Open science: Replicability, transparency, p-hacking

David Saldaña
@DavidSaldana_es
@DavidSaldana_en

Trondheim, 28th August 2019

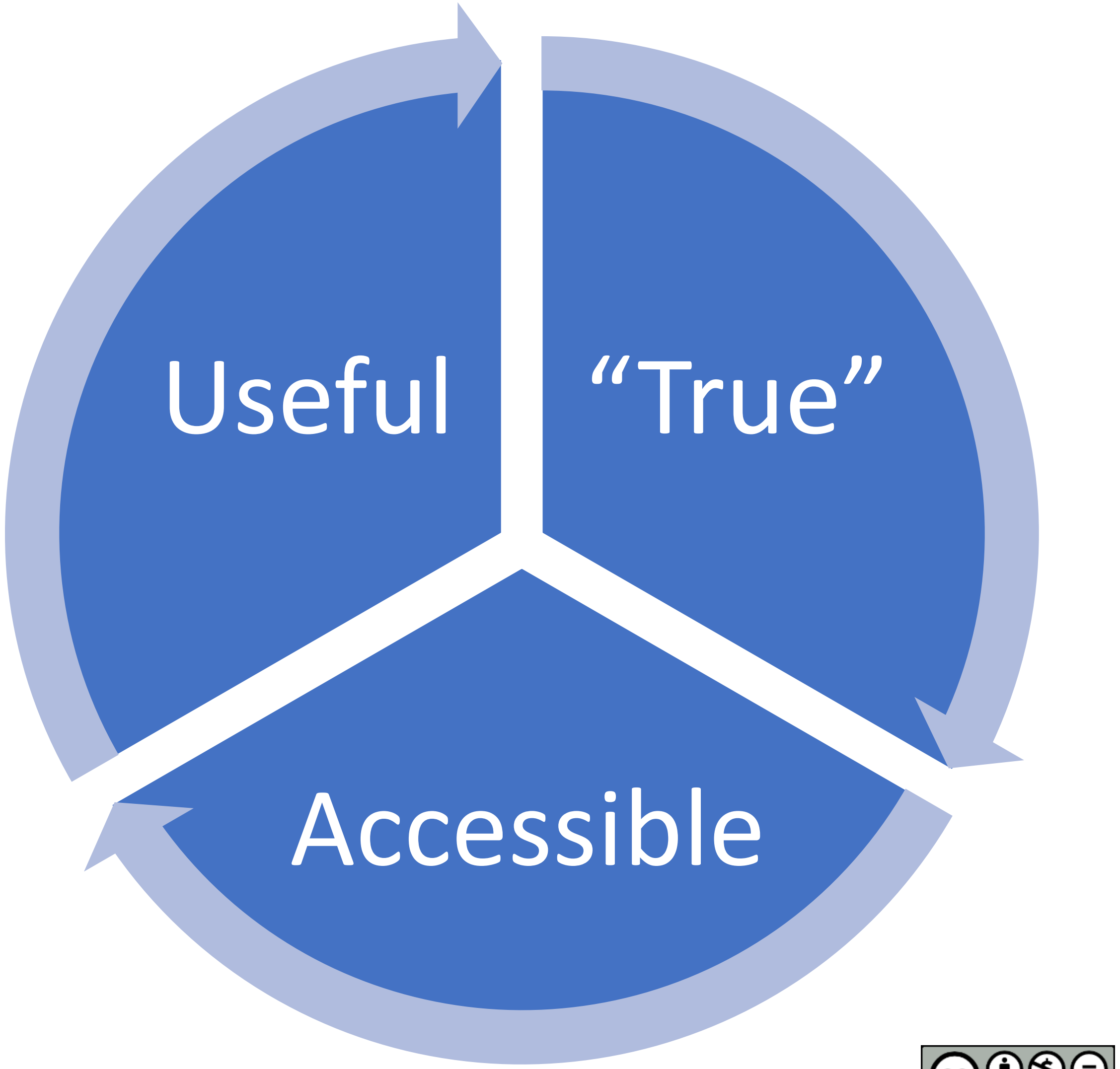
