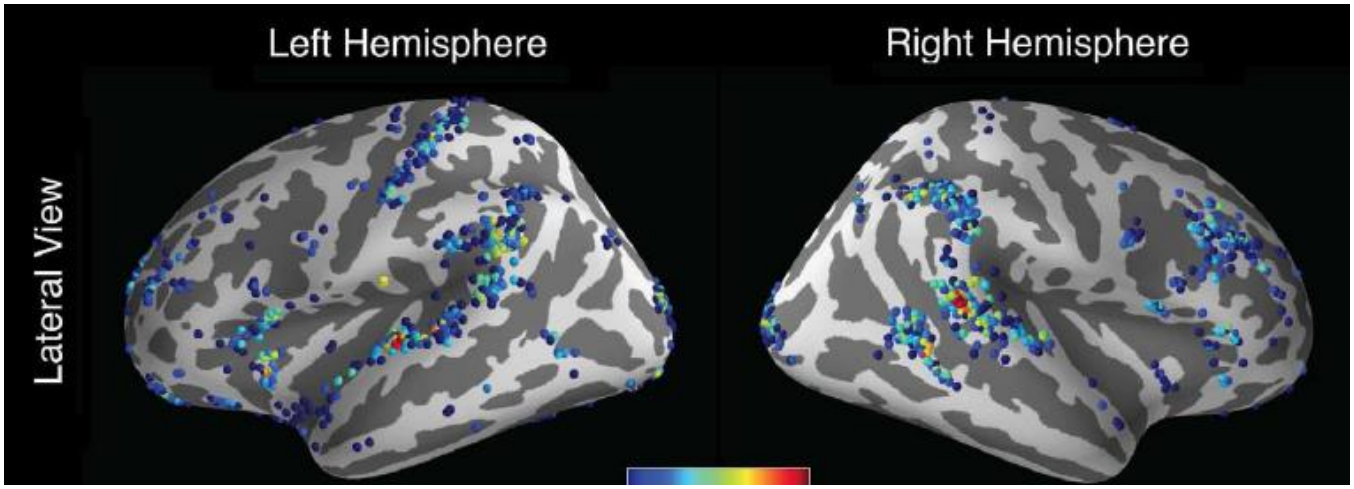
CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

Analytic flexibility in fMRI

Carp, Joshua. 'On the Plurality of (Methodological) Worlds: Estimating the Analytic Flexibility of fMRI Experiments'. *Frontiers in Neuroscience* 2012:

- Potential for false positives increases with analysis flexibility
- This study tested 34560 different ways of analysing one fMRI dataset
- Some outcomes were consistent across pipelines, others showed methods-related variability in activation strength, location, and extent



Colour indicates number of pipelines yielding activation for that location

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

OR: The perish of “Publish or Perish”

A short overview about

- What’s the problem?
  - the problems of p-values
  - **the reproducibility crisis**
  - questionable research practices

- Possible Solutions
  - Understanding the consequences of claiming $p < 0.05$
  - Incentives on replication
  - Preregistration
  - Blind analysis
  - Let us ignore “impact factors”
  - peer review 2.0 and new forms of publishing

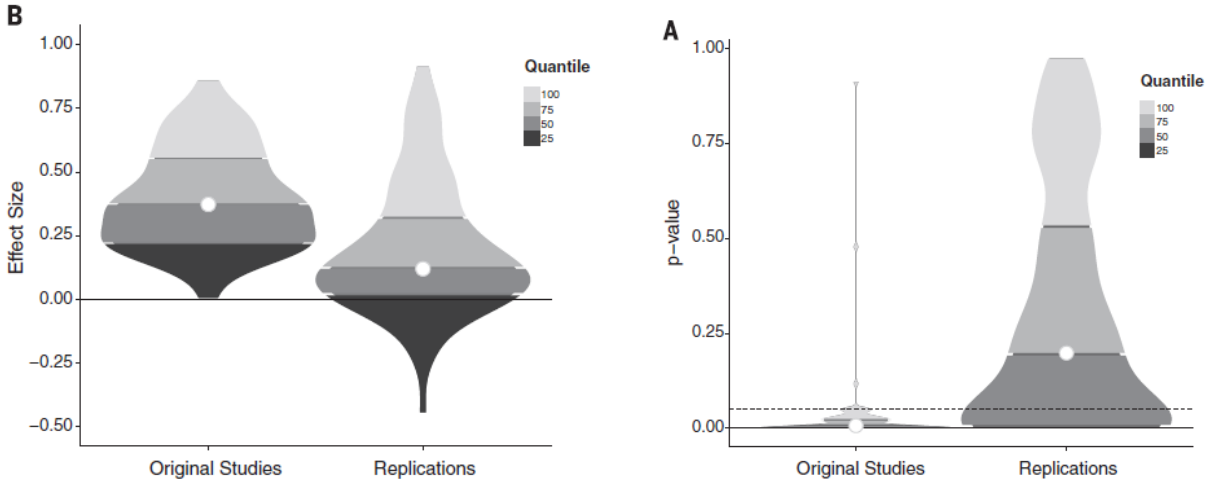
CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

