Hilmar Brohmer

Practicing
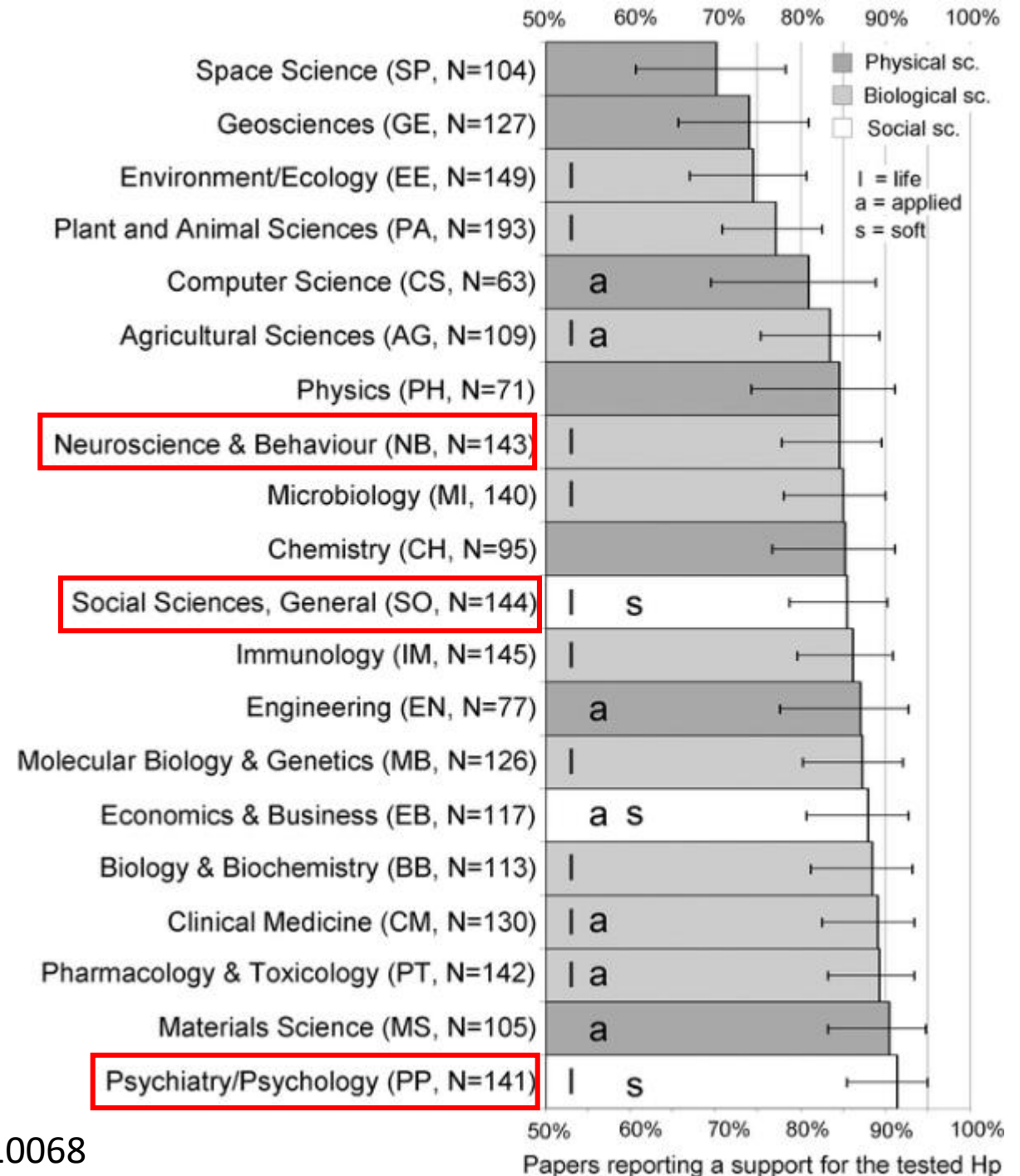# Open Science

A brief introduction

Contents

- Why Open Science (OS) is a