UNIVERSIDAD DE SEVILLA

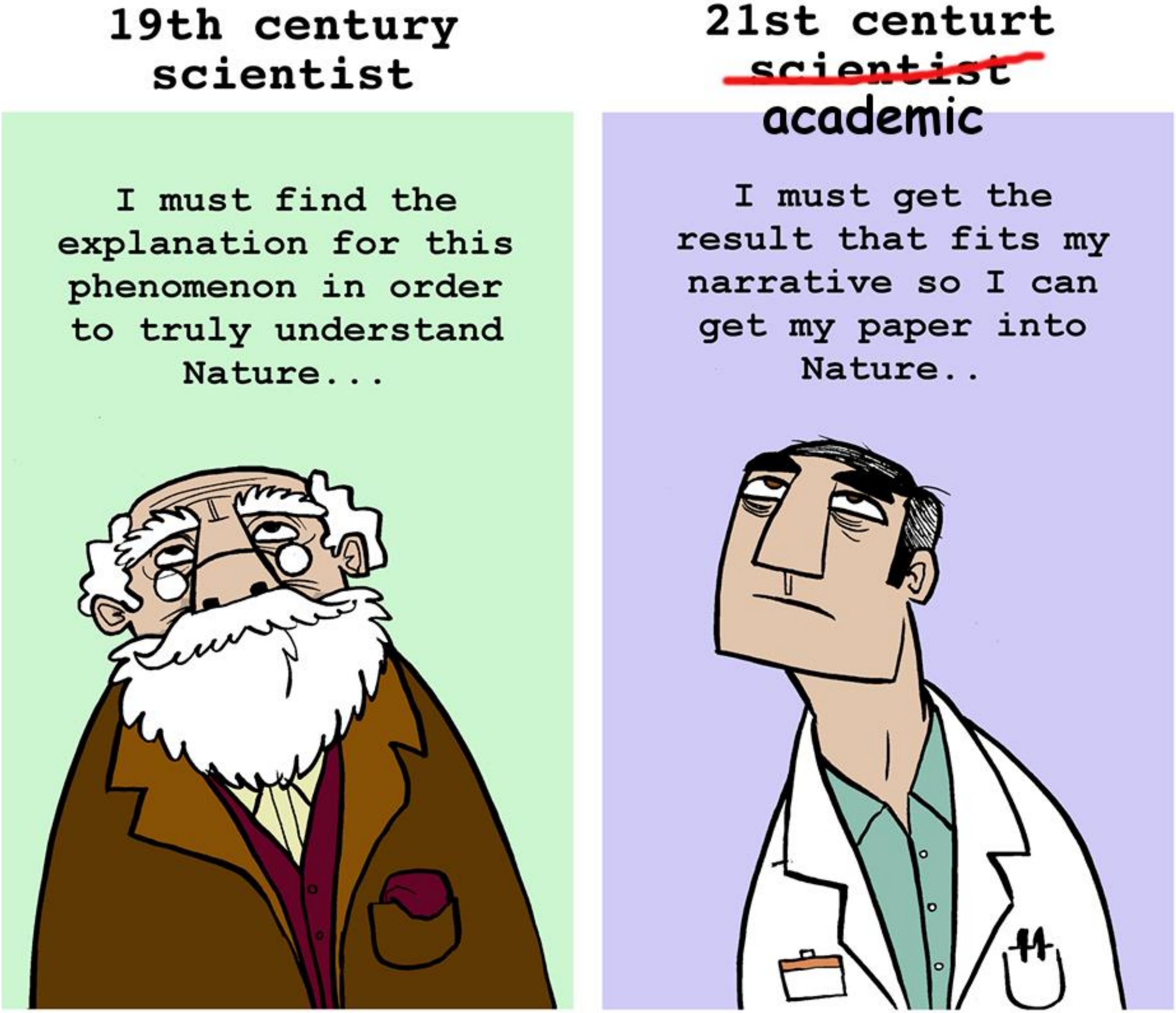
Laboratorio de Diversidad, Cognición y Lenguaje

Facultad de Psicología
UNIVERSIDAD Ð SEVILLA

What should we be aiming for?