Why is it a problem?

The replication/reproducibility "crisis"

- a large-scale (**100 studies**), collaborative effort to reproduce psychological science studies published in 3 journals was conducted
- The mean effect size of the replicated effects was **half the effect size** of the original effects
- **97%** of original studies had significant results ($p < .05$), but only **36%** of the replications had significant results



[Open Science Collaboration. 'Estimating the Reproducibility of Psychological Science'. *Science* 2015]

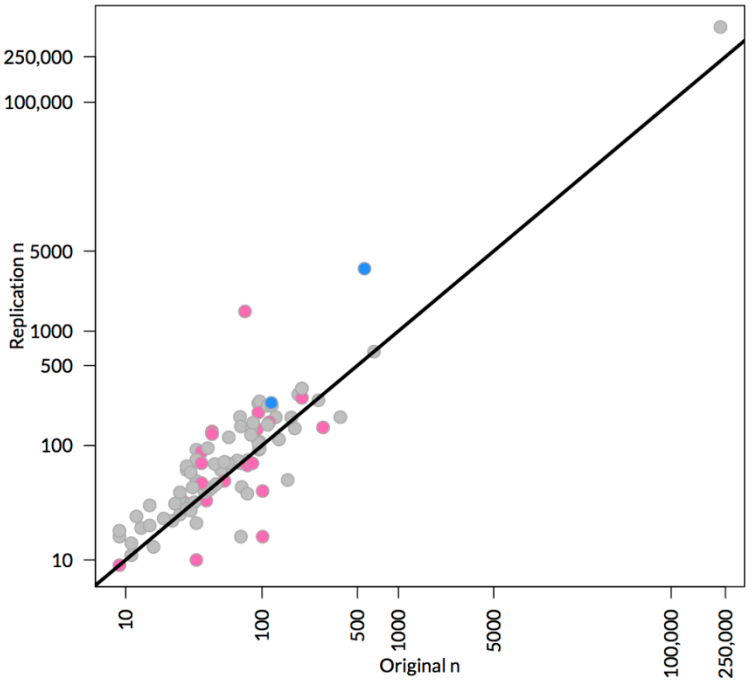
CRICOS Provider No 00025B

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
Create change

Why is it a problem?

The replication/reproducibility "crisis"

- only 36% of the studies were reproduced – this sounds very bad ...
- but: many of the studies being replicated had a small sample size, which is already a problem:
  - small sample size studies are almost impossible to replicate in the first place!
  - **The smaller the sample and effect size, the less likely the research findings are to be true.**



http://simplystatistics.org/?p=4336

Ioannidis, John P. A. 'Why Most Published Research Findings Are False'. *PLoS Med* 2005

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

OR: The perish of “Publish or Perish”

A short overview about

- What’s the problem?
  - the problems of p-values
  - the reproducibility crisis
  - **questionable research practices**

- Possible Solutions
  - Understanding the consequences of claiming $p < 0.05$
  - Incentives on replication
  - Preregistration
  - Blind analysis
  - Let us ignore “impact factors”
  - peer review 2.0 and new forms of publishing

CRICOS Provider No 00025B

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
Create change

Questionable Research Practices

- include p-hacking, HARKing (hypothesizing after results are known), lack of replication, publication bias (file drawer problem), low statistical power, lack of data sharing, selective reporting (cherry picking), selective stopping, …

why are we doing that?

