SLUB
Wir führen Wissen.

Lessons in Open Science

Preregistration and Registered Reports

Dr. Katharina Zinke
SLUB Dresden

Presentation Slides by David Riedel

16. September 2022

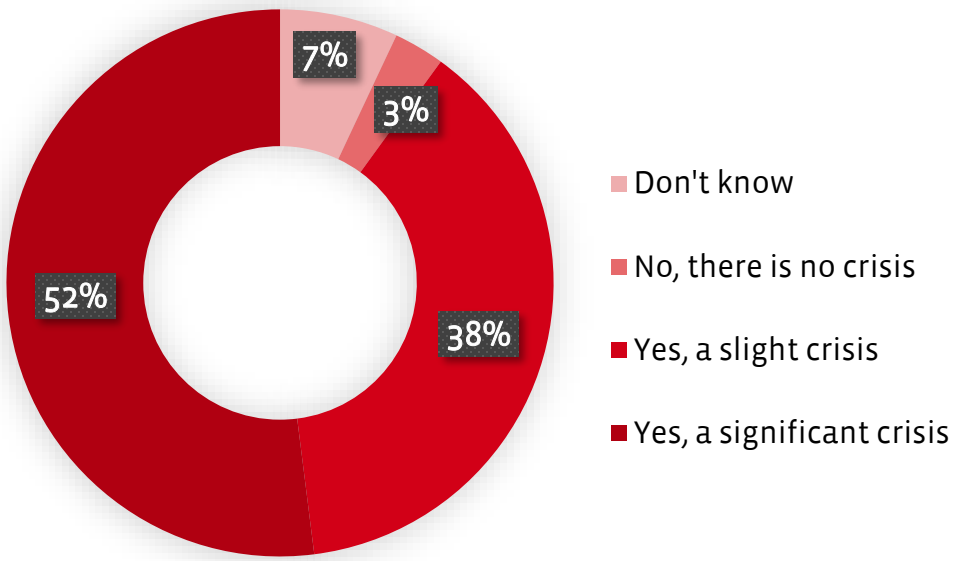
Agenda

- **Current Issues in Science**
  - *Credibility and trust*
  - *Prediction vs. postdiction*
  - *File Drawer, p-Hacking, HARKing, multiverse*
- Definition and Merits
  - *Benefits to science in general*
  - *Personal benefits*
  - *Adopting preregistration*
- Process of Preregistration
  - *Different methods*
  - *Registered Reports*
- Common Obstacles and Concerns

Credibility of Science?

- Neither laypersons nor professional scientists trust in the majority of published results in the life sciences

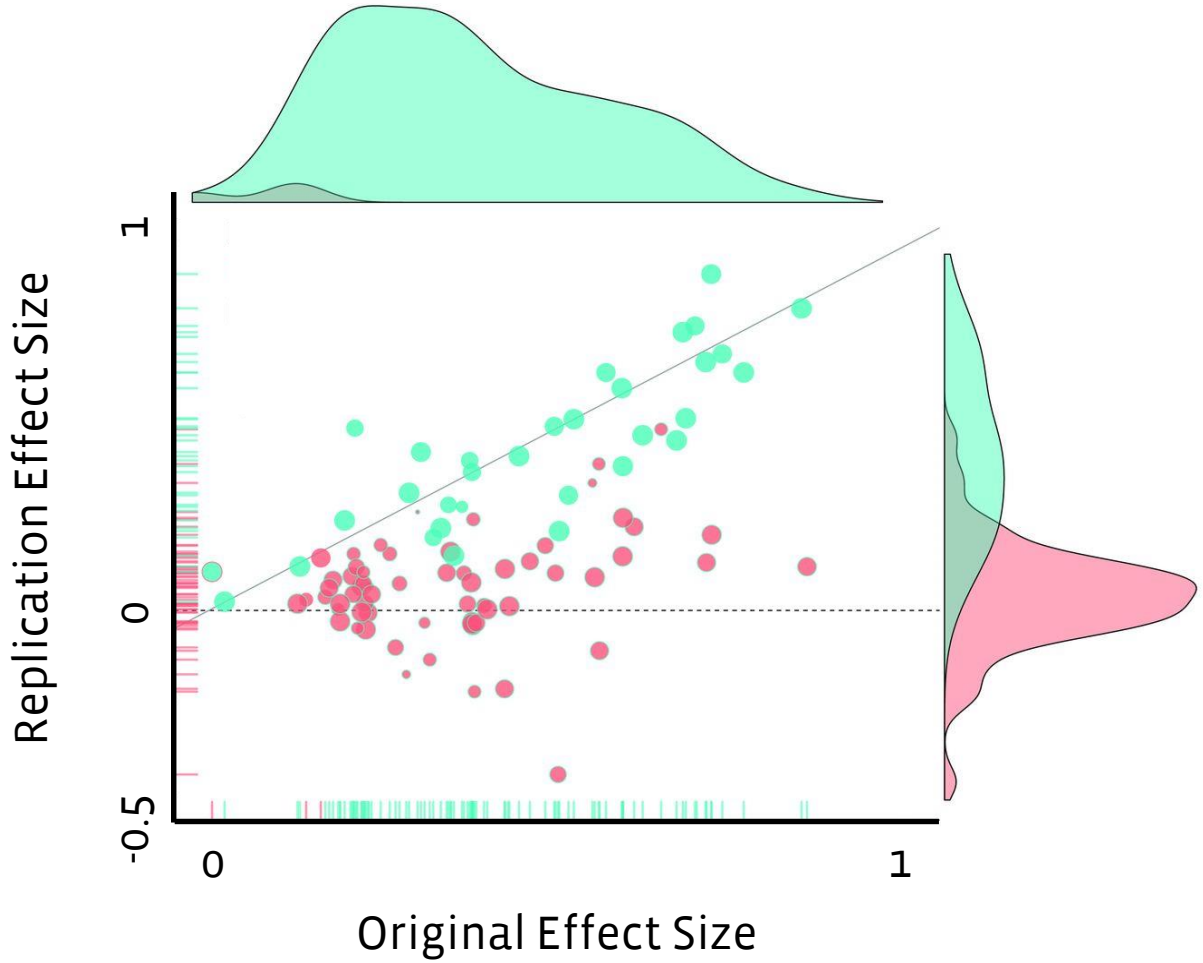
- Average statistical power is much too low (20-50%) considering the published positive rate (>90%) *(Nosek, 2018)*