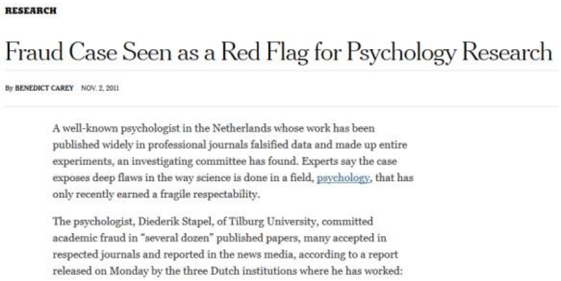
Valid and useful knowledge that is available to those who may need/want to use it




19th century scientist
I must find the explanation for this phenomenon in order to truly understand Nature...
21st centurt ~~scientist~~ academic
I must get the result that fits my narrative so I can get my paper into Nature..

facebook.com/pedromics

https://pandelisperakakis.info/wp-content/uploads/2017/05/scientist_vs_academic.png



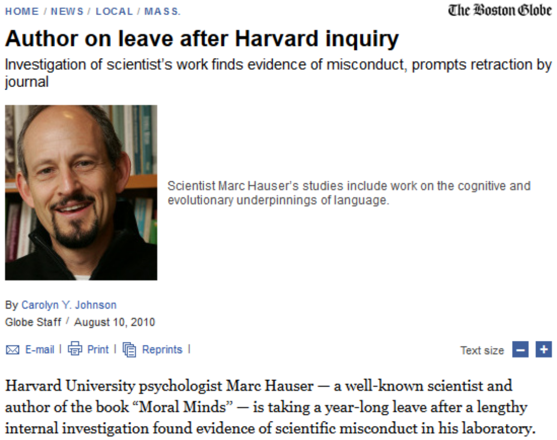
RESEARCH

Fraud Case Seen as a Red Flag for Psychology Research

By BENEDICT CAREY NOV. 2, 2011

A well-known psychologist in the Netherlands whose work has been published widely in professional journals falsified data and made up entire experiments, an investigating committee has found. Experts say the case exposes deep flaws in the way science is done in a field, psychology, that has only recently earned a fragile respectability.

The psychologist, Diederik Stapel, of Tilburg University, committed academic fraud in "several dozen" published papers, many accepted in respected journals and reported in the news media, according to a report released on Monday by the three Dutch institutions where he has worked:



HOME / NEWS / LOCAL / MASS.

The Boston Globe

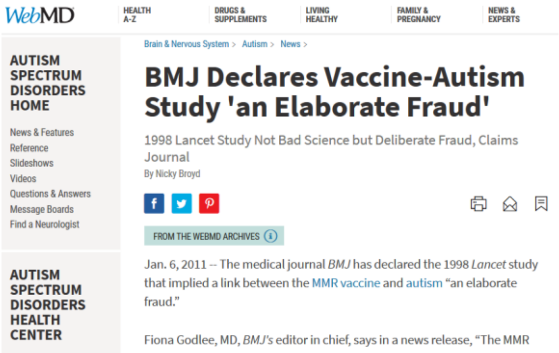
Author on leave after Harvard inquiry

Investigation of scientist's work finds evidence of misconduct, prompts retraction by journal

Scientist Marc Hauser's studies include work on the cognitive and evolutionary underpinnings of language.

By Carolyn Y. Johnson
Globe Staff / August 10, 2010

Harvard University psychologist Marc Hauser — a well-known scientist and author of the book "Moral Minds" — is taking a year-long leave after a lengthy internal investigation found evidence of scientific misconduct in his laboratory.