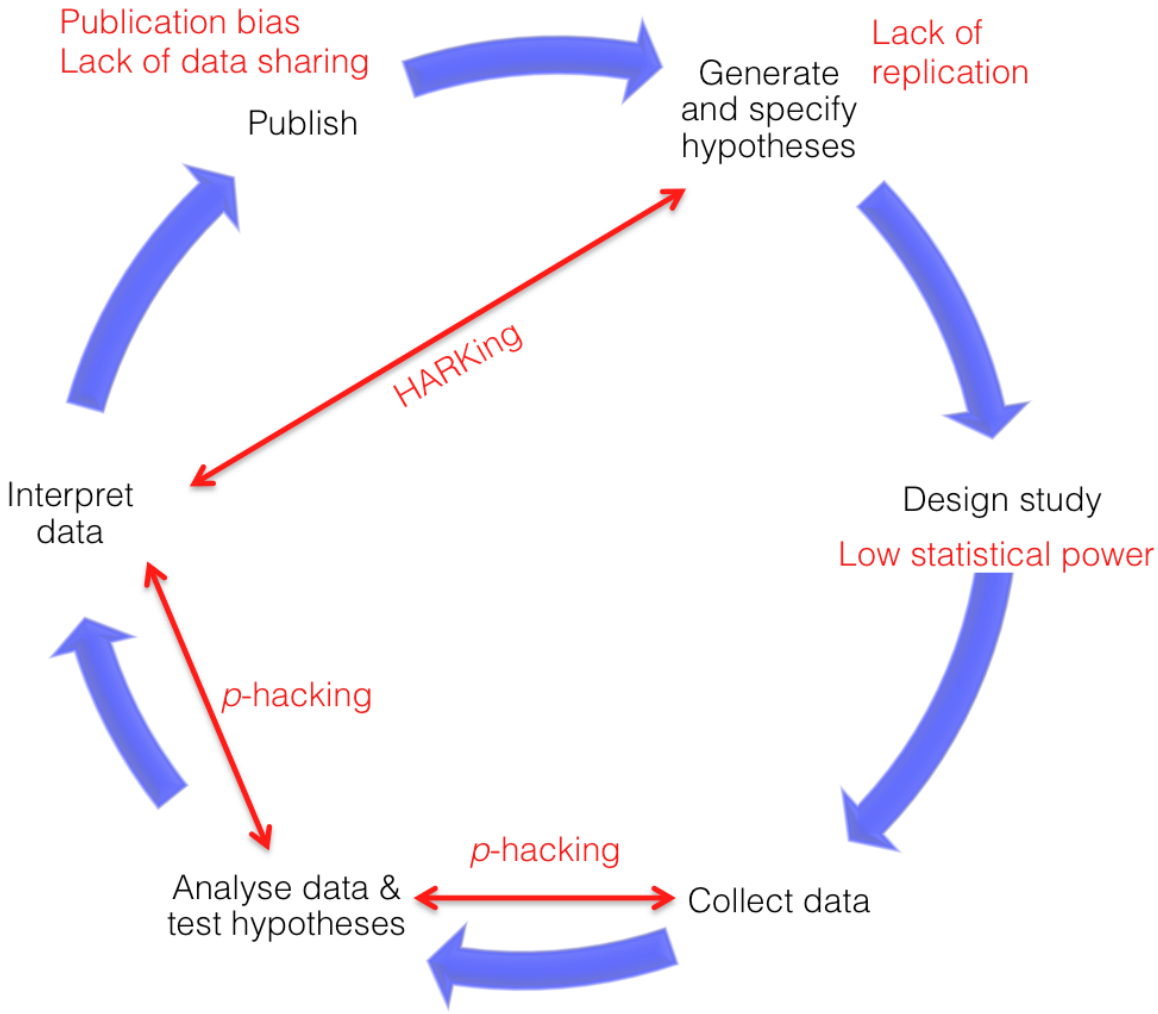
- incentives in academia conflict with what is good for science
  - the need to produce many novel and striking results
    - Frequent publication is one of the few methods to demonstrate academic talent – 'publish or perish'
  - we compete for a limited number of prestigious publication slots
  - we have to protect knowledge from competitors

http://figshare.com/articles/The_Resistable_Rise_of_Questionable_Research_Practices/1540908

http://orca.cf.ac.uk/59475/1/AN2.pdf

The University Of Queensland
AUSTRALIA
Create change

Questionable Research Practices



http://orca.cf.ac.uk/59475/1/AN2.pdf

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

OR: The perish of “Publish or Perish”

A short overview about

- What’s the problem?
  - the problems of p-values
  - the reproducibility crisis
  - questionable research practices

- Possible Solutions
  - **Understanding the consequences of claiming p < 0.05**
  - Incentives on replication
  - Preregistration
  - Blind analysis
  - Let us ignore “impact factors”
  - peer review 2.0 and new forms of publishing

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

How can we solve the problem?