Adapted from Baker, 2016

Replication Crisis in Psychology

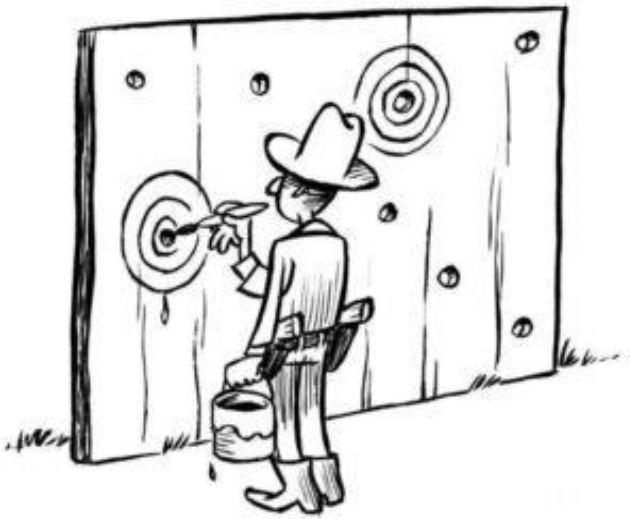
- Replication of 100 studies
- Significant results decreased from 97% to 36%
- Reasons: coincidence/error, biases, questionable research practices



Open Science Collaboration, 2015

HARKing and p-Hacking

Hypothesizing **A**fter
Results are **K**nown



https://ispgr.org/open-science-and-the-power-of-pre-registration/

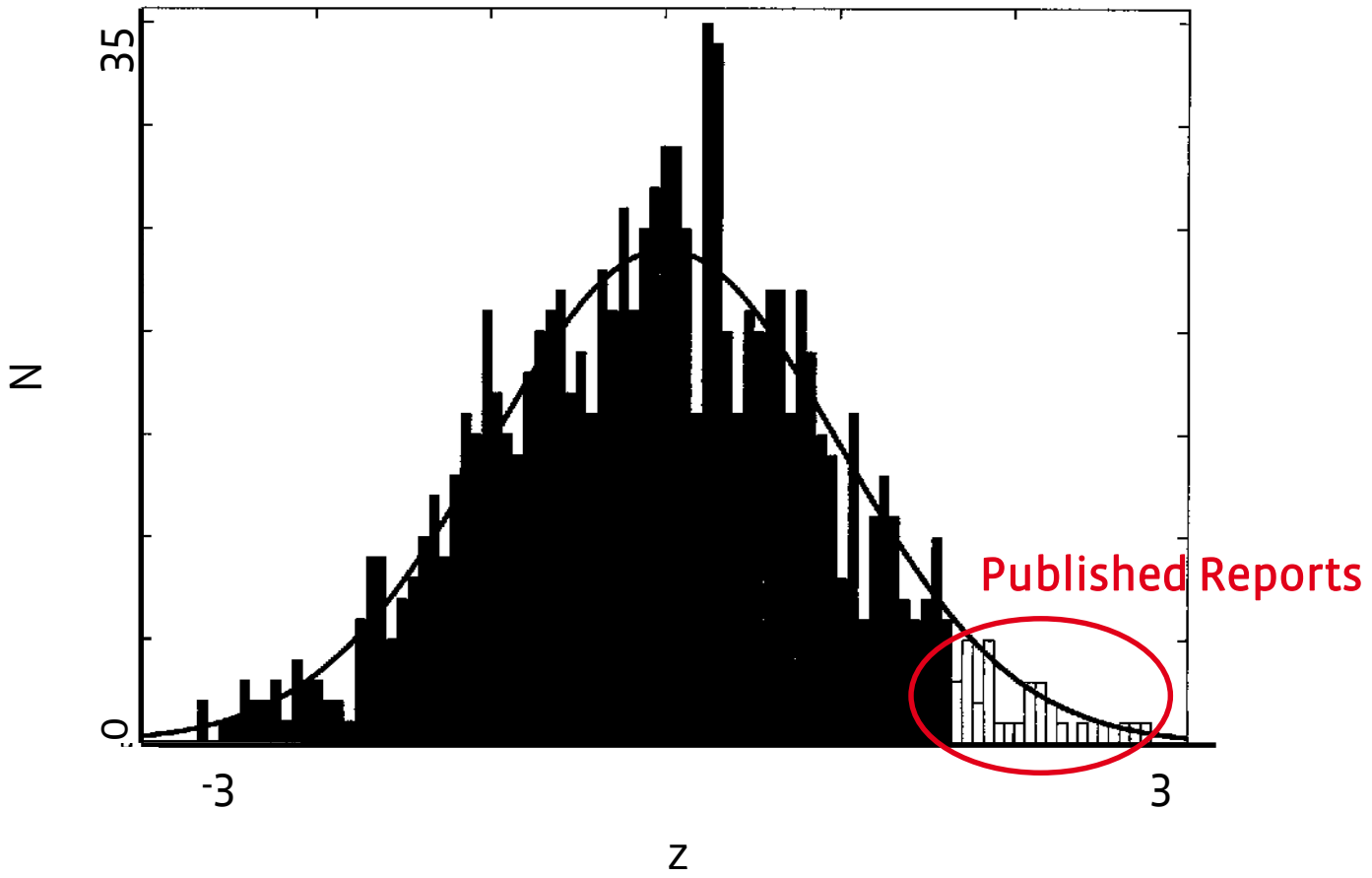
p-Hacking

"If you don't reveal some insights soon, I'm going to be forced to slice, dice, and drill!"