big topic in psychology (and beyond)
- Central OS practices
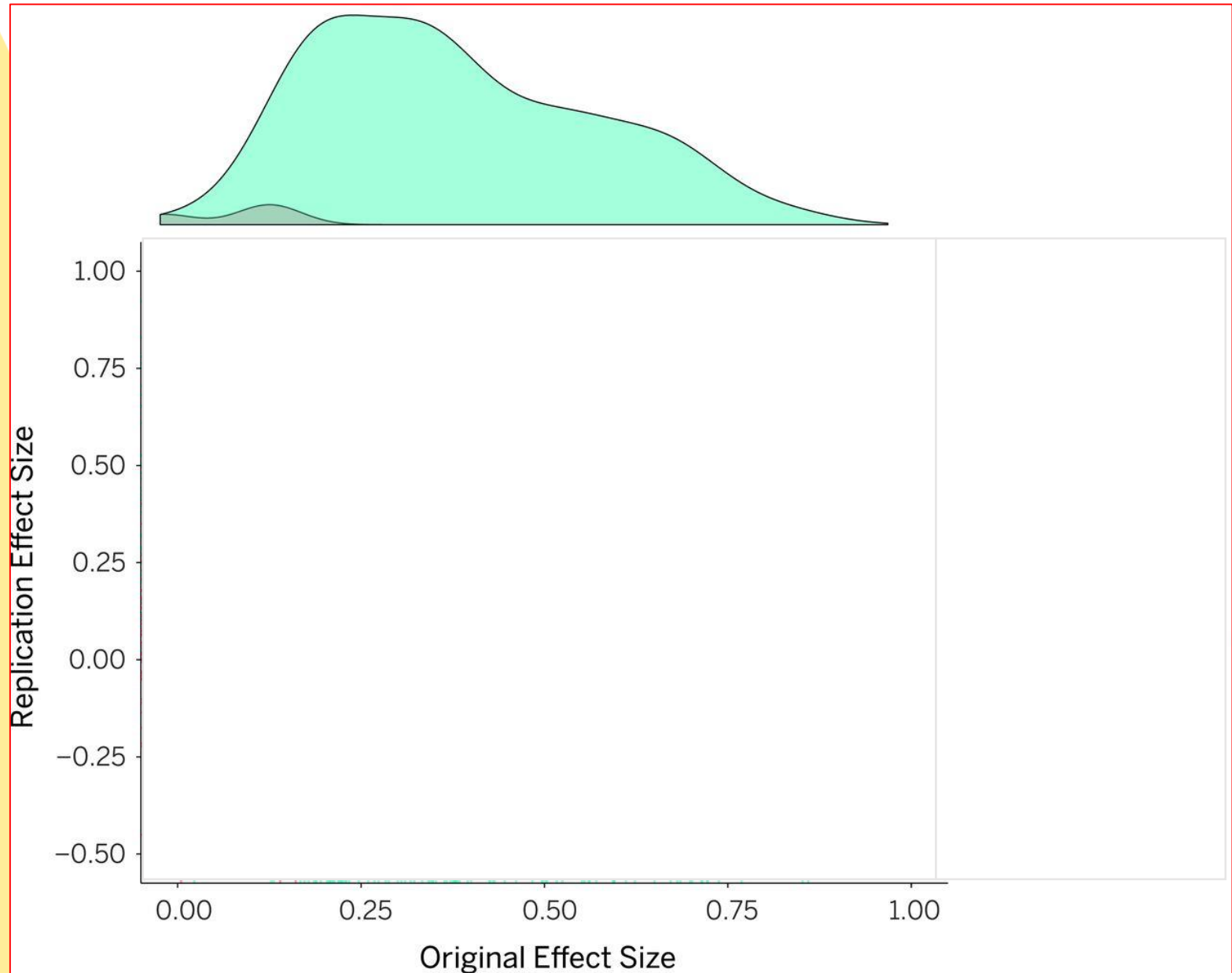- Suggestions how to start with OS practices

