



GEORG-AUGUST-UNIVERSITÄT  
GÖTTINGEN



# Open Science

vs.

# Bad Science

the replication crisis  
and possible reforms

[ruben.arslan@gmail.com](mailto:ruben.arslan@gmail.com)

 @\_r\_c\_a

blog: <http://the100.ci>

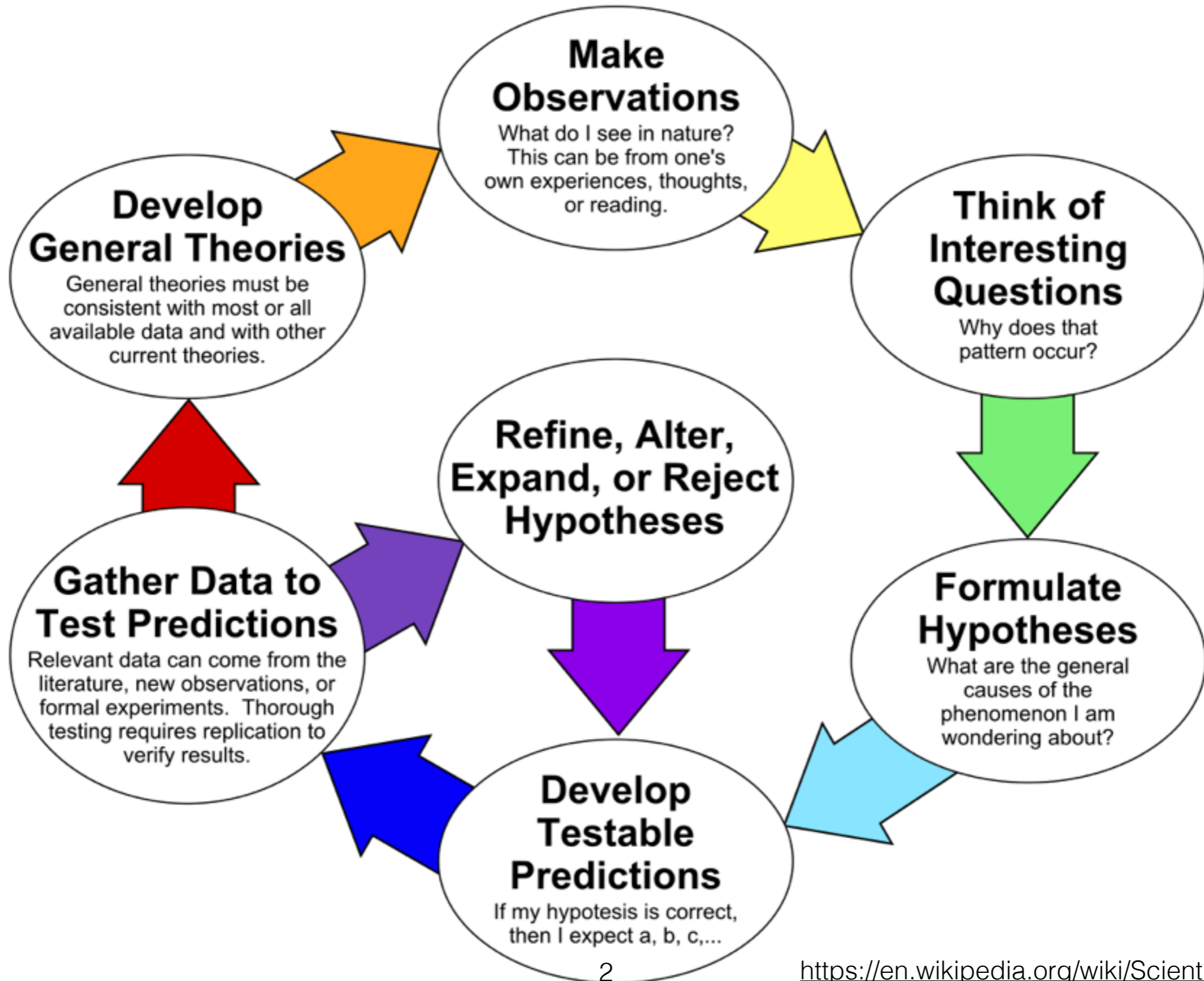


F E L L O W  
P R O G R A M M  
F R E I E S  