https://atozmarkets.com/news/untold-reality-of-p-hacking-in-finance/

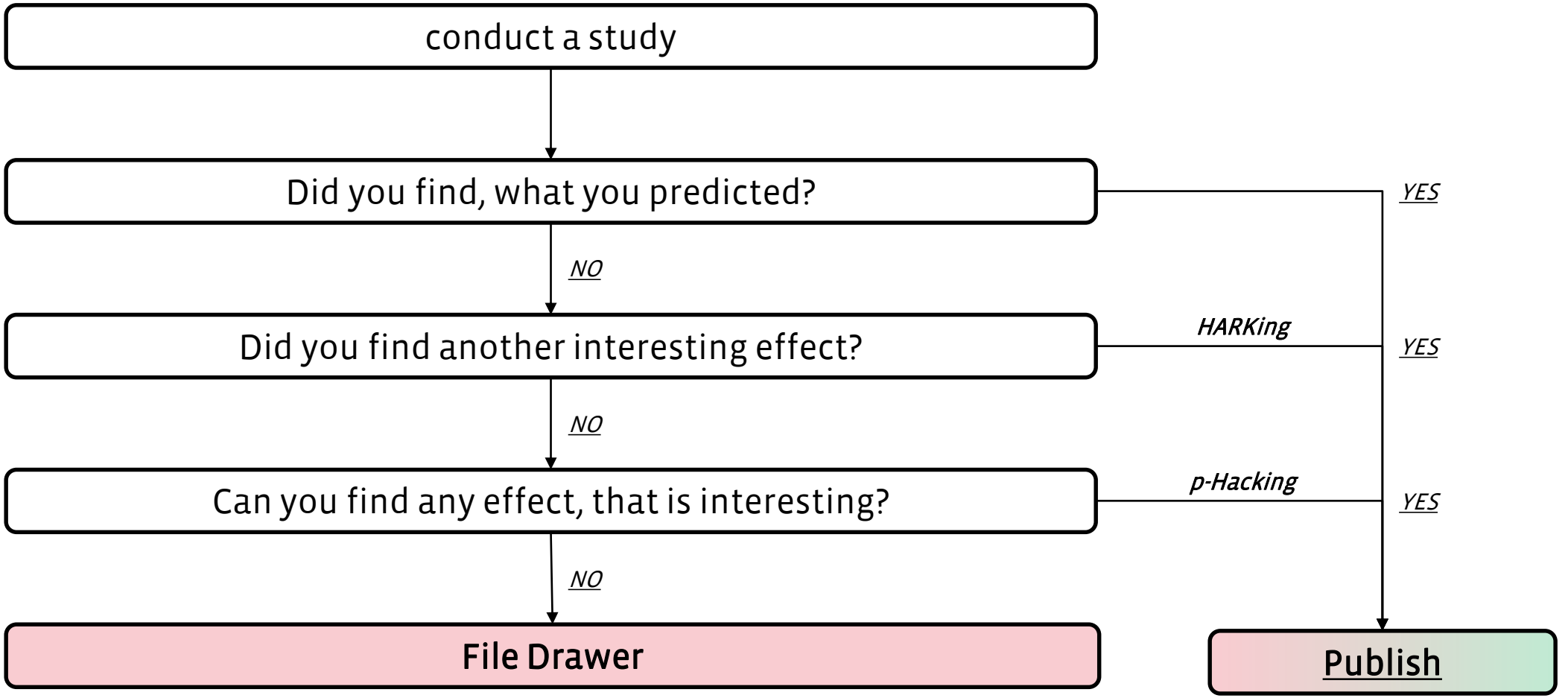
File Drawer

"A publication bias exists if the probability that a study reaches the literature [...] depends on the results of the study."