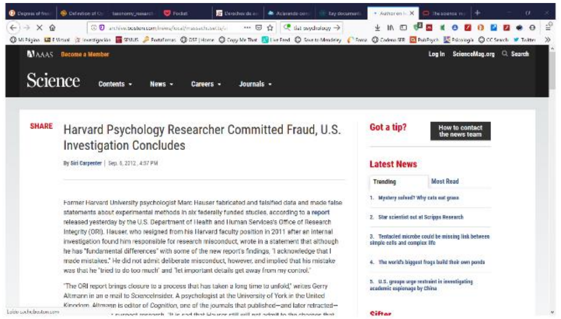
BMJ Declares Vaccine-Autism Study 'an Elaborate Fraud'

1998 Lancet Study Not Bad Science but Deliberate Fraud, Claims Journal

By Nicky Broyd

Jan. 6, 2011 -- The medical journal *BMJ* has declared the 1998 *Lancet* study that implied a link between the MMR vaccine and autism "an elaborate fraud."

Fiona Godlee, MD, *BMJ's* editor in chief, says in a news release, "The MMR



AAAS Become a Member
Log In ScienceMag.org Search
Science
Contents
News
Careers
Journals
SHARE
Got a tip?
How to contact the news team
Latest News
Trending
Most Read
1. Mystery solved? Why cats eat grass
2. Star scientist out at Scripps Research
3. Tentacled microbe could be missing link between simple cells and complex life
4. The world's biggest frogs build their own ponds
5. U.S. groups urge restraint in investigating academic espionage by China
Sifter

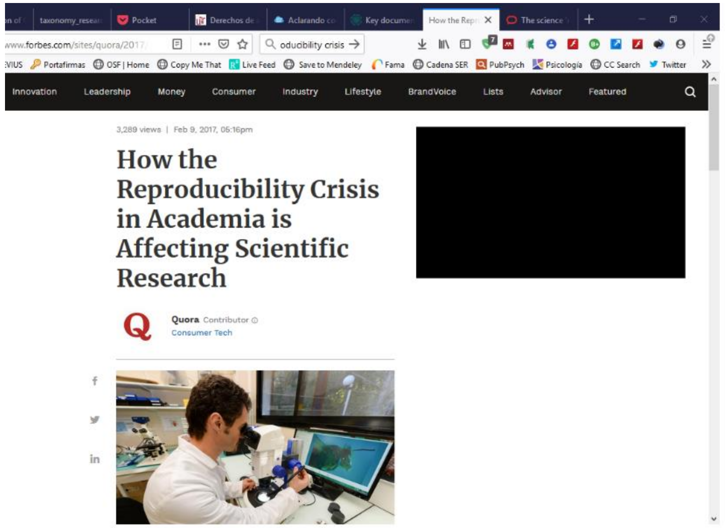
Harvard Psychology Researcher Committed Fraud, U.S. Investigation Concludes

By Siri Carpenter | Sep. 6, 2012 , 4:57 PM

Former Harvard University psychologist Marc Hauser fabricated and falsified data and made false statements about experimental methods in six federally funded studies, according to a report released yesterday by the U.S. Department of Health and Human Services's Office of Research Integrity (ORI). Hauser, who resigned from his Harvard faculty position in 2011 after an internal investigation found him responsible for research misconduct, wrote in a statement that although he has "fundamental differences" with some of the new report's findings, "I acknowledge that I made mistakes." He did not admit deliberate misconduct, however, and implied that his mistake was that he "tried to do too much" and "let important details get away from my control."