Most published research has positive results

If this is true, this implies that most researchers are mostly right with their hypotheses
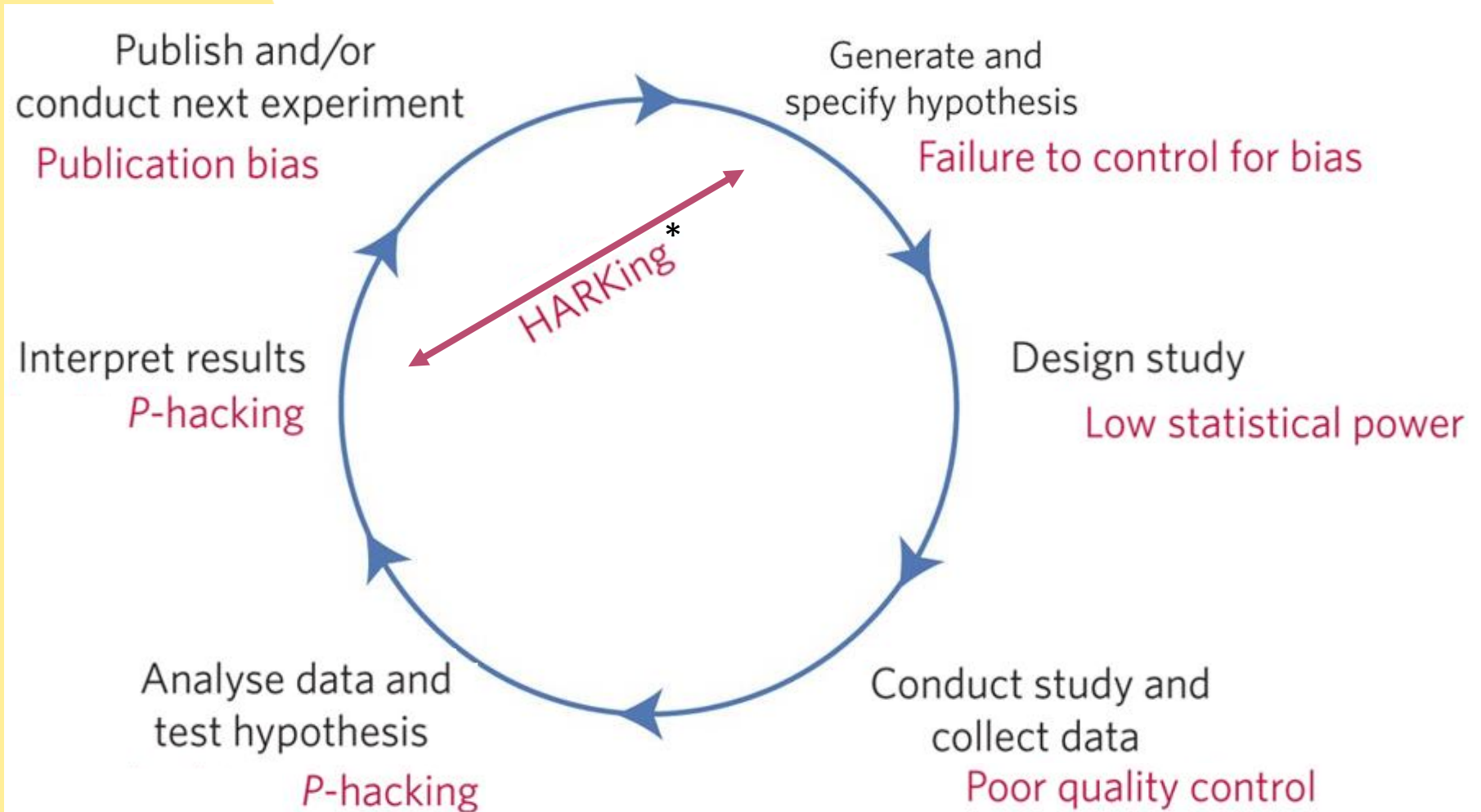
But is this really true?