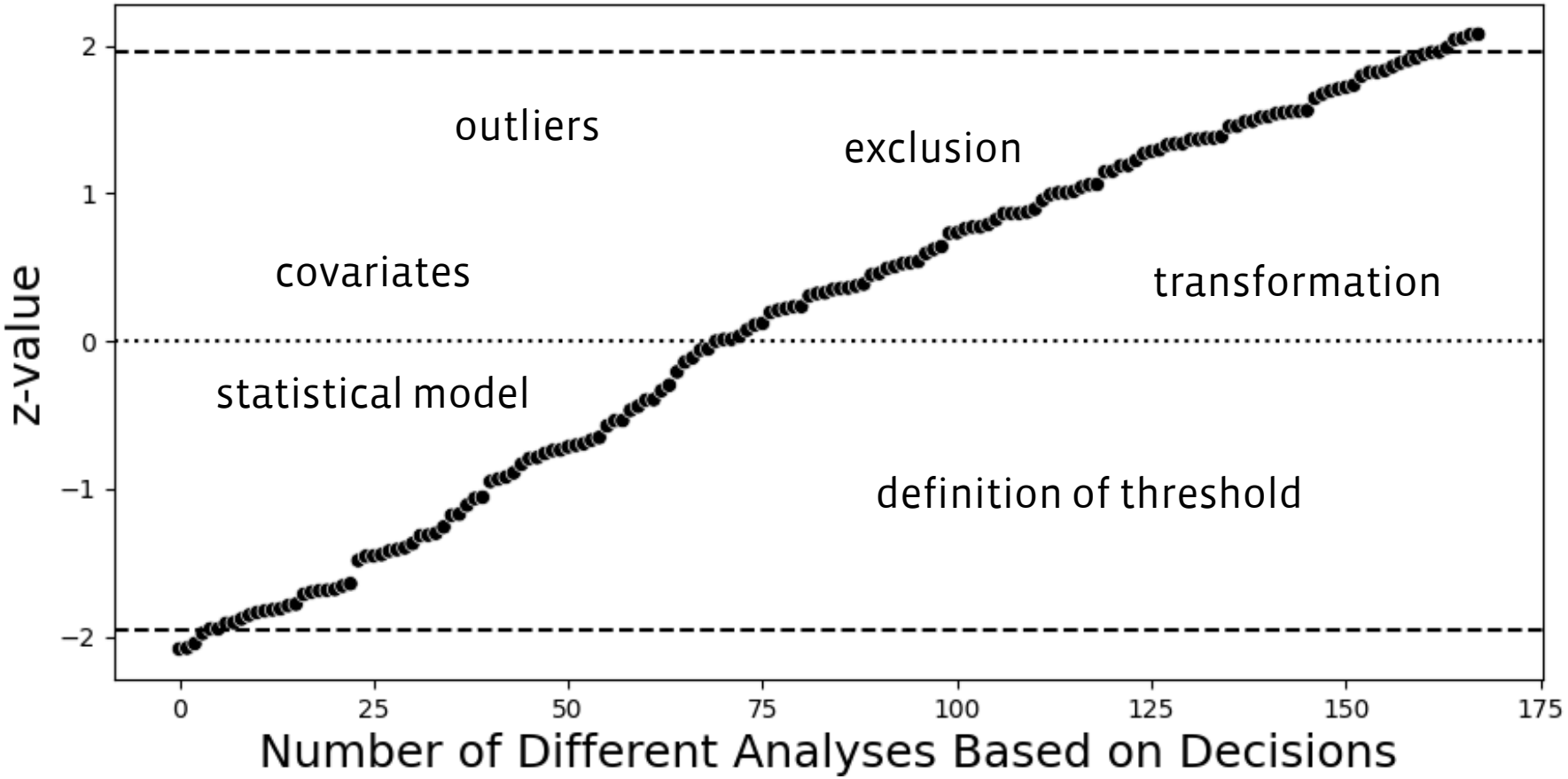
Adapted from Scargle, 1999

The Path to Incredibility

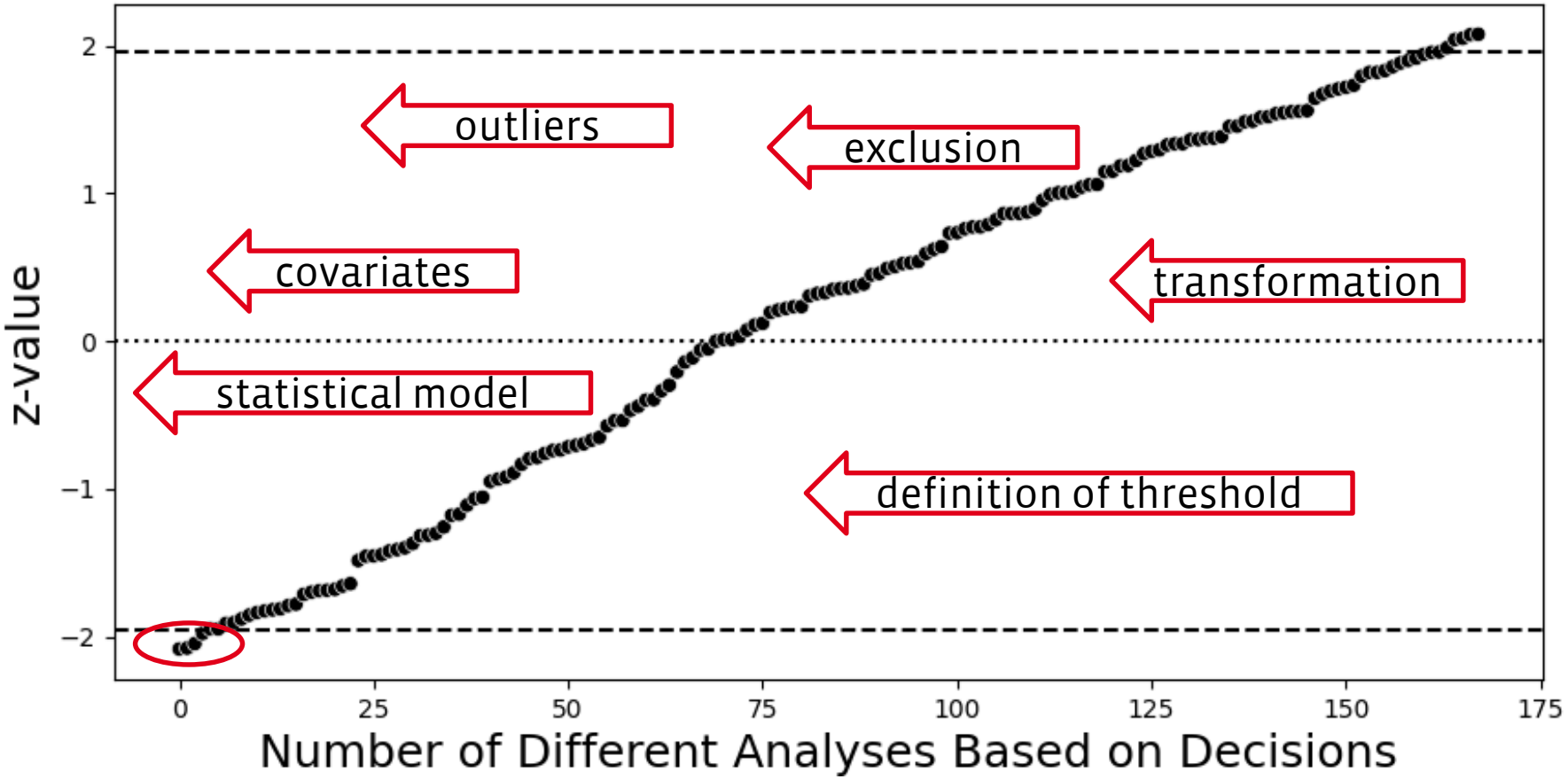


adapted from Brian Nosek, 2018: https://osf.io/9m6tx/

Multiverse


outliers
exclusion
covariates
transformation
statistical model
definition of threshold
z-value
2
1
0
−1
−2
0
25
50
75
100
125
150
175
Number of Different Analyses Based on Decisions

Multiverse


outliers
exclusion
covariates
transformation
statistical model
definition of threshold
z-value
2
1
0
−1
−2
0
25
50
75
100
125
150
175
Number of Different Analyses Based on Decisions

Prediction vs. “Postdiction”

property	prediction	“postdiction”
Influenced by investigation and outcome?	No	Yes
Null hypothesis significance testing (p-values) is fully applicable?	Yes	No
Confirmatory or exploratory?	confirmatory	exploratory
Valuable to the scientific process?	Yes	Yes

We only have to be able to differentiate between the two approaches.

Agenda

- Current Issues in Science
  - *Credibility and trust*
  - *Prediction vs. postdiction*
  - *File Drawer, p-Hacking, HARKing, multiverse*
- **Definition and Merits**
  - ***Benefits to science in general***
  - ***Personal benefits***
  - ***Adopting preregistration***
- Process of Preregistration
  - *Different methods*
  - *Registered Reports*
- Common Obstacles and Concerns

What is preregistration?