Acknowledge that p-values are a very unreliable measure for significance and often misinterpreted

- p-values are often equated with strength of relationship, but small effects can have low p-values and vice versa
  - use of effect size measures can help to judge importance
- p-values are highly unstable and variable with small changes in the sampling (see the dance of the p-values: https://www.youtube.com/watch?v=u5uN3drVSkE) !
- "If you use *p*=0.05 to suggest that you have made a discovery, you will be wrong at least 30% of the time. If experiments are underpowered, you will be wrong most of the time." (David Colquhoun)

Colquhoun, David. 'An Investigation of the False Discovery Rate and the Misinterpretation of P-Values'. *Royal Society Open Science* 2014

[Head, et a. 'The Extent and Consequences of P-Hacking in Science'. *PLoS Biol* 2015]

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

Is the following statement true, or false?

We investigate an effect which is real in 30 % of the cases with a test which has a power of 35 % and

set our significance threshold to a p-value to 0.05

If we claim that we have found a significant effect,

the probability that we are wrong is 5 %

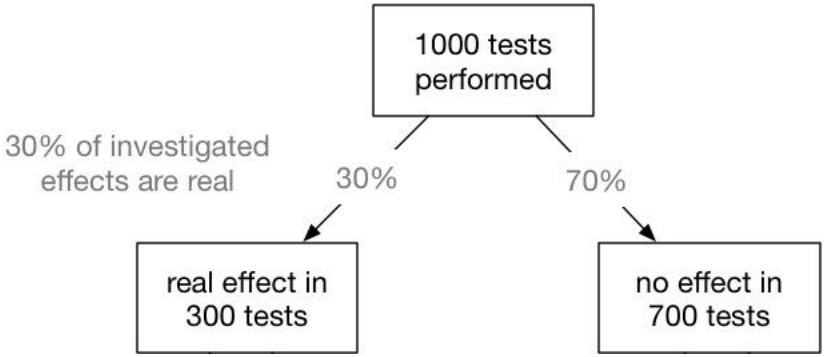
TRUE OR FALSE?

http://www.nicebread.de/whats-the-probability-that-a-significant-p-value-indicates-a-true-effect/

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

Is the following statement true, or false?

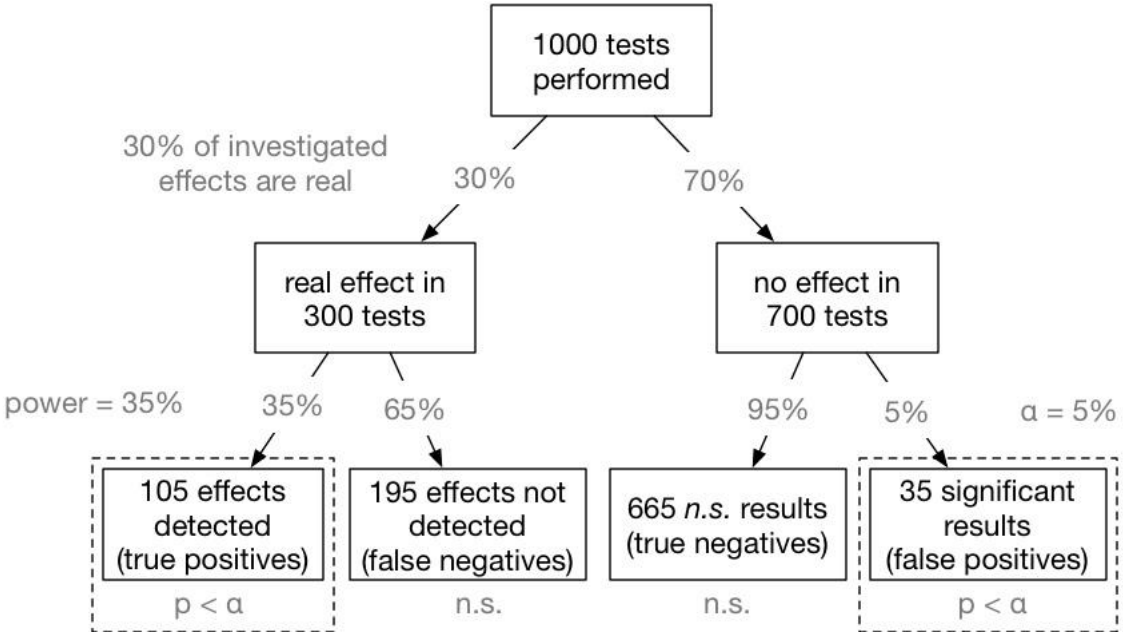


http://www.nicebread.de/whats-the-probability-that-a-significant-p-value-indicates-a-true-effect/

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

Is the following statement true, or false?