W I S S E N

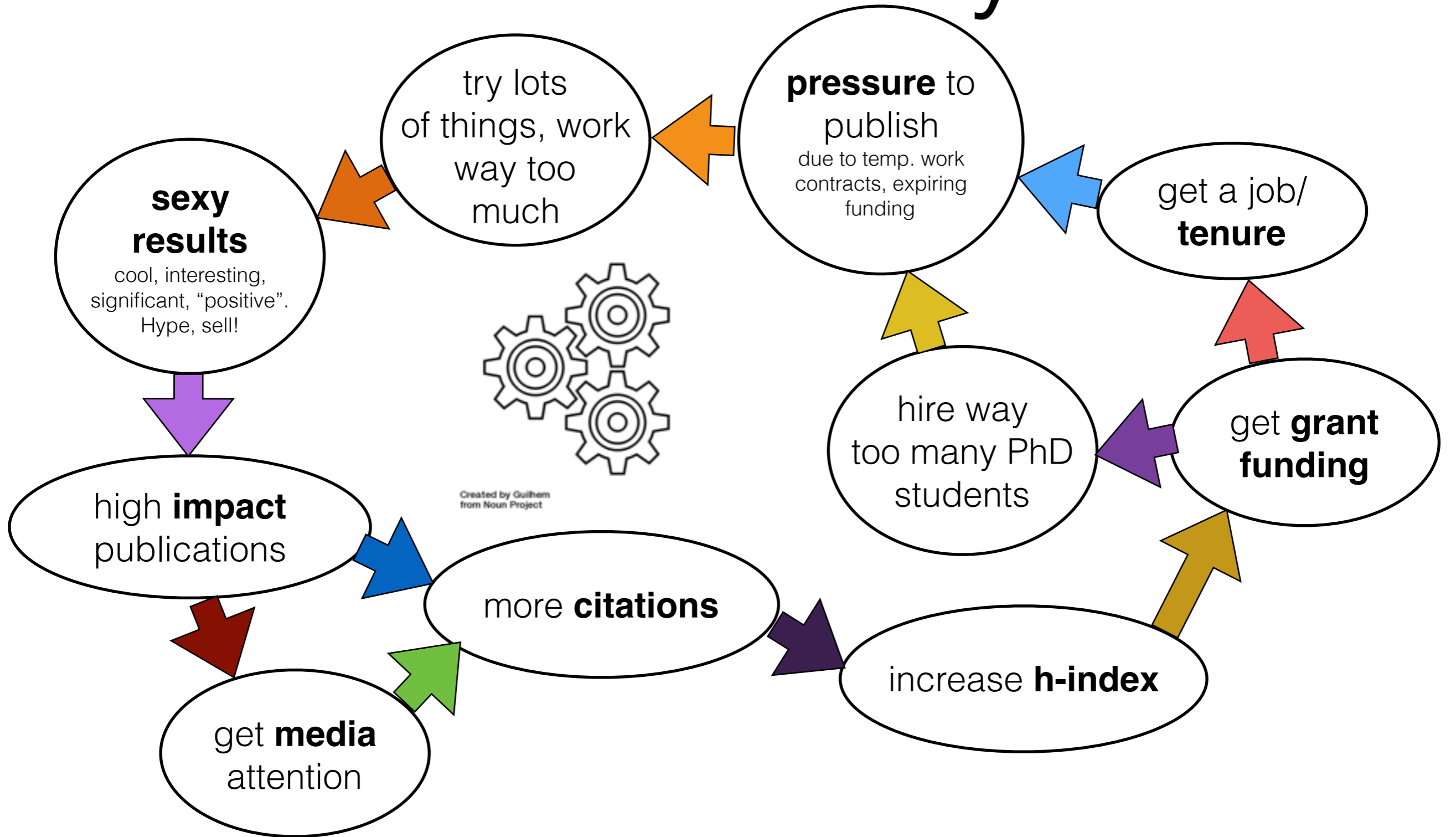


STIFTERVERBAND  
Bildung. Wissenschaft. Innovation.

# The Scientific Method as an Ongoing Process



# Science as a system

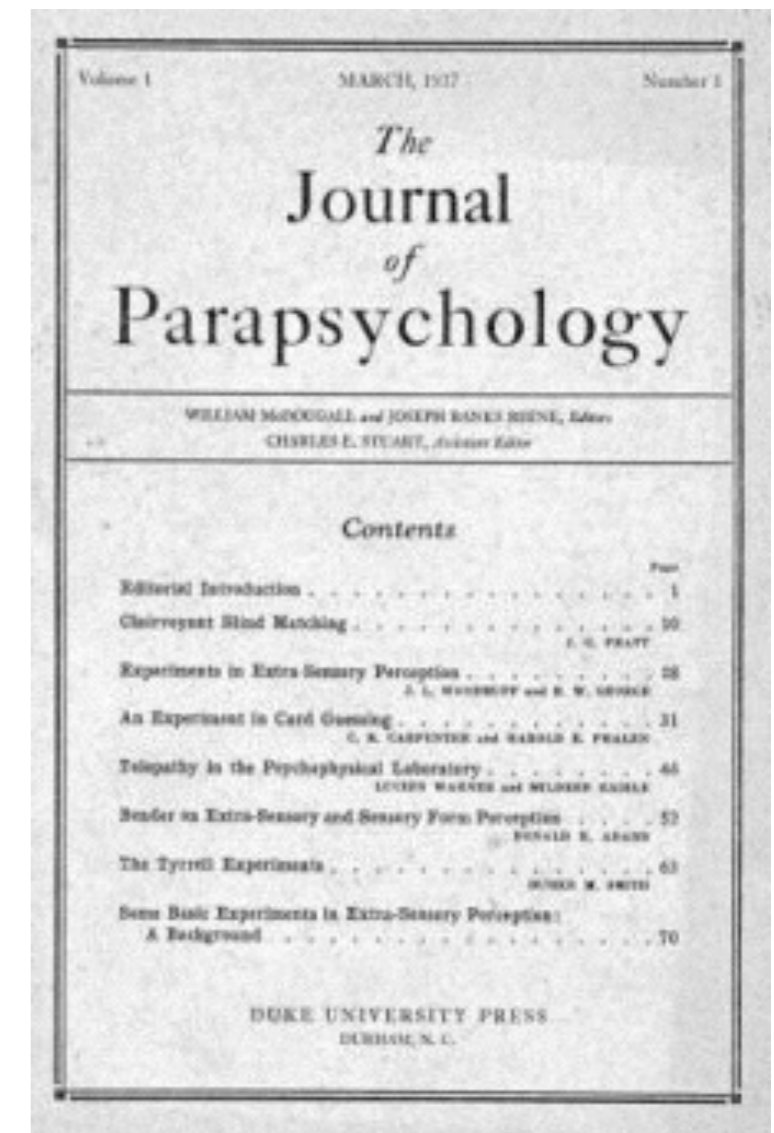
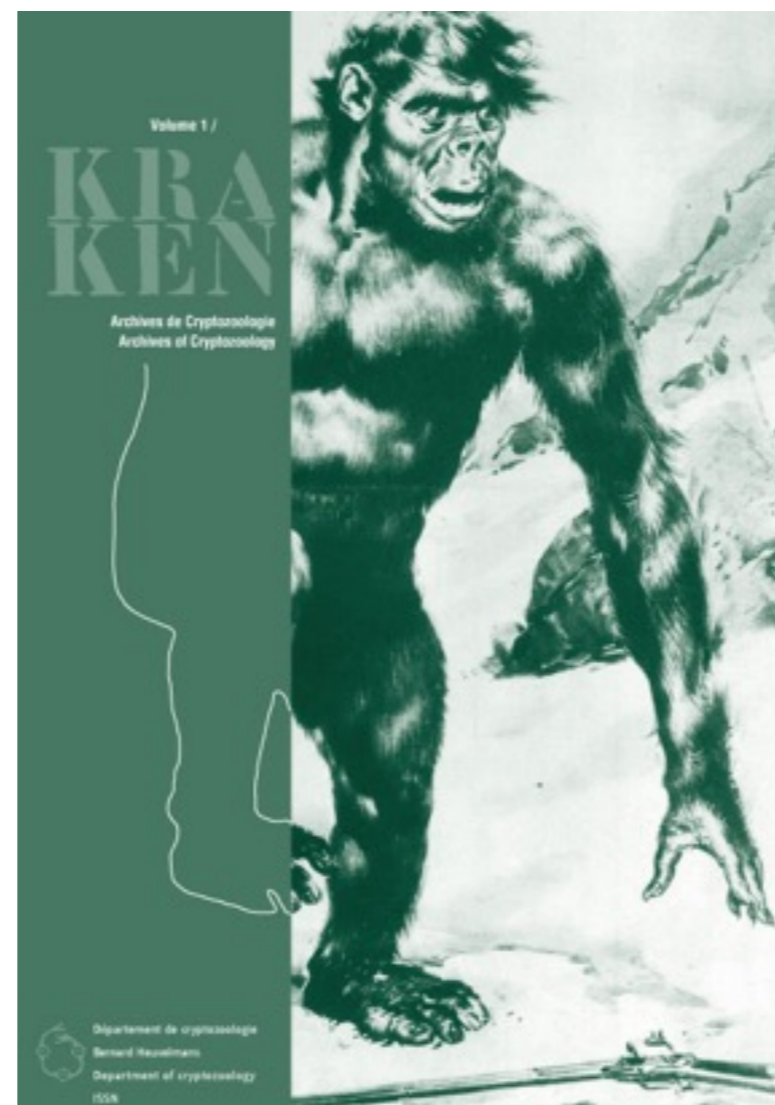
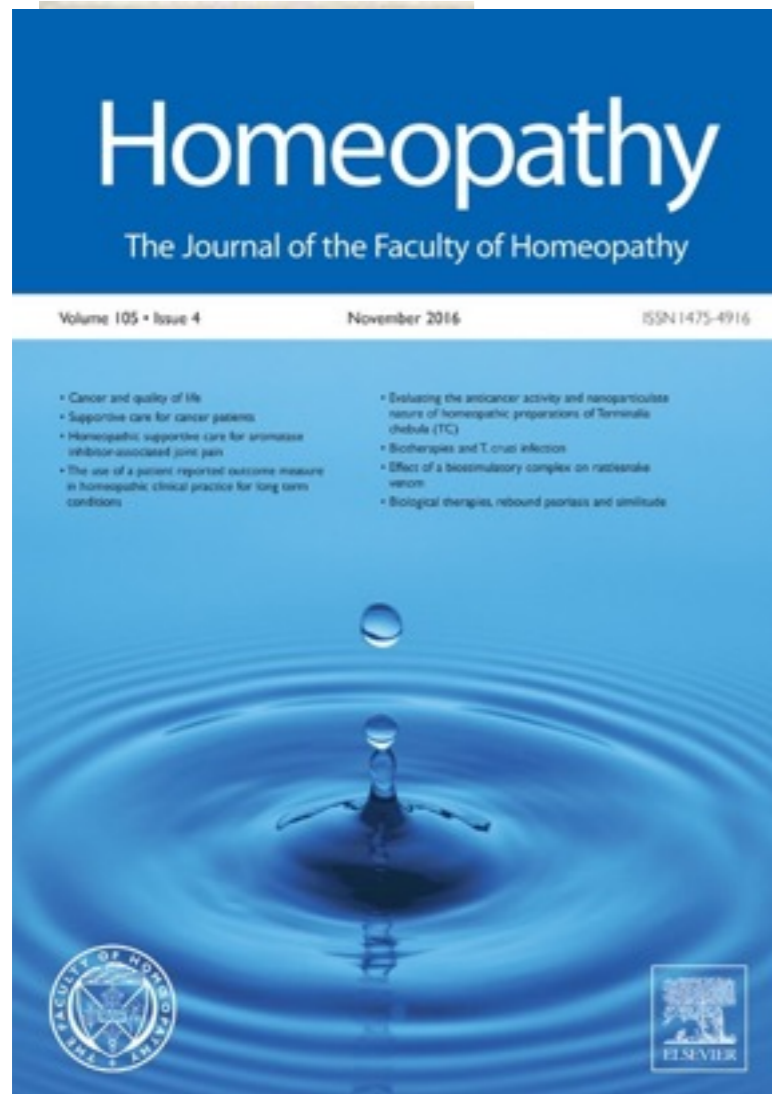