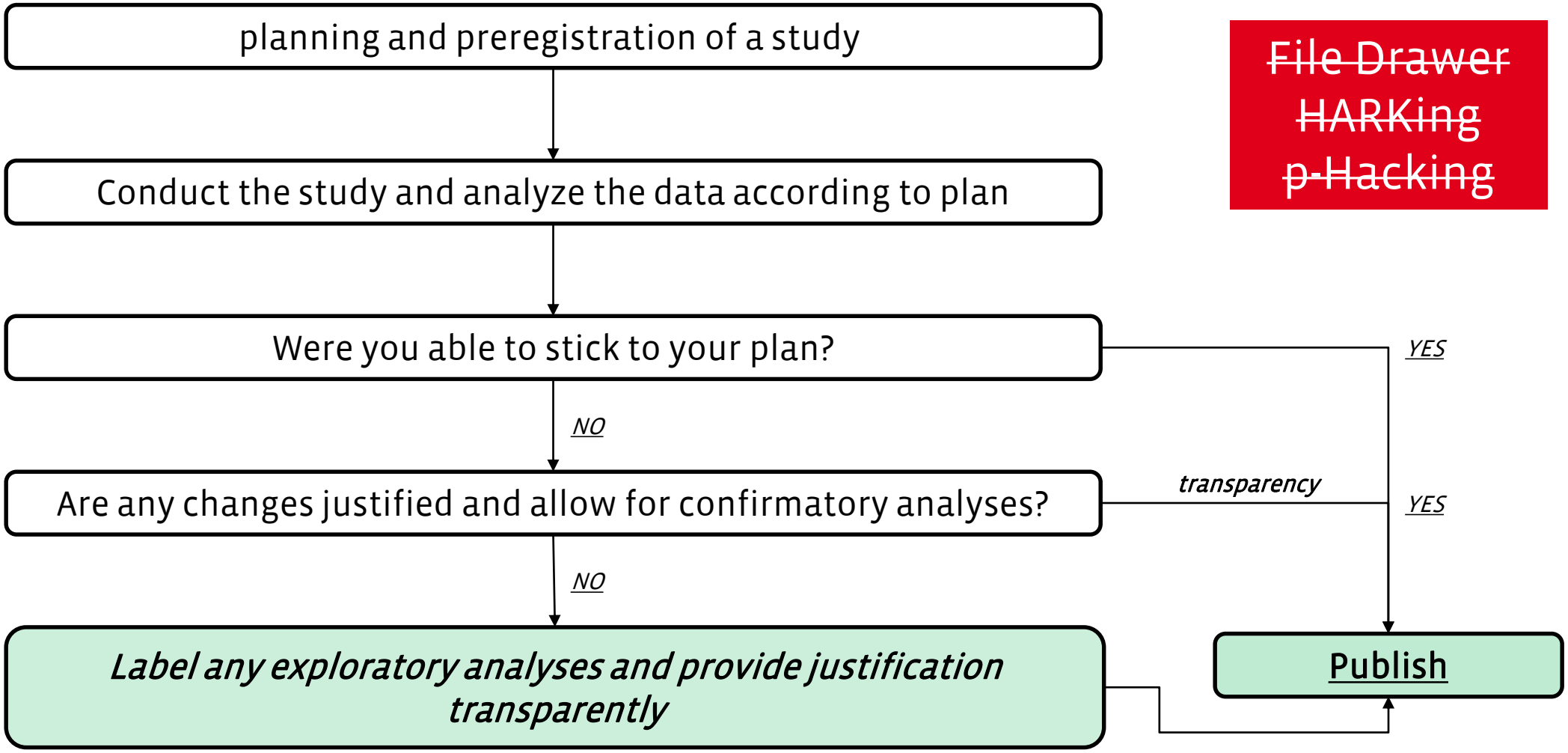
- A priori documentation/publication of information on planned investigation, such as:
    - *study design, methods, hypotheses, analysis plan*
- Independent party registry
    - *E.g. Open Science Framework (OSF)*
    - *Embargo upon submission is possible*
- Time Stamp
- Immutable
    - *Can not be changed or deleted (without record)*

Key principles

- Avoid bias/influence by making important decisions and careful planning in advance
- Be as transparent as possible
- Changes and exploratory analyses are no problem, as long as you report and are able to justify them

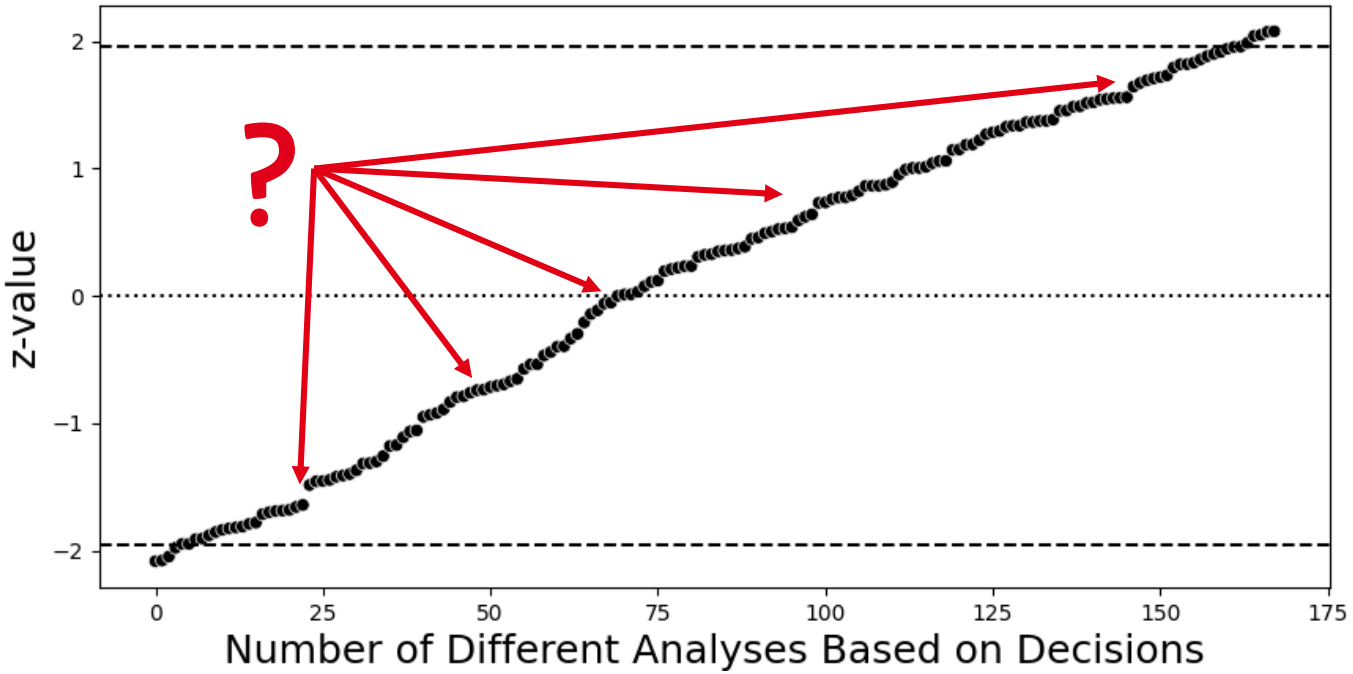
Be aware of these principles and of what you are trying to achieve. If you are honest, preregistration will come naturally to you!

The Path to Credibility



~~File Drawer~~
~~HARKing~~
~~p-Hacking~~

Multiverse



By making analytical decisions in advance, preregistration ensures arbitrary results within the multiverse.