Space Science (SP, N=104)
Geosciences (GE, N=127)
Environment/Ecology (EE, N=149)
Plant and Animal Sciences (PA, N=193)
Computer Science (CS, N=63)
Agricultural Sciences (AG, N=109)
Physics (PH, N=71)
Neuroscience & Behaviour (NB, N=143)
Microbiology (MI, 140)
Chemistry (CH, N=95)
Social Sciences, General (SO, N=144)
Immunology (IM, N=145)
Engineering (EN, N=77)
Molecular Biology & Genetics (MB, N=126)
Economics & Business (EB, N=117)
Biology & Biochemistry (BB, N=113)
Clinical Medicine (CM, N=130)
Pharmacology & Toxicology (PT, N=142)
Materials Science (MS, N=105)
Psychiatry/Psychology (PP, N=141)

Physical sc.
Biological sc.
Social sc.

l = life
a = applied
s = soft

Papers reporting a support for the tested Hp

NO!

How come there is this descrepancy?

Questionable Research Practices
- these mostly happen without bad intent
- our **confirmation bias** is strong in us (Bishop, 2020, https://doi.org/10.1038/d41586-020-02275-8)

Munafo et al., 2017, https://doi.org/10.1038/s41562-016-0021

Publish and/or conduct next experiment
Publication bias

Generate and specify hypothesis
Failure to control for bias

Interpret results
P-hacking

HARKing *

Design study
Low statistical power

Analyse data and test hypothesis
P-hacking

Conduct study and collect data
Poor quality control

* Hypothesizing After Results are Known

A non-exaustive list of OS practices

https://opensocialwork.org/research/open-science/
https://osf.io/tvyxz/wiki/home/

Open Data & Code

OPEN DATA

Heycke, 2019, https://osf.io/djp2f/
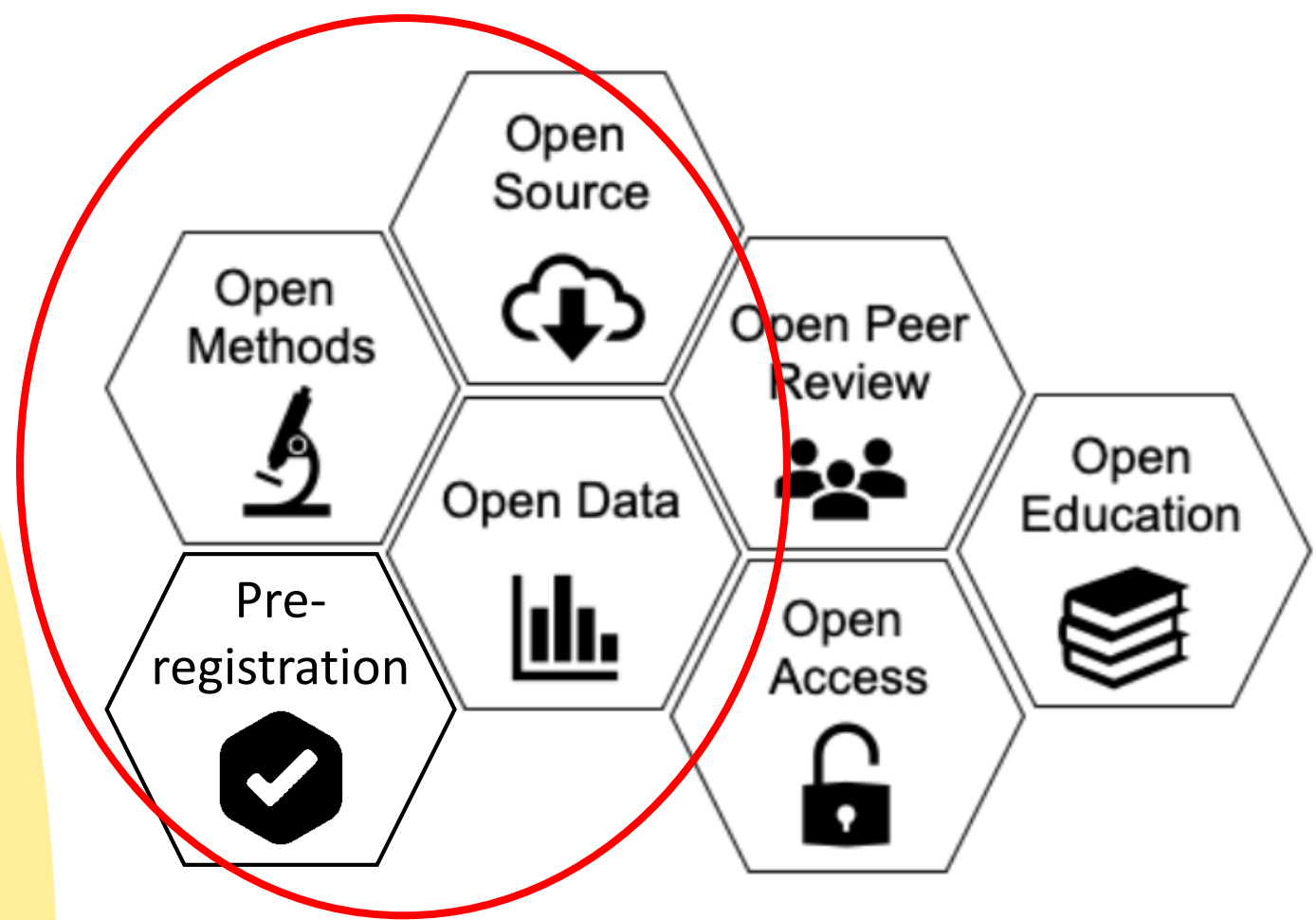
- Basic assumption:
- When I run the **same analysis** with the **same data** and the **same code**, I should get the **same results** as reported in the paper
- → *statistical reproducibility*