"The ORI report brings closure to a process that has taken a long time to unfold," writes Gerry Altmann in an e-mail to ScienceInsider. A psychologist at the University of York in the United Kingdom, Altmann is editor of *Cognition*, one of the journals that published—and later retracted—

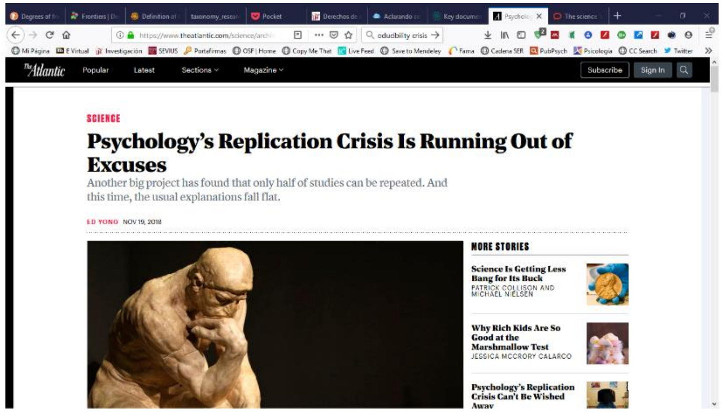
Threats to "truthfulness" of our results



3,289 views | Feb 9, 2017, 05:16pm

How the Reproducibility Crisis in Academia is Affecting Scientific Research

Quora Contributor
Consumer Tech



SCIENCE

Psychology's Replication Crisis Is Running Out of Excuses

Another big project has found that only half of studies can be repeated. And this time, the usual explanations fall flat.

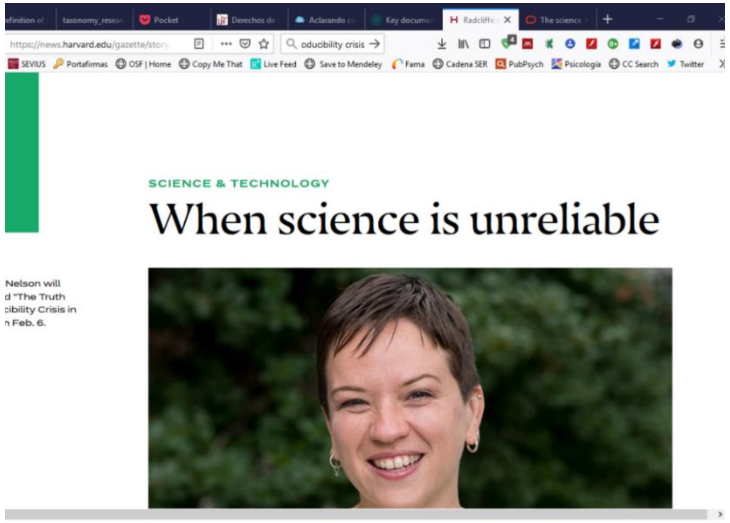
ED YONG NOV 19, 2018

MORE STORIES

Science Is Getting Less Bang for Its Buck
PATRICK COLLISON AND MICHAEL NIELSEN

Why Rich Kids Are So Good at the Marshmallow Test
JESSICA MCCRORY CALARCO

Psychology's Replication Crisis Can't Be Wished Away



SCIENCE & TECHNOLOGY

When science is unreliable

Nelson will
d "The Truth
cibility Crisis in
n Feb. 6.

Valid…means as “true” as possible

Some concepts

REPRODUCIBILITY:

Refers to the ability of a researcher to duplicate the results of a prior study using the same materials and procedures as were used by the original investigator

- E.g., a researcher uses the same raw data, builds same analysis files, and same statistical procedures to make sure that same results obtained as in published study
- Differences could be due to:

o Processing (e.g., treatment missing data) of data
o Application of statistical method (e.g., different defaults)
o Accidental errors in original analysis (or follow-up analysis)
o Reproducibility is a minimum necessary condition for a finding to be believable and informative.

Report of the Subcommittee on Replicability in Science of the SBE Advisory Committee to the National Science Foundation
13 May 2015 Presentation
at SBE AC Spring Meeting by K. Bollen

Some concepts

REPRODUCIBILITY:

Refers to the ability of a researcher to duplicate the results of a prior study using the same materials and procedures as were used by the original investigator

- E.g., a researcher uses the same raw data, builds same analysis files, and same statistical procedures to make sure that same results obtained as in published study
- Differences could be due to:

o Processing (e.g., treatment missing data) of data
o Application of statistical method (e.g., different defaults)
o Accidental errors in original analysis (or follow-up analysis)
o Reproducibility is a minimum necessary condition for a finding to be believable and informative.