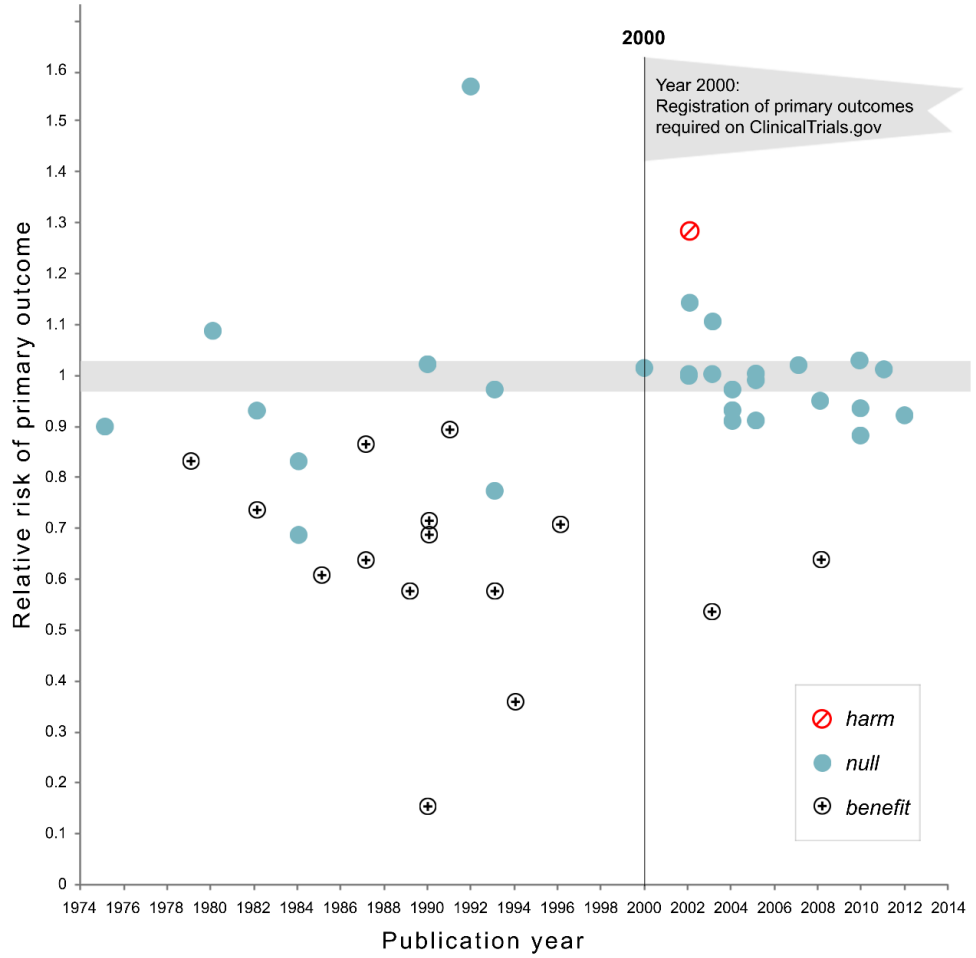
Is preregistration effective?

Rate of positive results dropped from 57% to 8% with the introduction of obligatory preregistration.

Kaplan & Irvin, 2015

"The strongest factor associated with the false positive or true positive study outcome was if the study had a specific a priori hypothesis."

Swaen, Teggeler & van Amelsvoort, 2001



Benefits on an individual level

- Reduce implicit bias
  - *Improve quality of research*
- Distinction between confirmatory and exploratory investigations
- Increases focus on project management
  - *Preregistration involves increased careful planning*
- Initialization of collaborations
  - *Involvement of all project partners*
- Open Science Badges

https://osf.io/tvyxz/wiki/1.%20View%20the%20Badges/

Personal Benefits

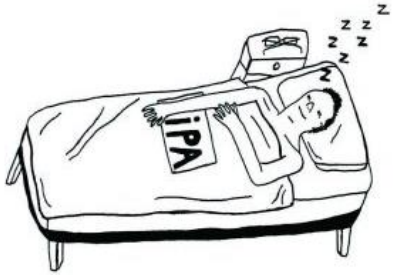
1. Take credit for your predictions.

4. Profit from online resources.

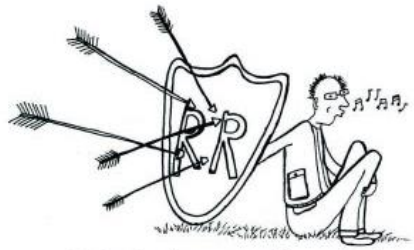
5. Increase your reputation and self-image.

2. Experience the excitement.

3. Prevent the data from taking you hostage.



6. Await your results without fear with in-principle acceptance.

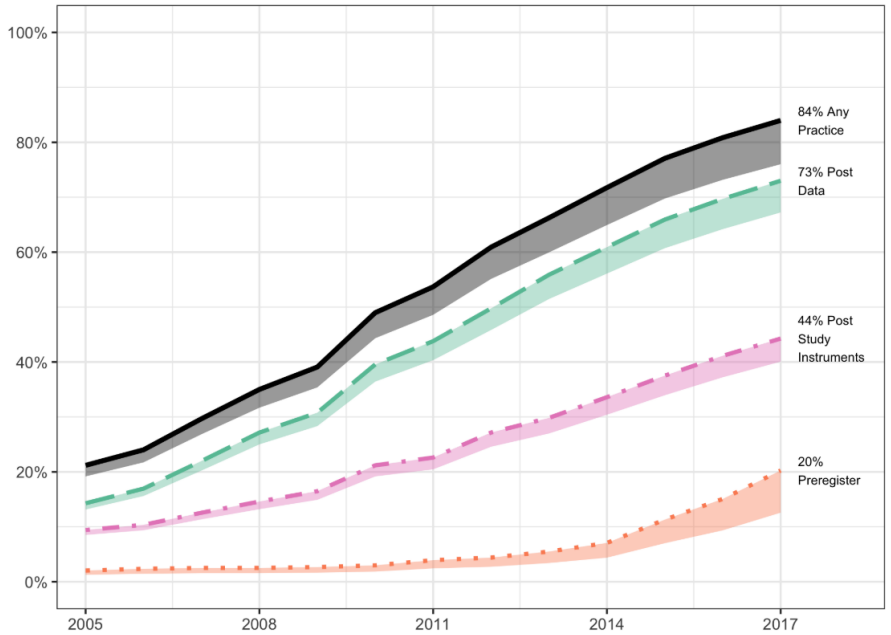


7. Protect yourself against post-hoc critique.

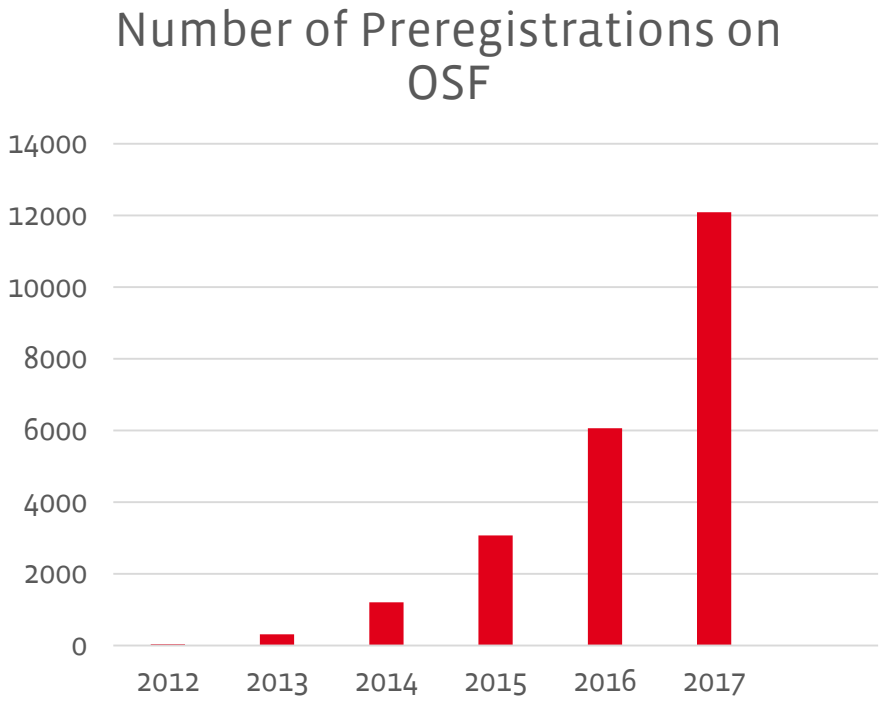
Illustrations by Stella de Kort, www.stelladekort.nl