- I can't check this basic assumption because data is mostly not reported

Open Data & Code

- Resistance to data sharing

"To what extent do you agree with the
following statements about barriers related to data sharing?"

Statement	Disagree	Neutral	Agree
"Sharing data is not a common practice in my field"	18%	14%	68%
"I prefer to share data upon request"	26%	17%	57%
"Preparing data is too time-consuming"	33%	12%	55%
"I never learned to share data online"	35%	10%	54%

And > 40% fear that data will be **misinterpreted** or that they are **getting scooped**.

Open Materials

- Similar arguments have been stated for sharing materials

(i.e., how exactly an experiment or study was set up)

➔ ***But materials sharing is vital for replication studies***

Without the exact materials, you will only achieve a „conceptual" replication, which may yield completely different results

Starting with open data and open materials

- **Open Science Framework:**
  https://osf.io/
- **Austrian Social Science Data Achieve:**
  https://aussda.at/
- **OpenNEURO: https://openneuro.org/**
- **GitHub: https://github.com/**

- Check out online repositories
- Prepare your data and materials (i.e., make them understandable and keep GDPR in mind)
- Upload them and give appropriate descriptions (e.g., read mes, meta data, etc)
- DEFINITELY provide a link to them in the article (e.g., in the final version)

BTW: instead of "open", data and materials should be assured to be "FAIR" (findable, accessible, interoperable, reusable). See https://www.tugraz.at/sites/rdm/the-fair-principles/

Preregistration & Registered Reports

PREREGISTERED

- A *prereg* (as the cool kids say) is a written piece, where your research idea, research design, sample size / data generation, and analysis plan is specified

- The clue: it's done **BEFORE** you have looked into the data or conducted a study

It's NOT about putting you in chains (or dismiss data exploration)!

- We want to know what was known, assumed, and hypothesized before the data (prediction)

- But we also want to know what was thought after the data was inspected and how it relates to the predictions (postdiction)

➔ *Preregs is for distinguishing prediction from postdiction (mainly in confirmatory analyses)*

Nosek et al., 2018, https://doi.org/10.1073/pnas.1708274114
Nosek et al., 2019, https://www.doi.org/10.1016/j.tics.2019.07.009

Important distinction!

Preregs:

- time-stamped and saved on a repository before data collection

- A clear guideline mainly for your future self

- Not peer-reviewed (although some reviewers want to check it out and you want to provide a **link** for it in your paper)

- https://www.cos.io/initiatives/prereg

Registered Reports (RRs)

- have a peer-review before data collection (Stage 1)

- not in a prereg form, but written as if you write an article (i.e., comprehensive theory and methods)

- A positive Stage 1 usually guarantees a later publication independent of the results