Can we use the *scientific method* to check how well **Science** works?

# A control group for **Science**

“ What we really need is a **control group for science** – people who will behave exactly like scientists, doing experiments, publishing journals, and so on, but whose field of study is completely empty: one in which the null hypothesis is always true.

# A control group for Science



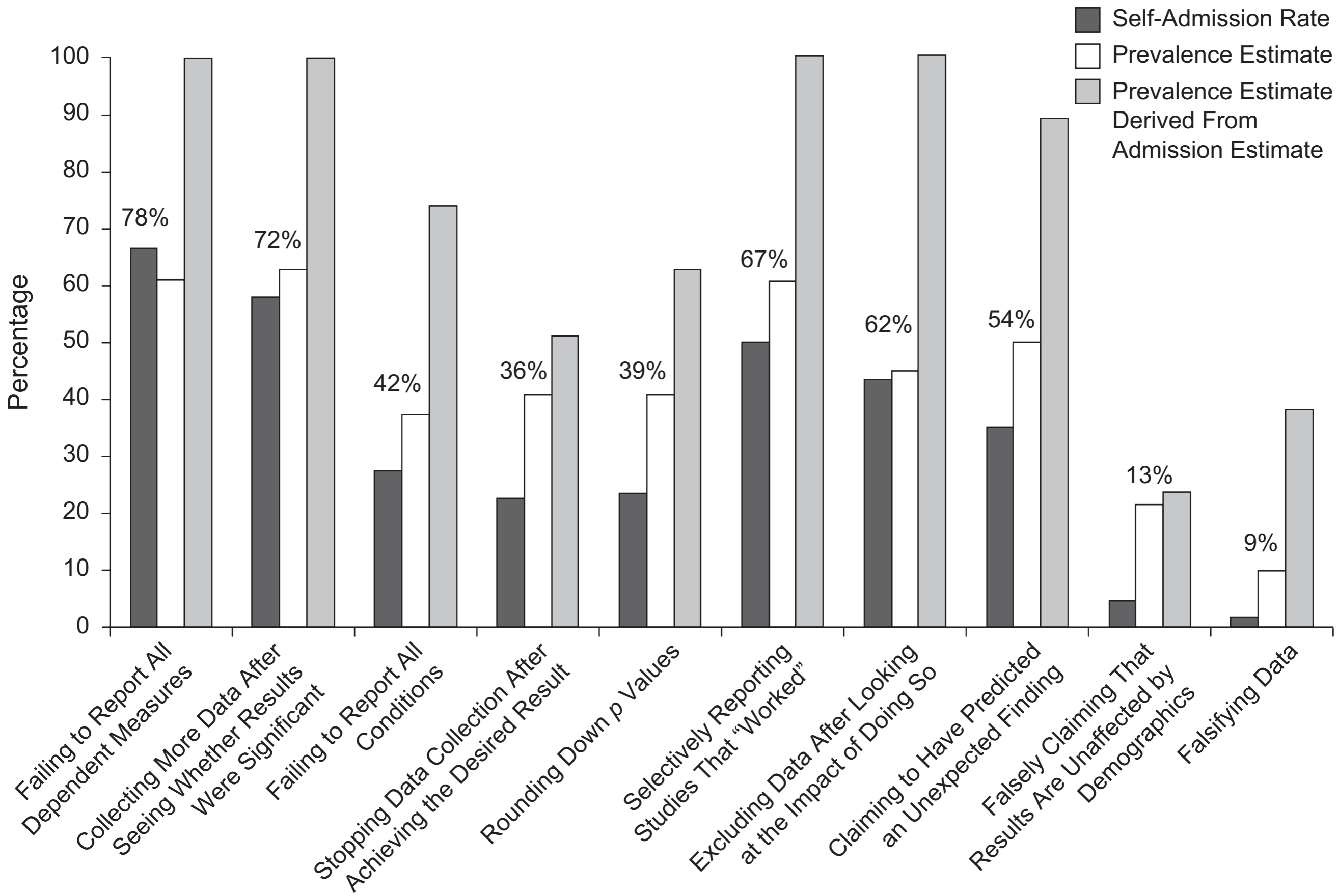
Quote [http://lesswrong.com/lw/1ib/parapsychology\\_the\\_control\\_group\\_for\\_science/](http://lesswrong.com/lw/1ib/parapsychology_the_control_group_for_science/)