Report of the Subcommittee on Replicability in Science of the SBE Advisory Committee to the National Science Foundation
13 May 2015 Presentation
at SBE AC Spring Meeting by K. Bollen

Some concepts

REPLICABILITY

Refers to the ability of a researcher to duplicate the results of a prior study if the same procedures are followed but new data are collected

• a failure to replicate occurs when one study documents relations and a subsequent attempt with new data fails to yield the same relations

• null results or nonzero results could be replications

o E.g., failure to find intervention to work in two different data sets is a replication as would be the finding of positive effect

• Same researcher performing second study more likely to replicate
  o Fully aware of procedures

• Second researcher in another location less likely because:
o Did not directly observe the first study
o Relies on text description of first study
o Critical details not fully understood or described
o Failure to replicate might be due to different procedures

Report of the Subcommittee on Replicability in Science of the SBE Advisory Committee to the National Science Foundation
13 May 2015 Presentation
at SBE AC Spring Meeting by K. Bollen

Some concepts

GENERALIZABILITY

refers to whether the results of a study apply in other contexts or populations that differ from the original one

- degree to which found relations apply in different situations
- E.g., do findings based on college students apply to adult population of the United States?
- E.g., does an experiment that uses one type of persuasive message work when researcher tries other types of persuasive messages?
- Failure to generalize directs attention to operation of limiting conditions on relationship

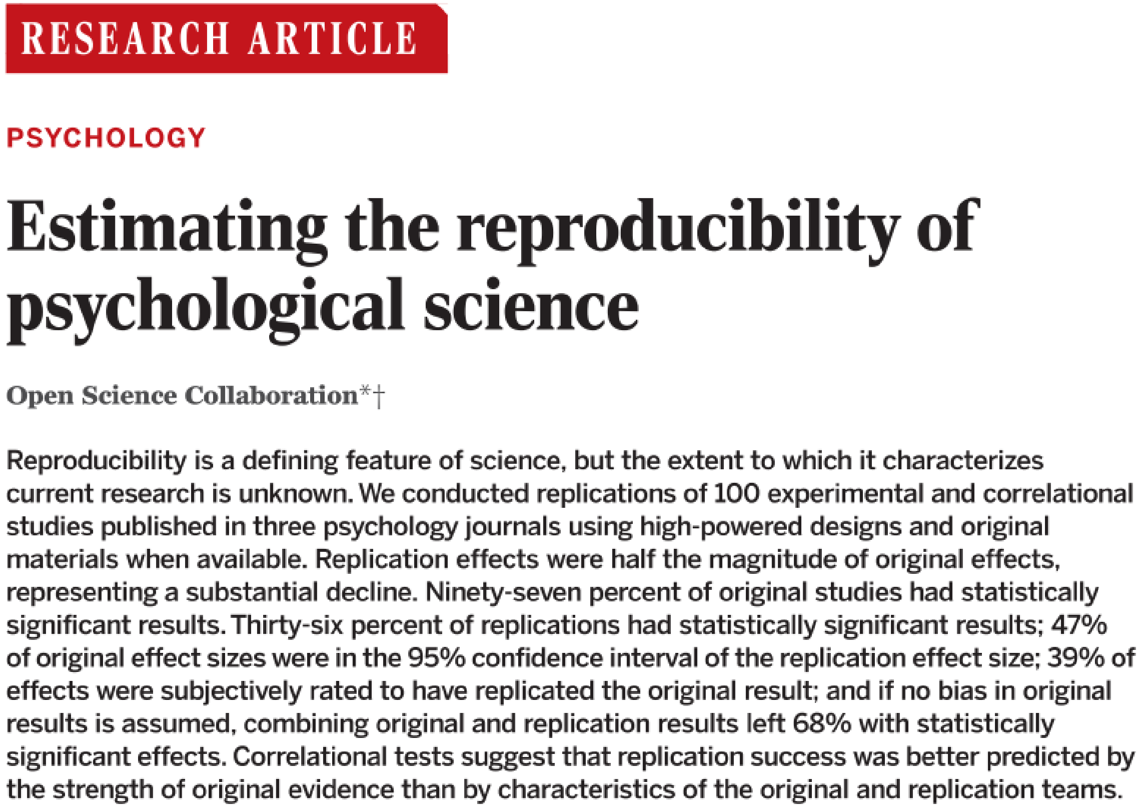
o Chance to advance theory as these limiting conditions are uncovered