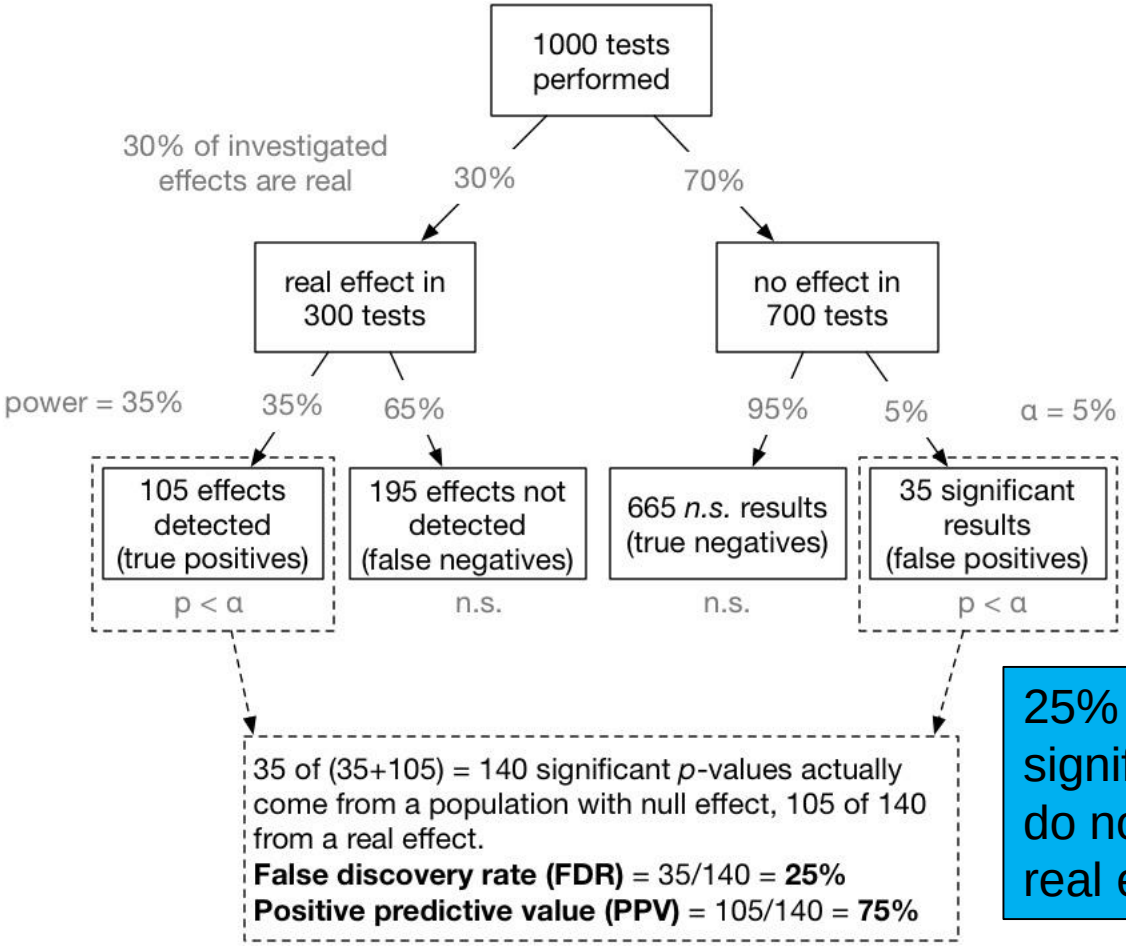


http://www.nicebread.de/whats-the-probability-that-a-significant-p-value-indicates-a-true-effect/

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

Is the following statement true, or false?



25% of all significant *p*-values do not indicate a real effect!

http://www.nicebread.de/whats-the-probability-that-a-significant-p-value-indicates-a-true-effect/

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

How can we solve the problem?

Acknowledge that p-values are a very unreliable measure for significance and often misinterpreted!