# A control group for Science

1. Nothing is pre-registered, extremely flexible study design and lax methods ensure statistical significance. Researcher is usually not intentionally committing fraud.
2. **p < .05** in several studies
3. Top journal publishes study with senior author and attention-grabbing result
4. Wild media attention (but closed-access, so most people can't read paper)
5. Top journal refuses to publish non-replication, two middling journals reject it (once after inviting Bem as reviewer)
6. Unusually: Many statistical and methodological criticisms of Bem are blogged and published
7. Unusually: media attention for the non-replications
8. After all this criticism, Bem publishes a meta-analysis of 90 replications in 33 laboratories in 14 countries.  $p < 1.2 * 10^{-10}$





“**Everyone knew it was wrong**, but they thought it was wrong the way it’s wrong to **jaywalk**. We decided to write 'False-Positive Psychology' when simulations revealed it was wrong the way it’s wrong to **rob a bank**.

(Simmons, Nelson, & Simonsohn, 2017)

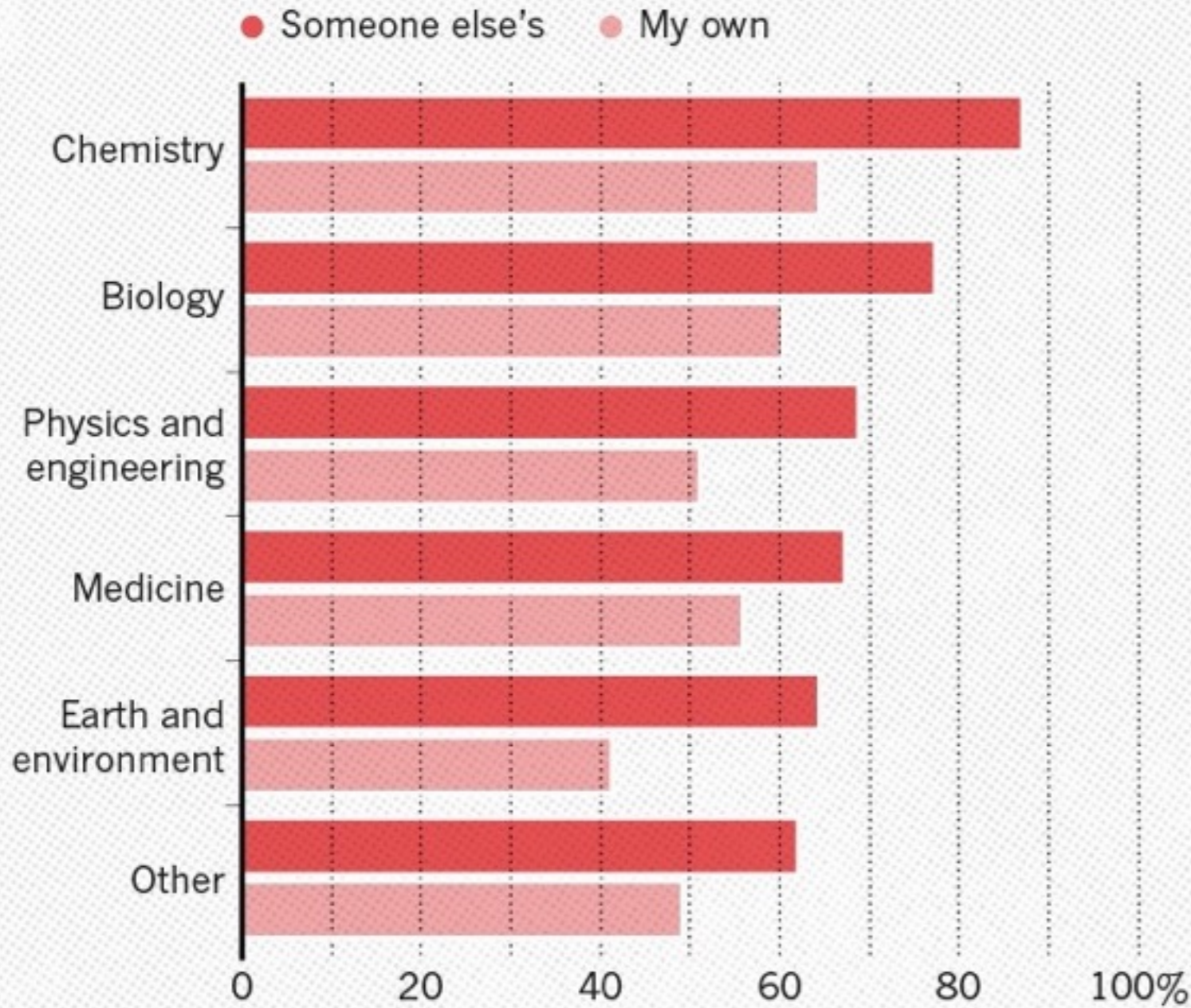
# False-positive psychology

**Table 1.** Likelihood of Obtaining a False-Positive Result

Researcher degrees of freedom	Significance level		
	$p < .1$	$p < .05$	$p < .01$
Situation A: two dependent variables ( $r = .50$ )	17.8%	9.5%	2.2%
Situation B: addition of 10 more observations per cell	14.5%	7.7%	1.6%
Situation C: controlling for gender or interaction of gender with treatment	21.6%	11.7%	2.7%
Situation D: dropping (or not dropping) one of three conditions	23.2%	12.6%	2.8%
Combine Situations A and B	26.0%	14.4%	3.3%
Combine Situations A, B, and C	50.9%	30.9%	8.4%
Combine Situations A, B, C, and D	81.5%	60.7%	21.5%