Wagenmakers & Dutilh, 2016

Adopting preregistration

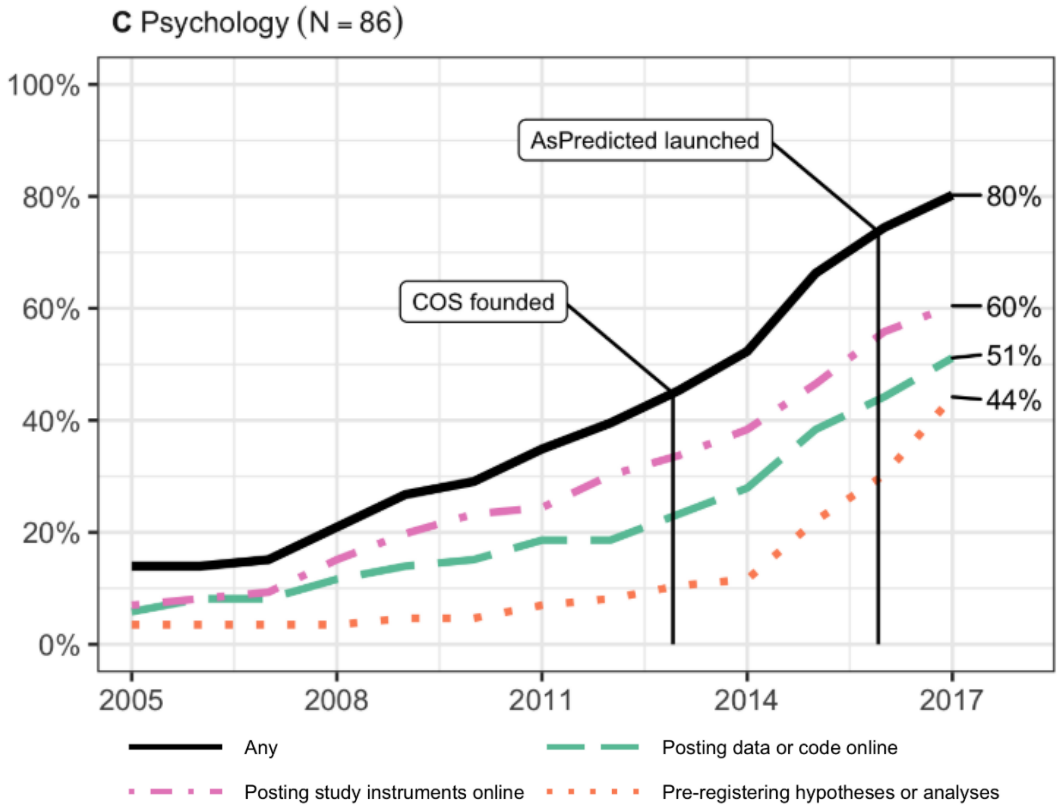


Christensen et al., 2020

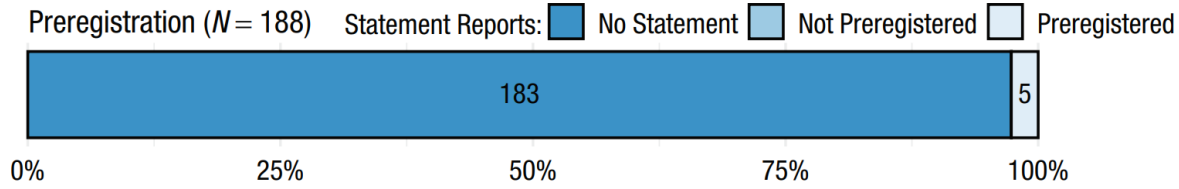


Adapted from Nosek & Lindsay 2018 https://www.psychologicalscience.org/observer/preregistration-becoming-the-norm-in-psychological-science

Adopting preregistration (Psychology)



Christensen et al., 2020



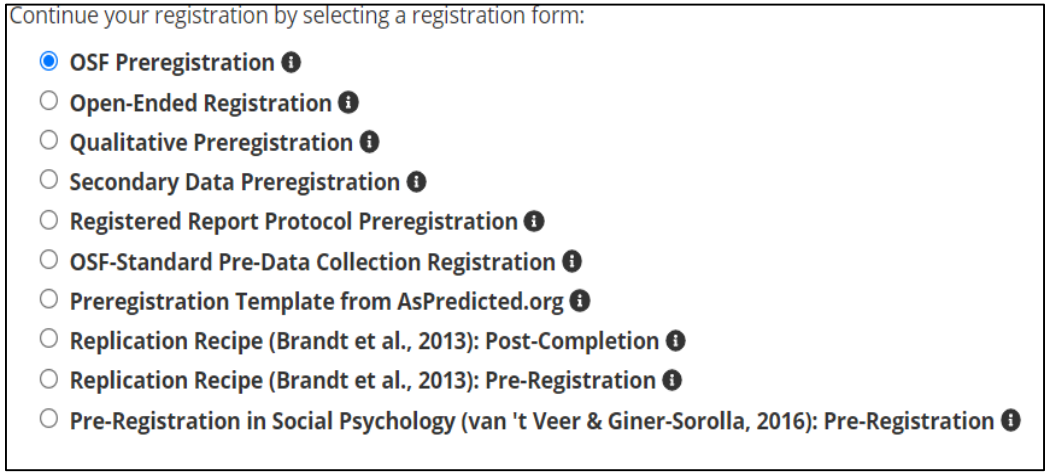
Hardwicke et al., 2021

Agenda

- Current Issues in Science
  - *Credibility and trust*
  - *Prediction vs. postdiction*
  - *File Drawer, p-Hacking, HARKing, multiverse*
- Definition and Merits
  - *Benefits to science in general*
  - *Personal benefits*
  - *Adopting preregistration*
- **Process of Preregistration**
  - ***Different methods***
  - ***Registered Reports***
- Common Obstacles and Concerns