- non-parametric randomisation tests and Bayesian statistics can help

- **However:**
  - Banning p-values (1) will not solve the problem, as publication bias will happen also with the alternatives (confidence intervals, Bayesian credible intervals, …)
  - Effect size measures only partly help as they are often inflated in small sample size studies!

(1) http://www.nature.com/news/psychology-journal-bans-p-values-1.17001

[Head, et a. 'The Extent and Consequences of P-Hacking in Science'. *PLoS Biol* 2015]

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

OR: The perish of “Publish or Perish”

A short overview about

- What’s the problem?
  - the problems of p-values
  - the reproducibility crisis
  - questionable research practices

- Possible Solutions
  - Understanding the consequences of claiming $p < 0.05$
  - **Incentives on replication**
  - Preregistration
  - Blind analysis
  - Let us ignore “impact factors”
  - peer review 2.0 and new forms of publishing

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

How can we solve the problem?

Incentivise replication of effects instead of novelty

- Perform your own replication studies in-house in independent samples and publish failed replications
  - E.g.: Lane et al. ‘Failed Replication of Oxytocin Effects on Trust: The Envelope Task Case’. *PLoS ONE* 2015
- Publish failed experiments to reduce the publication bias (e.g. in a self publishing manner if journals do not accept the work)

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

OR: The perish of “Publish or Perish”

A short overview about

- What’s the problem?
  - the problems of p-values
  - the reproducibility crisis
  - questionable research practices

- Possible Solutions
  - Understanding the consequences of claiming $p < 0.05$
  - Incentives on replication
  - **Preregistration**
  - Blind analysis
  - Let us ignore “impact factors”
  - peer review 2.0 and new forms of publishing

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

How can we solve the problem?

Preregistration of studies

- Define analysis pipeline before analysing the data
- Peer review before study is conducted based on hypotheses & methods
- studies will be published regardless of the results they show
- Circumvents post-hoc hypothesis generation / cherry-picking

- Dan Simons: “It keeps us from convincing ourselves that an exploratory analysis was a planned one.”

http://www.dansimons.com/index.html

http://www.nature.com/news/how-to-make-biomedical-research-more-reproducible-1.18684?WT.mc_id=TWT_NatureNews

http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

How can we solve the problem?

Preregistration of studies

- Journals already offering registered reports:
  - Cortex
  - AIMS Neuroscience
  - Attention, Perception & Psychophysics
  - Perspectives on Psychological Science
  - Experimental Psychology
  - Drug and Alcohol Dependence
  - Social Psychology

http://www.theguardian.com/science/head-quarters/2014/may/20/psychology-registration-revolution

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

OR: The perish of “Publish or Perish”

A short overview about

- What’s the problem?
  - the problems of p-values
  - the reproducibility crisis
  - questionable research practices

- Possible Solutions
  - Understanding the consequences of claiming $p < 0.05$
  - Incentives on replication
  - Preregistration
  - **Blind analysis**
  - Let us ignore “impact factors”
  - peer review 2.0 and new forms of publishing

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

“Hide results to seek the truth”

Since the early 2000s physicists actively fight bias and error

- After realizing that new estimates of physical constants were closer to already published values (“confirmation bias”)
- By: blinding analysis, removing data labels, altering data values
- Consequence of blinding: Analytical decisions have been completed and debugged before the final result is visible to the researcher

Easy transfer to fMRI analysis possible:

- data acquired from controls and patients and we are interested in the differences between the two groups
- Why not hiding the group membership and treating all the same until analysis pipeline refined?
- If pipeline is done: reveal the group membership and look at results, but do not change pipeline

http://www.nature.com/news/blind-analysis-hide-results-to-seek-the-truth-1.18510

CRICOS Provider No 00025B

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
Create change

OR: The perish of "Publish or Perish"

A short overview about

- What's the problem?
  - the problems of p-values
  - the reproducibility crisis
  - questionable research practices

- Possible Solutions
  - Understanding the consequences of claiming $p < 0.05$
  - Incentives on replication
  - Preregistration
  - Blind analysis
  - **Let us ignore "impact factors"**
  - peer review 2.0 and new forms of publishing

CRICOS Provider No 00025B

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
Create change

How can we solve the problem?