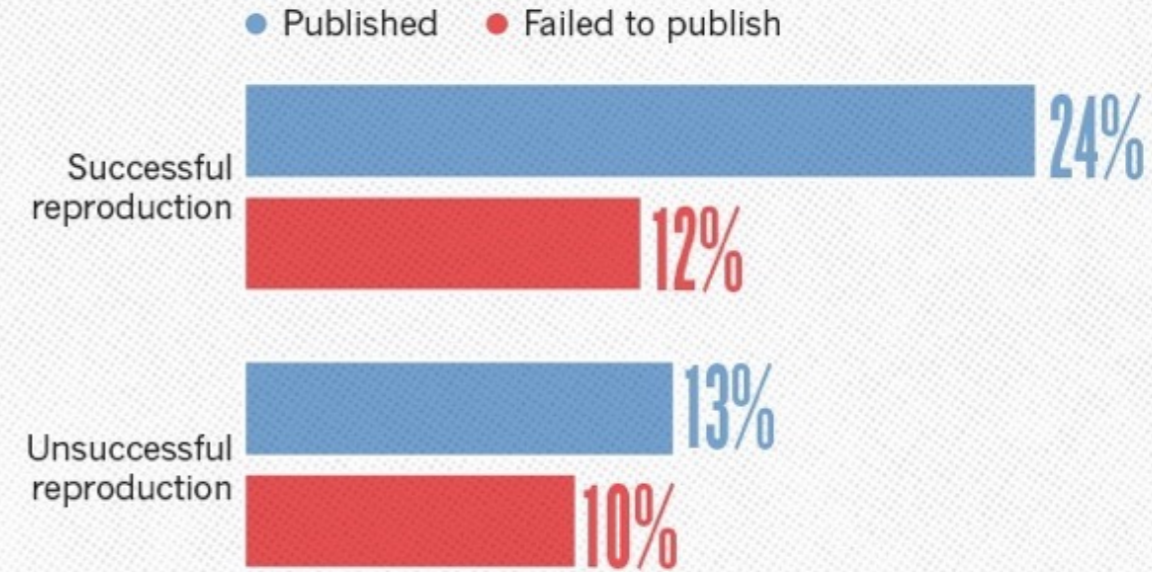
# HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

Most scientists have experienced failure to reproduce results.



# HAVE YOU EVER TRIED TO PUBLISH A REPRODUCTION ATTEMPT?

Although only a small proportion of respondents tried to publish replication attempts, many had their papers accepted.



Number of respondents from each discipline:  
 Biology **703**, Chemistry **106**, Earth and environmental **95**,  
 Medicine **203**, Physics and engineering **236**, Other **233**

©nature

# Other fields haven't checked themselves so systematically

- only 49% out of 59 papers in **economics** could be reproduced *with the same data & with help from the authors*
- 11% (6 out of 53) landmark studies in **preclinical cancer** research replicated
- 45.2% of 25,927 finished pre-registered **clinical trials** are missing results (buried)
- **Neuroscience** has a median statistical power of 18%

Begley & Ellis 2012

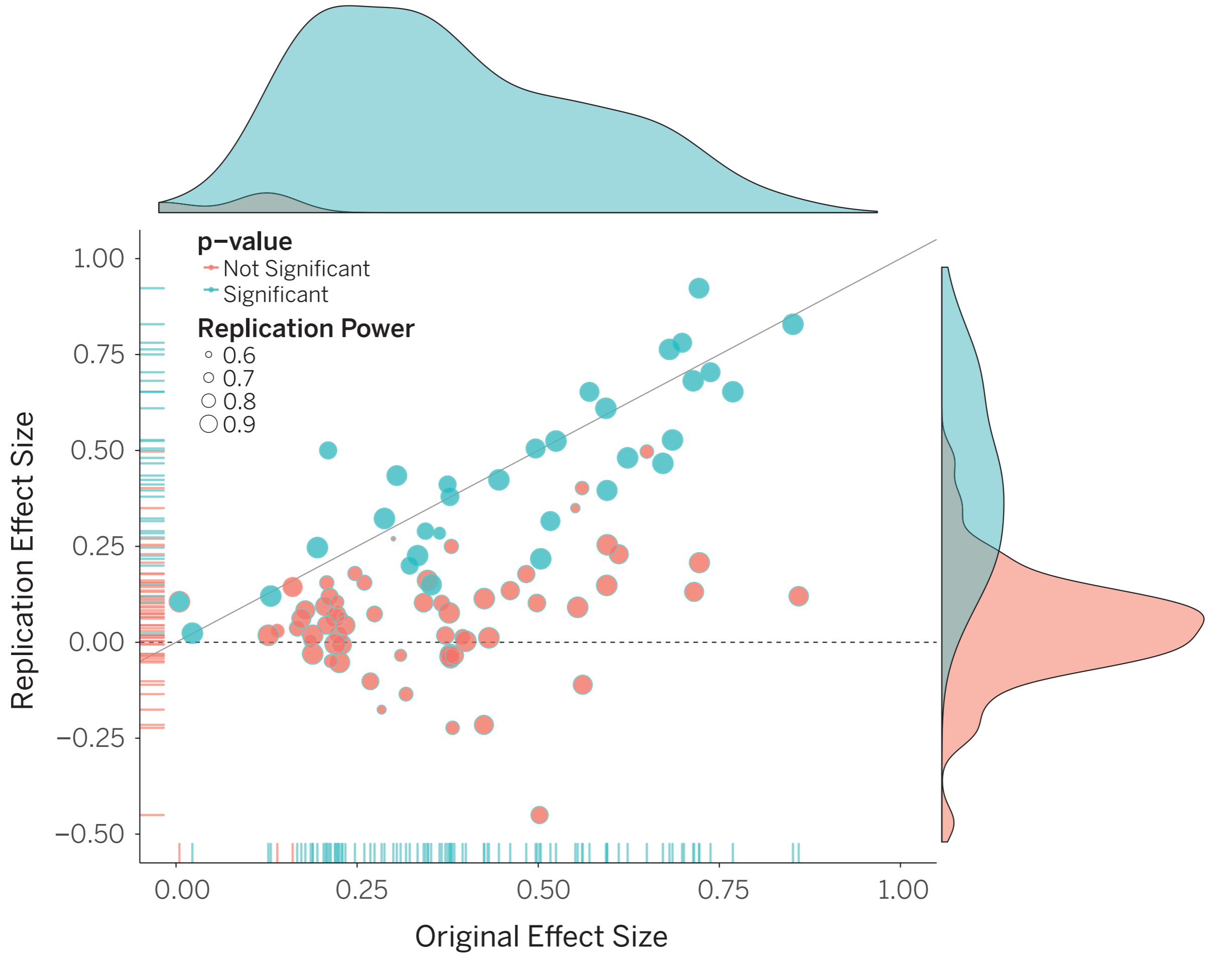
Prinz et al. 2011

<https://trialstracker.ebmdatalab.net/>

Button et al. 2013

# In my studies I was taught

- “ego depletion” is a real thing
- elderly priming makes you walk slow
- power posing increases testosterone
- facial expressions affect emotions
- judges are less lenient when hungry
- fish smell makes you suspicious
- ovulation affects voting







# Transparency as a disinfectant



Created by Tomas Knopp  
from Noun Project

“Nullius in verba”, Take nobody’s word for it  
— motto of the Royal Society

# What can we do?

- A lot!
  - as cogs in the system/incentive structure
  - as colleagues
  - as authors
  - as peer reviewers
  - as editors

# Incentive structure

- Transparency is not its own reward
- It takes time, resources and it can turn exciting results boring
- But: People will trust you more.
- And: Think of it as future-proofing your research. Times are changing. Be ready.

# As colleagues

- Form local open science initiatives (NOSI)
- Join global initiatives: Society for the Improvement of Psychological Science



**SOCIETY FOR THE  
IMPROVEMENT OF  
PSYCHOLOGICAL SCIENCE**

 View

Wiki Version: (Current) Tobias Rothmund: 2017-

## Wiki Pages

e

macht bei meiner Uni mit?