- https://www.cos.io/rr
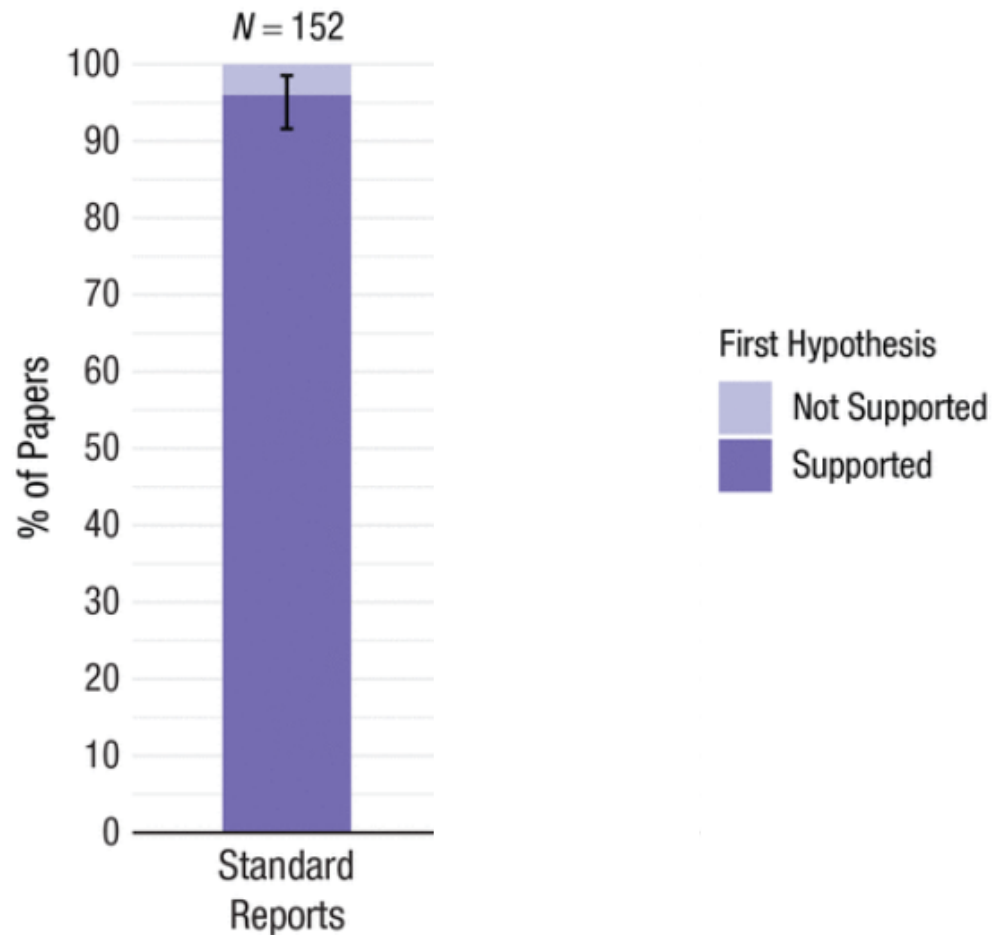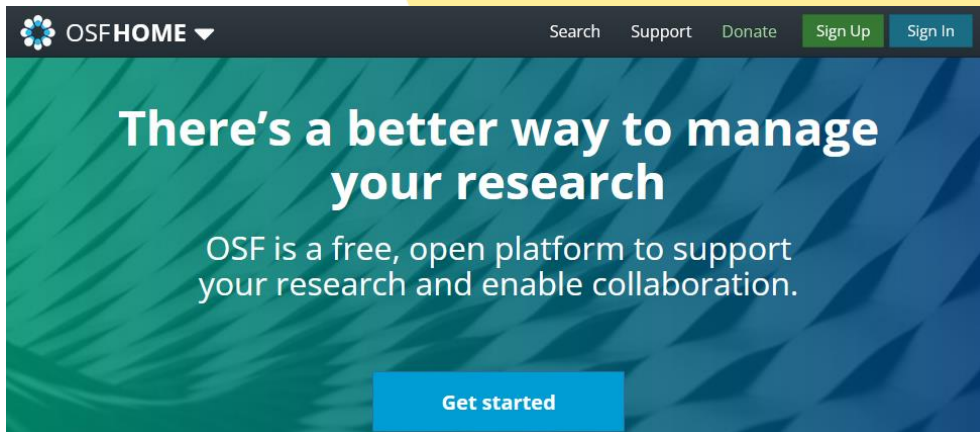
- RRs are considered to be a „vaccine" against publication bias
- Reason: incentives for QRPs are eradicated
- Caveat: they might be time-intensive due to multiple peer-review rounds (better plan them early, e.g., in your 1st PhD year)

Scheel et al., 2021, https://doi.org/10.1177/25152459211007467

But let's go one step back...