Report of the Subcommittee on Replicability in Science of the SBE
Advisory Committee to the National Science Foundation
13 May 2015 Presentation
at SBE AC Spring Meeting by K. Bollen

Replication study of the Open Science Collaboration

- 2008 articles of: Psychological Science (PSCI), Journal of Personality and Social Psychology (JPSP), and Journal of Experimental Psychol- ogy: Learning, Memory, and Cognition (JEP: LMC)

- 100 replications by 270 contributing authors



RESEARCH ARTICLE

PSYCHOLOGY

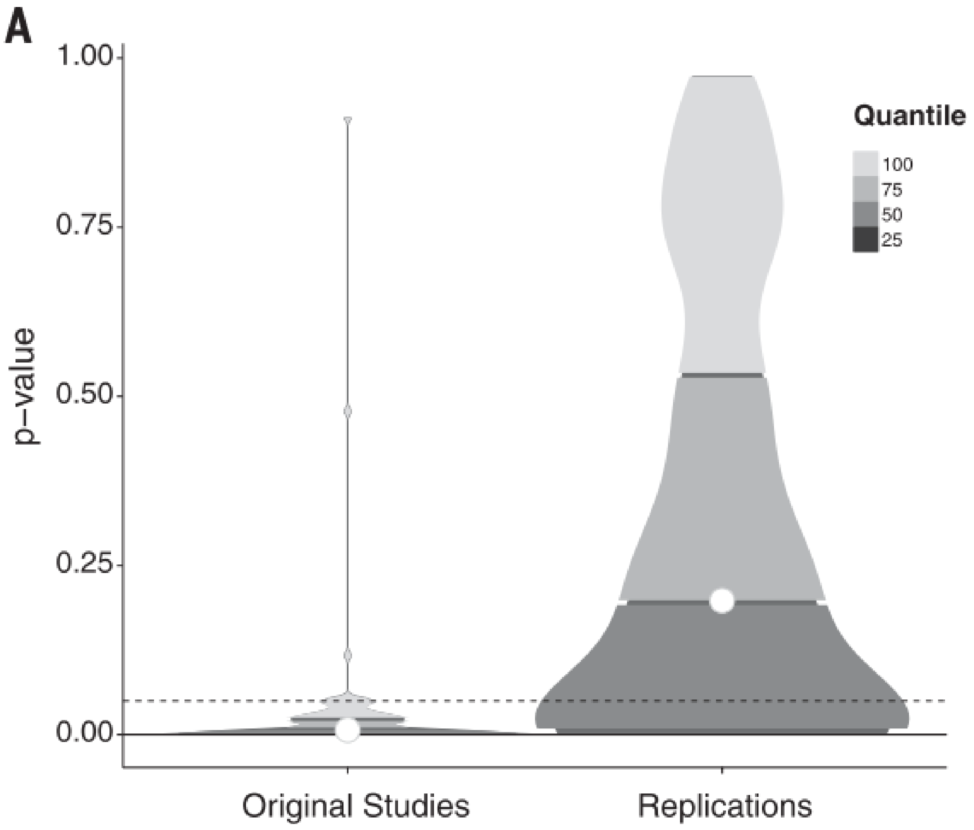
Estimating the reproducibility of psychological science