## Recent Wiki Pages

## Bestehende Open-Science-Initiativen

Chronologisch nach Gründungsdatum, mit den jeweiligen Ansprechpartnern:

- Ludwig-Maximilians-Universität München (Felix Schönbrodt [felix@nicebread.de](mailto:felix@nicebread.de))
- Universität Koblenz-Landau (Tobias Rothmund [rothmund@uni-landau.de](mailto:rothmund@uni-landau.de))
- FernUniversität Hagen (Andreas Glöckner [andreas.gloeckner@fernuni-hagen.de](mailto:andreas.gloeckner@fernuni-hagen.de))
- Universität Münster (Mitja Back [mitja.back@wwu.de](mailto:mitja.back@wwu.de))
- Universität Göttingen (Ruben Arslan [ruben.arslan@gmail.com](mailto:ruben.arslan@gmail.com))
- Universität Wien (Martin Voracek [martin.voracek@univie.ac.at](mailto:martin.voracek@univie.ac.at))
- Humboldt-Universität Berlin (Kai Horstmann [horstmak@cms.hu-berlin.de](mailto:horstmak@cms.hu-berlin.de))

## Emailliste

Zur Kommunikation untereinander nutzen wir eine Emailliste bei Google Groups. Alle Interessierten sind herzlich eingeladen!  
<https://groups.google.com/d/forum/nosi-de>

### Weitere Links

Nature Artikel zur [Open Science Initiative im Bereich NeuroScience an der McGill University, Kanada](#)

# As authors



# As authors

## The Preregistration Challenge

Preregistration is an unfamiliar practice in many areas of science. We will give 1,000 awards of \$1,000 each to promote education and experience with preregistration. Award winners will have completed a preregistration using the Open Science Framework, a free scholarly commons, and published the results in an eligible journal. Read through the [steps to earn a prize](#).