Reduce the influence of impact factors on science careers

- Impact factors are calculated annually as the mean number of citations to articles published in any given journal in the two preceding years
  - Problem: distribution is highly skewed (85% of the articles have *fewer* citations than the average) -> **mean is inappropriate**

- It has become one of the most determinant factor in the award of grants and promotions (ARC Australia: “H-index and number of publications are discussed in the panel, with nature, science and PNAS papers holding most weight (1)”)

- Most scientists play the impact ladder game: start in high impact journals and go down, instead of aiming for a journal based on the readership and trying to publish fast
  - time-wasting and demoralising rounds of rejection and delay of information distribution

https://www.science.org.au/emcr-pathways-issue-5/behind-closed-doors-observing-arc-selection-meeting

http://occamstypewriter.org/scurry/2012/08/13/sick-of-impact-factors/

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

How can we solve the problem?

Reduce the influence of impact factors on science careers

- Right now scientists' success is measured mainly based on the amount of high-impact articles
- To increase the impact factor of journals: results that are positive, novel, eye-catching, surprising, and simple to understand are favoured
- Therefore: negative results, complicated results, or replication studies are difficult to publish and it is often not worth the effort (especially for early career scientists)
- journals fear that pre-registration reduces their impact factor and are therefore opposing it?
- impact factors positively correlate with retractions due to fraud or error (Fang et al. 2012)

Fang, et al. 'Misconduct Accounts for the Majority of Retracted Scientific Publications'. *PNAS* 2012

http://www.theguardian.com/science/blog/2013/jun/05/trust-in-science-study-pre-registration

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

OR: The perish of “Publish or Perish”

A short overview about

- What’s the problem?
  - the problems of p-values
  - the reproducibility crisis
  - questionable research practices

- Possible Solutions
  - Understanding the consequences of claiming $p < 0.05$
  - Incentives on replication
  - Preregistration
  - Blind analysis
  - Let us ignore “impact factors”
  - **peer review 2.0 and new forms of publishing**

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

Peer review

Peer review is crucial for science:

- identifies weaknesses
- encourages authors to provide more evidence (and do more work)
- Ensures that details for others to understand and replicate experiments are presented
- leads to greater reproducibility and fewer retractions

But:

- publication is heavily incentivized -> publication rates grow dramatically -> more work for reviewers, but no incentives to do good peer review
- Some scientists do not review at all or delay review for months
- Often, no discussion between authors and reviewers happens, as the journals reject quickly if there is some negative feedback
- Reviewer comments are mostly not published

[https://academickarma.wordpress.com/]

CRICOS Provider No 00025B

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
Create change

Peer review 2.0

One solution proposed by http://academickarma.org:

- Academic Karma is a journal independent peer-review network by Lachlan Coin and Louis Stowasser from Brisbane
- to get your paper reviewed you need to review other papers

- The review is open and transparent and the reviewing work can now be used to evaluate academics
  - Right now publishing papers with high impact counts in grants a lot, but delivering high quality reviews does not count at all …
- Reviews can be send to different journals avoiding unnecessary reviewer work

[https://academickarma.wordpress.com/]

Peer review 2.0

Another solution proposed by https://publons.com/

- allows to record, verify, and showcase peer review contributions in a format you can include in job and funding applications (without breaking reviewer anonymity)

Others:

- http://www.rubriq.com/
  - independent peer review service attempting to improve the publishing process, pay 100$ per review
- https://www.peerageofscience.org/

http://www.nature.com/news/the-scientists-who-get-credit-for-peer-review-1.16102

Gasparyan et al. 'Rewarding Peer Reviewers: Maintaining the Integrity of Science Communication'. Journal of Korean Medical Science 2015

CRICOS Provider No 00025B

The University Of Queensland AUSTRALIA

Create change

Post-Publication Peer-Review

- Could offer a better debate about scientific work
- Getting credit for reviews

examples:

- http://www.ncbi.nlm.nih.gov/pubmedcommons/
- https://pubpeer.com/
- https://publons.com/
- http://www.ploslabs.org/openevaluation/
- https://www.researchgate.net/publicliterature.OpenReviewInfo.html
- http://f1000research.com/about

https://www.researchgate.net/post/What_are_your_thoughts_on_and_experiences_with_open_post-publication_peer-review#view=561e67326307d94eba8b45a1

The University Of Queensland
AUSTRALIA
Create change

New forms of publishing

publication portals / megajournals for fast publication (e.g. for null-results):

- http://f1000research.com/
  - author-led process, publishing all scientific research within a few days
  - Open, invited peer review of articles is conducted after publication, focusing on scientific soundness rather than novelty or impact
- https://peerj.com/
  - 26 days until published including peer-review
  - offers publication of pre prints: draft of an article to get feedback