How to do a prereg?

- Choose a prereg form (from very basic to very comprehensive)

- Fill out a prereg form

- Upload it on a repository that offers prereg (e.g., OSF.io, ZPID)
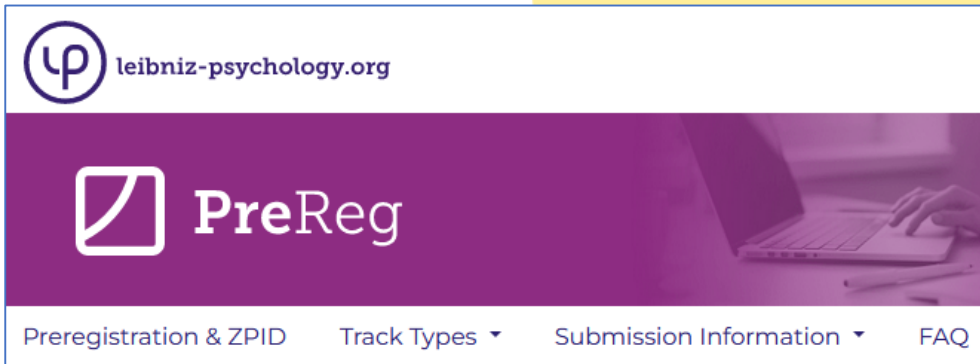
Repositories

- The OSF is most suitable for preregs! (osf.io)
- The ZPID offers a sufficient alternative and has better data protection policies (https://prereg-psych.org/index.php/rrp)
- On other repositories (e.g., Zenodo.org) there is only the option to upload a prereg in form of a preprint

Let's check out prereg forms

- OSF Prereg: https://osf.io/zab38/wiki/home/
- Beginners prereg: AsPredicted form (do it on the OSF)

- PRP Quant: https://prereg-psych.org/index.php/rrp/templates

Take away

- Open data and materials is important for reproducibility and replicability

- Choose a repository and start preparing your data (maybe from a smaller project)

- If you set up a new study, try to write a prereg before data collection (e.g., a small AsPredicted)

June 30

OPEN SCIENCE

HS 02.01

Spritzer-stand

Free entrance!

DAY
OF THE INSTITUTE
OF PSYCHOLOGY

Anne Scheel from VU Amsterdam & CWTS Leiden

| Time: | 11.45 – 13.30 |
| Location: | HS 02.01 |

Register for a seat and a free lunch bag until June 27: **tinyurl.com/OSDreg22**

Welcome address by Prof. Roland Grabner; followed by a short panel discussion

Keynote by Anne M. Scheel, MSc.

A young researcher's guide to the credibility revolution

The replication crisis has fundamentally shaken psychology. Its primary victim were not studies that failed to replicate — far worse, it was our trust in the rigor of standard research practice. This talk reviews what went wrong, how transparent and reproducible practices address the problem, and how early-career researchers can navigate the fault line between old incentives and new standards.

tinyurl.com/ OSDreg22

June 30

OPEN SCIENCE

Zach

DAY
OF THE INSTITUTE
OF PSYCHOLOGY

Clemens Brunner from the Institute of Psychology

| Time: | 13.45 – 15.15 |
| Location: | PC room Zach |

Workshop hosted by Clemens Brunner, Dr. techn.

Getting started with R & RStudio

June 30

OPEN SCIENCE

Zach

DAY
OF THE INSTITUTE
OF PSYCHOLOGY

Hilmar Brohmer from the Institute of Psychology

| Time: | 15.30 – 17.00 |
| Location: | PC room Zach |

Workshop hosted by Hilmar Brohmer, Dr. rer.nat.

Preregistration for Beginners

Thanks!