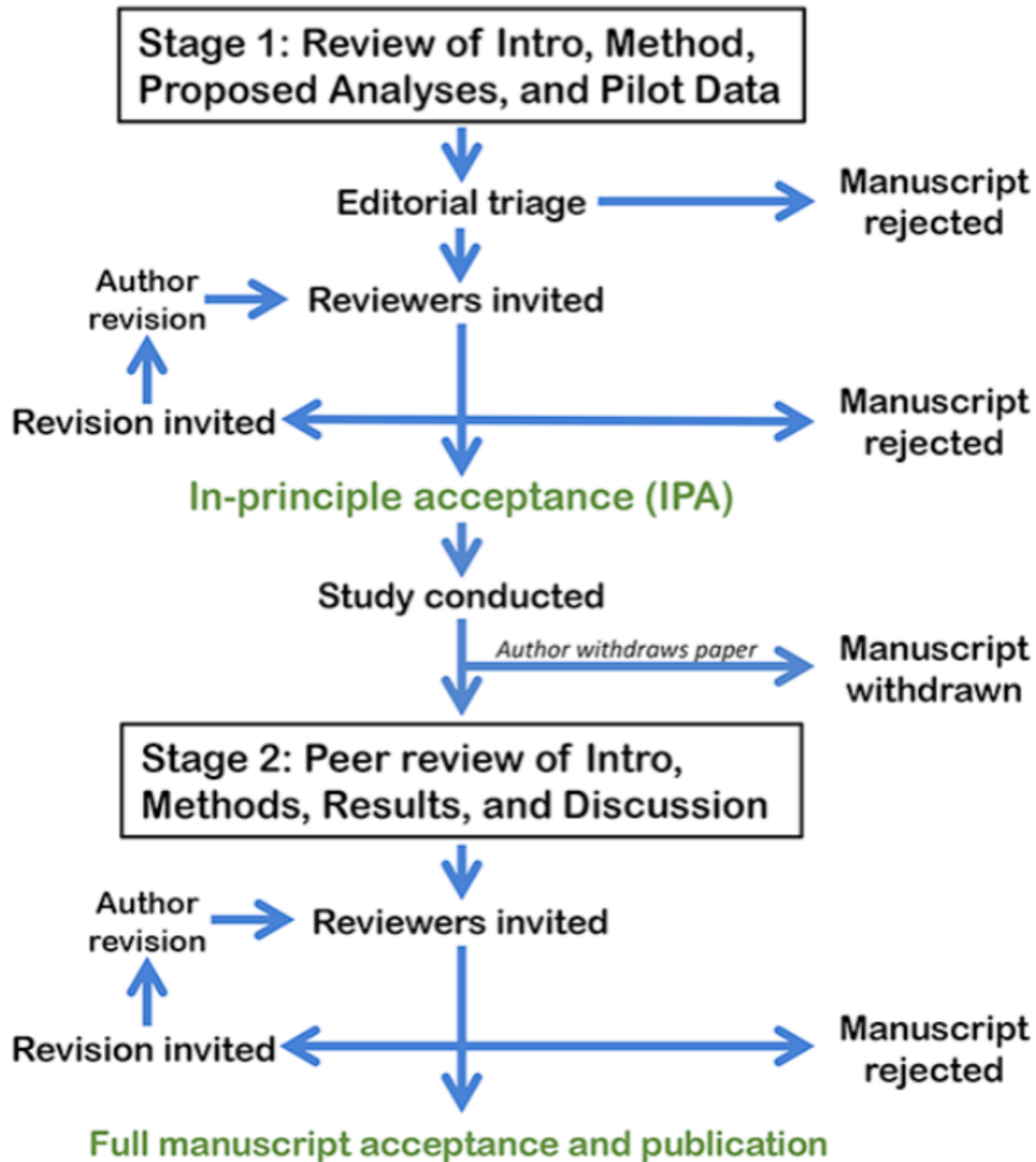
# Registered Reports



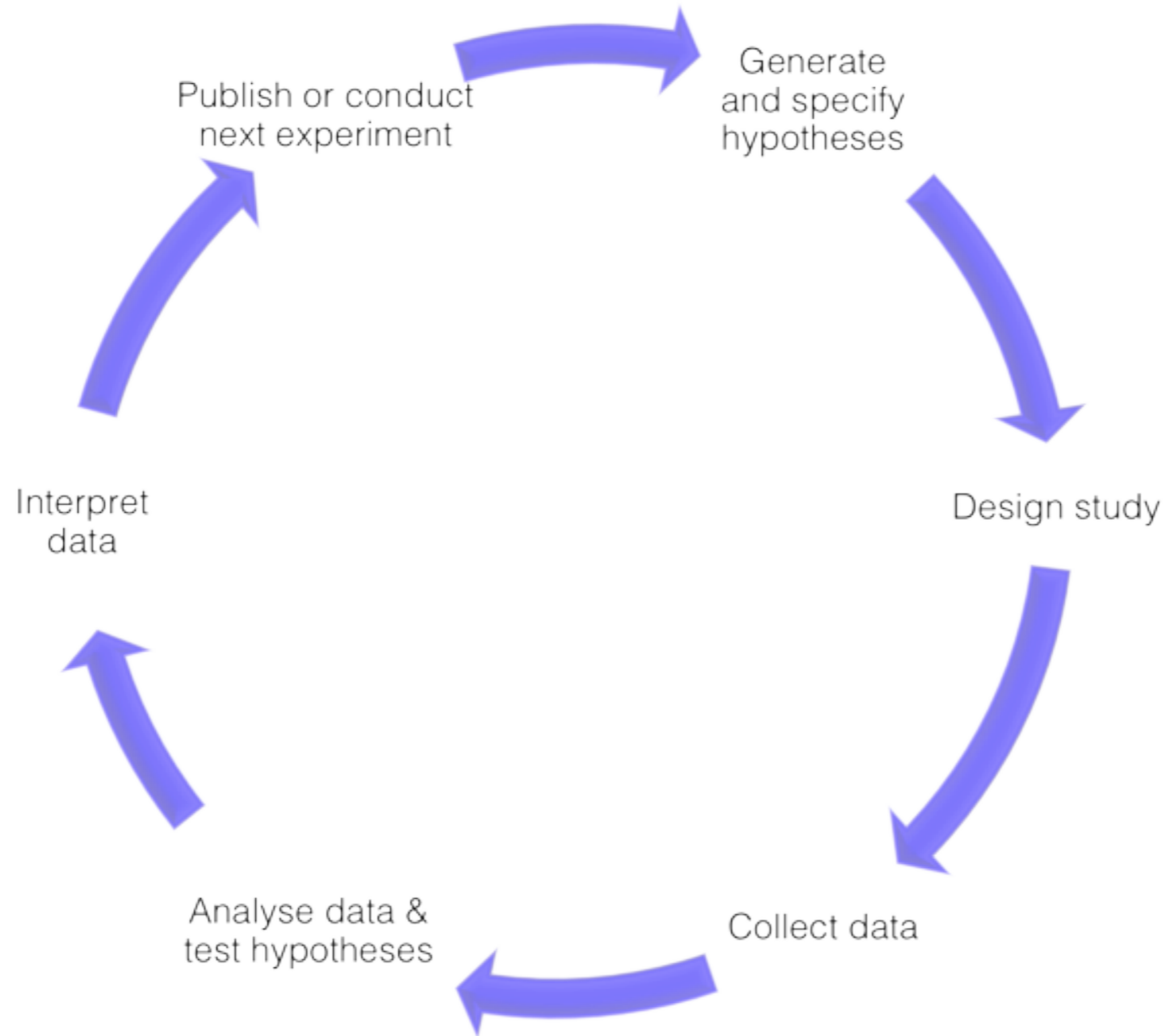
**52 journals and counting**, e.g. *Cortex*, *Nature Human Behaviour*, *Royal Society Open Science*