- Others: sciencematters.io, thewinnower.com, *PLOS ONE, BMJ Open, SAGE Open, Scientific Reports, Open Biology, SpringerPlus, …*

http://blog.impactstory.org/the-3-dangers-of-publishing-in-megajournals-and-how-you-can-avoid-them/

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

New forms of publishing

publication portals / megajournals for fast publication (e.g. for null-results):

- Advantages
  - Excellent research has been published in these journals
  - Boost citation and readership
  - They publish fast (PLOS ONE: 6months, PeerJ: 51 days, F1000: few days)
  - They are cheap (PeerJ: from 99 $)

- Disadvantages
  - They do not always have a good reputation (‘article dumping’ …)
  - Low impact factors (which shouldn’t be a disadvantage, as articles should be judged by their own merits …)

http://blog.impactstory.org/the-3-dangers-of-publishing-in-megajournals-and-how-you-can-avoid-them/

Self-Archiving of Articles

Should we pay additional money to make our articles open-access?

- Maybe not.

- there are existing repositories and maybe we should use them instead of paying additional fees for open access publishing in the established journals
- transparent feedback on our work from the whole community “post” publication adds additional value
- The output of publicly funded work has to be accessible for everyone and this should not cost additional money for society!

https://pandelisperakakis.wordpress.com/2015/09/09/how-to-negotiate-with-publishers-an-example-of-immediate-self-archiving-despite-publishers-embargo-policy/

Self-Archiving of Articles

Your are allowed to self-archive your paper after the last review step (i.e. not yet layouted by the journal)

- However, some publishers request an embargo time of 6 or 12 months (e.g. Springer)
  - This embargo time can be negotiated to 0 with the publisher: https://pandelisperakakis.wordpress.com/2015/09/09/how-to-negotiate-with-publishers-an-example-of-immediate-self-archiving-despite-publishers-embargo-policy/

To find information about the journals' policies:
http://www.sherpa.ac.uk/romeo/

https://pandelisperakakis.wordpress.com/2015/09/09/how-to-negotiate-with-publishers-an-example-of-immediate-self-archiving-despite-publishers-embargo-policy/

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change

Self-Archiving of Articles

Your are allowed to publish your work as pre-print and then later submit it to normal Journals

- This would allow that your article is read, even before peer-review and you can get feedback from the community
  - E.g. 10% of the articles on bioRxiv have comments

- http://www.theguardian.com/science/occams-corner/2015/sep/07/peer-review-preprints-speed-science-journals
- https://peerj.com/blog/post/115284878302/dorothy-bishop-on-her-preprint-experiences-at-peerj/
- https://en.wikipedia.org/wiki/List_of_academic_journals_by_preprint_policy

Self-Archiving of Articles

- http://www.zenodo.org/
  - Developed by CERN and funded by EU project FP7
- http://arxiv.org/
  - document server operated by Cornell University for pre-prints in physics, maths, engineering
- http://biorxiv.org/
  - document server operated by Cold Spring Harbor Laboratory for pre-prints in biology, neuroscience …
- http://riojournal.com/
  - publish project proposals, data, methods, workflows, software, project reports and research articles
- http://figshare.com/
  - repository where users can make all of their research outputs available in a citable, shareable and discoverable manner

How could academic publishing look like in the future?

1. replicate a study to see whether an effect from an exploratory analysis is reliable and publish as '**replication study**' regardless of result
2. improve experiment and submit the methods to a journal independent peer-review (e.g. Academic Karma) where the **reviewers get credit** for the important reviewing work they do
3. the manuscript and experiment is peer-reviewed based on the methods and the methods can be improved further and accepted in a journal as a '**pre-registered study**'
4. then the study is conducted and analysed as planned and published regardless of the results as a '**pre-registered study**'
5. then the data should be shared publicly and explored further, but the results from these analyses are published as '**exploratory studies**'
6. GOTO 1

during 1 to 5 all published work and reviewing comments should be made publically accessible independent of journals in a self-archiving fashion

CRICOS Provider No 00025B

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
Create change

Thank you for your attention.

Contact: **cai.uq.edu.au/bollmann**
Twitter: **@stebo85**
Funding: **UQPRF, NIF**

CRICOS Provider No 00025B

The University Of Queensland
AUSTRALIA
Create change