Content and Templates

- Metadata
  - *Title, Contributors, Subject, Tags…*
- Study Design
  - *Type, Blinding, Randomization…*
- Sampling
  - *Data collection, Subjects, Sample size…*
- Analysis
  - *Variables, Manipulation, Aggregation, Statistics, outlier/missing data policy…*

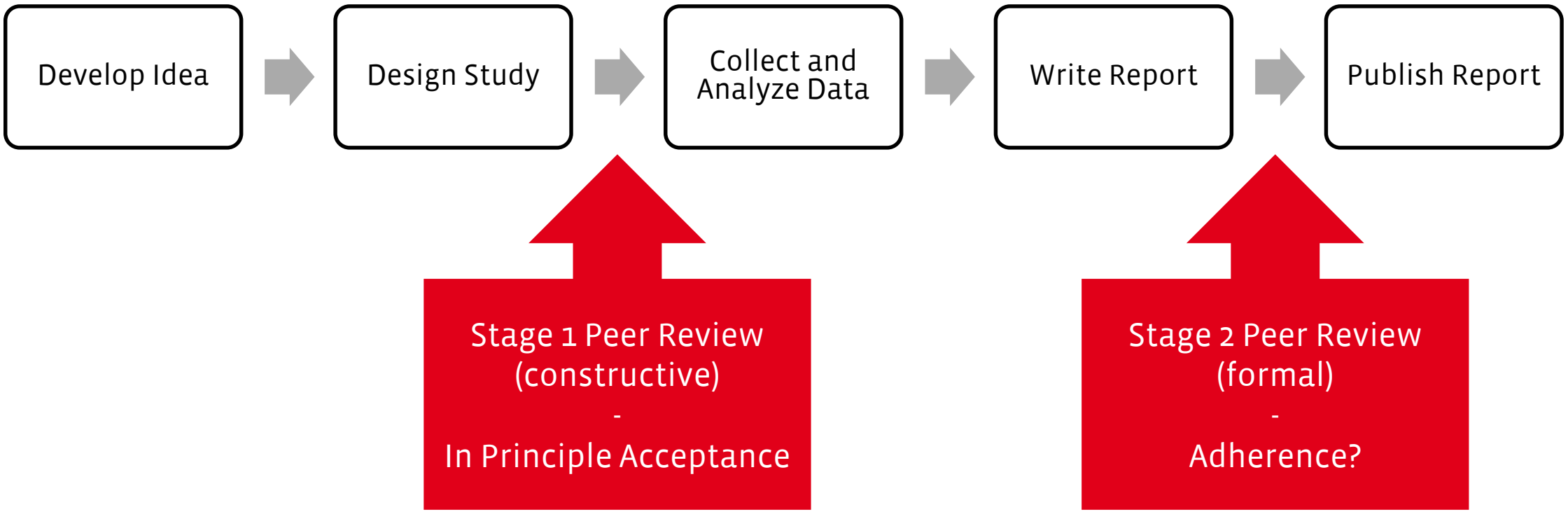


https://osf.io

How to preregister

- Online registries
  - *e.g. OSF, AsPredicted, Prospero*
- Registered Reports
- Other publication methods
  - *Study protocol*
  - *Poster at conference*

Registered Report



Adapted from https://osf.io/rr/

Methods

Method	Pro	Con
Online Registry	• Easy and fast • Most widely used • Stepwise approach easily applicable	• No quality assurance • Harder to discern between good and bad registration
Registered Reports	• Gold standard • “In Principle Acceptance” status • Constructive review process • Counteracts File Drawer	• Takes much more time • Less predictable
Other publication of study protocol (article, conference poster …)	• Additional publication	• Focusses on the dissemination of methods rather than decreasing degrees of freedom

Agenda

- Current Issues in Science
  - *Credibility and trust*
  - *Prediction vs. postdiction*
  - *File Drawer, p-Hacking, HARKing, multiverse*
- Definition and Merits
  - *Benefits to science in general*
  - *Personal benefits*
  - *Adopting preregistration*
- Process of Preregistration
  - *Different methods*
  - *Registered Reports*
- **Common Obstacles and Concerns**

Obstacles and Concerns

- Too much extra work!
  - *It is not going to be wasted, but worth it*
- Too restrictive, science needs freedom!
  - *You are free to explore, just label accordingly*
- It‘s not going to stop fraud anyway!
  - *No, but it is still helpful and might make it harder*
- Someone is going to steal my ideas!
  - *Embargo*

Nosek et al., 2018

Thank you for your attention!

Further information in the OPAL-Tutorial (in English and German):

https://bildungsportal.sachsen.de/opal/auth/RepositoryEntry/31574065165?0

Questions / Support : openscience@slub-dresden.de

References

- Christensen, G., Wang, Z., Levy Paluck, E., Swanson, N., Birke, D., Miguel, E., & Littman, R. (2020). Open Science Practices are on the Rise: The State of Social Science (3S) Survey. *UC Berkeley: Center for Effective Global Action.* https://escholarship.org/uc/item/0hx0207r
- Wagenmakers, E.-J., & Dutilh, G. (2016). Seven Selfish Reasons for Preregistration. *APS Observer*, *29*(9), 13–14.
- Aarts, A. A., Anderson, J. E., Anderson, C. J., Attridge, P. R., Attwood, A., Axt, J., Babel, M., Bahník, Š., Baranski, E., Barnett-Cowan, M., Bartmess, E., Beer, J., Bell, R., Bentley, H., Beyan, L., Binion, G., Borsboom, D., Bosch, A., Bosco, F. A., … Zuni, K. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251). https://doi.org/10.1126/science.aac4716
- Kaplan, R. M., & Irvin, V. L. (2015). Likelihood of null effects of large NHLBI clinical trials has increased over time. *PLoS ONE*, *10*(8), 1–12. https://doi.org/10.1371/journal.pone.0132382
- Scargle, J. D. (1999). *Publication Bias (The "File-Drawer Problem") in Scientific Inference*. http://arxiv.org/abs/physics/9909033
- Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(11), 2600–2606. https://doi.org/10.1073/pnas.1708274114
- Swaen, G. G. M. H., Teggeler, O., & Van Amelsvoort, L. G. P. M. (2001). False positive outcomes and design characteristics in occupational cancer epidemiology studies. *International Journal of Epidemiology*, *30*(5), 948–954. https://doi.org/10.1093/ije/30.5.948
- Baker, M. (2016). *IS THERE A REPRODUCIBILITY CRISIS?* 452–454.