& talks with funders/dataset managers

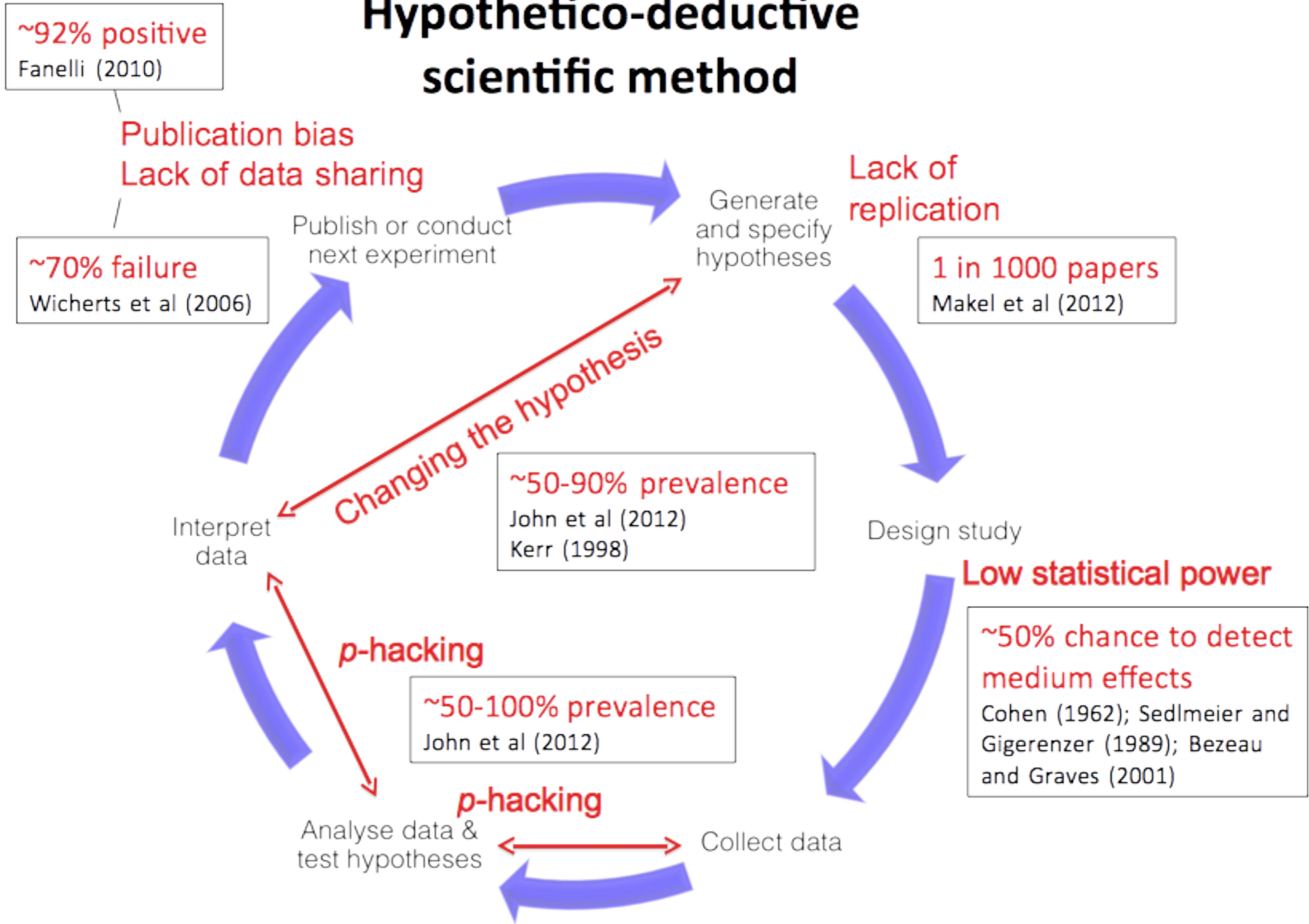


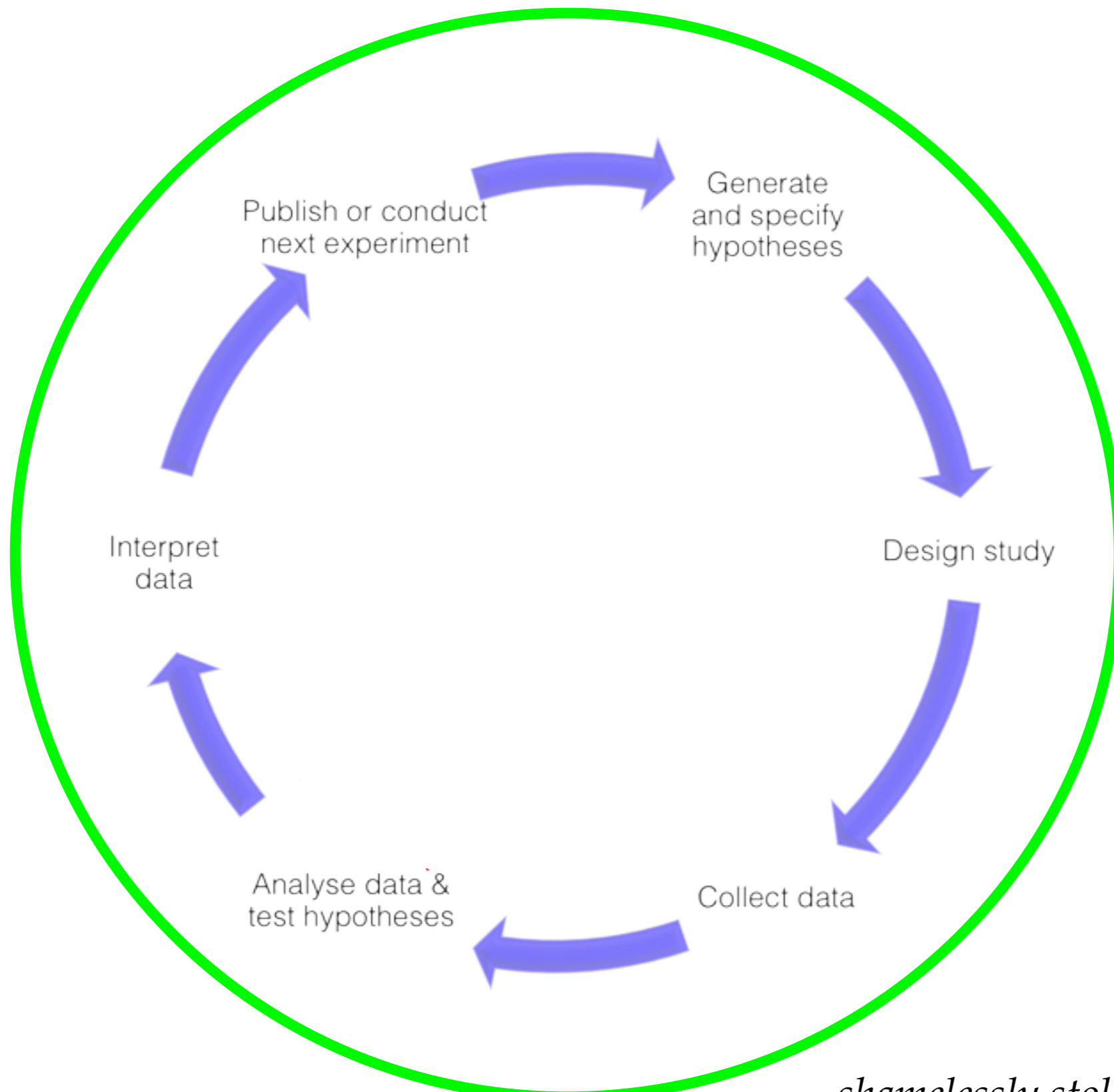


# Hypothetico-deductive scientific method



# Hypothetico-deductive scientific method





# As peer reviewers



- require transparency, reproducibility from authors (or a justification why not in the paper)
- withhold review until this happens
- 446 signatories

statcheck

# As peer reviewers

- perform post-publication peer review



**PUBPEER**  
The online journal club

**COMMONS**

PubMed



A forum for scientific discourse

- can stay anonymous
- to comment non-anonymously requires a publication
- includes Pubmed Commons

- with your good name attached
- visible in Pubmed
- requires a publication in Pubmed

# As colleagues

- Join the conversation or just “lurk”
  - on Twitter
  - on Facebook  
(Psychological Methods and Discussion Group,  
PsychMAP)
  - on the many, many blogs and podcasts

# As editors

- Implement Registered Reports
- Adopt TOP (>2900 journals and orgs signed up)

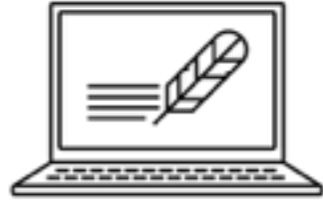
Eight Standards at Three Levels of Increasing Rigor

**Standards:** Data Citation | Data, Materials, and Code Transparency | Design and Analysis | Preregistration | Replication

**Levels:** Disclose, Require, or Verify

[cos.io/top](https://cos.io/top)





Created by dlayorganci  
from Noun Project

# Blog Roll

- [andrewgelman.com](http://andrewgelman.com) - Columbia Bayesianischer Statistiker, schreibt fast jeden Tag, oft über die Krise
- [sometimesimwrong.typepad.com](http://sometimesimwrong.typepad.com) - Simine Vazire - Psychologin im Executive Committee der Society for Improving Psychological Science
- [slatestarcodex.com](http://slatestarcodex.com) - Scott Alexander, über ihn unter anderem die Betrachtung der Wissenschaft mit Placebo-Kontrollgruppe
- [neuroskeptic.com](http://neuroskeptic.com) - pseudonymer Blogger, nicht nur Neurowissenschaft, hat viele der Replikationskrisenthemen schon früh angeregt
- [deevybee.blogspot.com](http://deevybee.blogspot.com) - Dorothy Bishop, eminente Verhaltensgenetikerin, Entwicklungspsychologin
- <https://www.facebook.com/groups/853552931365745/> - Psychological Methods Discussion Group
- <https://www.facebook.com/groups/psychmap/> - PsychMAP wie oben "lite" (moderiert für Tonfall)
- <http://the100.ci> - The 100% Confidence Interval (Eigenwerbung)
- Podcasts: Everything Hertz, The Black Goat

# Remaining problems

- Citation/attention bias
- “Natural selection” of Bad Science (Smaldino & McElreath, 2016)
- Peer review is unreliable, reviewers are overwhelmed
- Lack of validity
- Fraud
- How not to overfit to data when preregistration impossible

# Thank you!

blog: <http://the100.ci>



link list: [tiny.cc/ope](http://tiny.cc/ope)

Ruben C. Arslan



@\_r\_c\_a

[rubenarslan@gmail.com](mailto:rubenarslan@gmail.com)