Open Science Collaboration*†

Reproducibility is a defining feature of science, but the extent to which it characterizes current research is unknown. We conducted replications of 100 experimental and correlational studies published in three psychology journals using high-powered designs and original materials when available. Replication effects were half the magnitude of original effects, representing a substantial decline. Ninety-seven percent of original studies had statistically significant results. Thirty-six percent of replications had statistically significant results; 47% of original effect sizes were in the 95% confidence interval of the replication effect size; 39% of effects were subjectively rated to have replicated the original result; and if no bias in original results is assumed, combining original and replication results left 68% with statistically significant effects. Correlational tests suggest that replication success was better predicted by the strength of original evidence than by characteristics of the original and replication teams.

Gilbert et al. (2016). Estimating the reproducibility of psychological science. *Science*, *351*(6277). https://doi.org/10.1126/science.aad7243

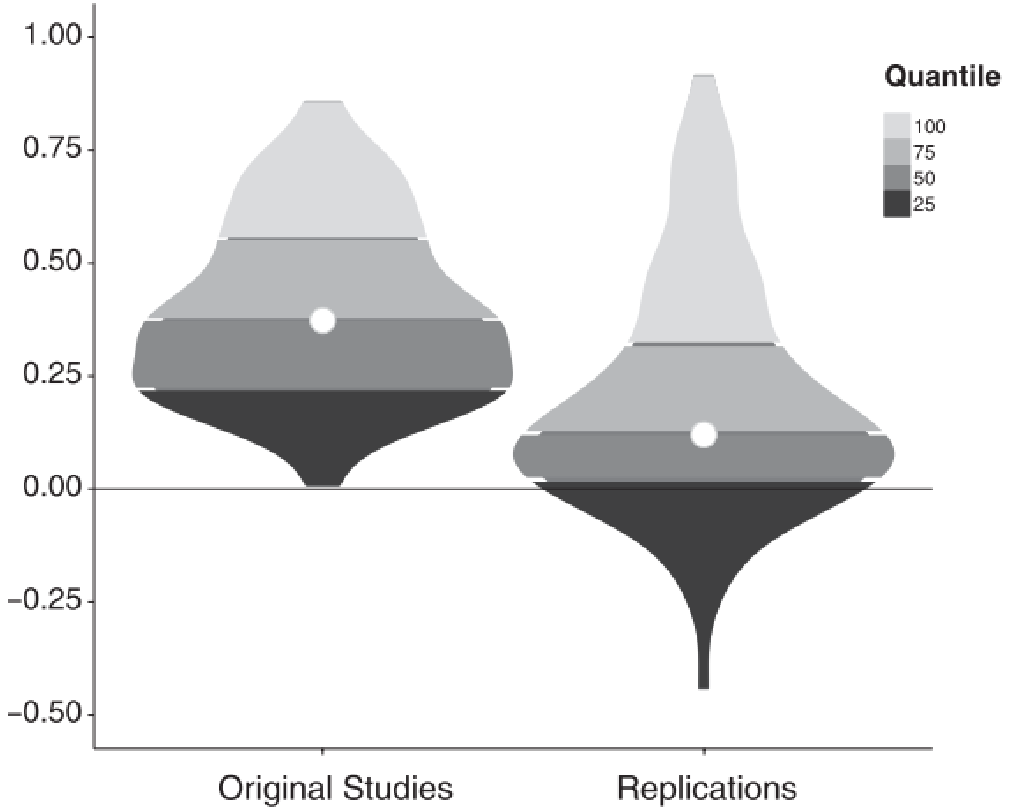
Gilbert et al. (2016)



Does replication show a p < .05 in the same direction?

35 replication studies showed significant effects versus 98 in the original studies

Gilbert et al. (2016)



Is the original effect size within 95 % CI of the replication?

It is in 48 % of the studies

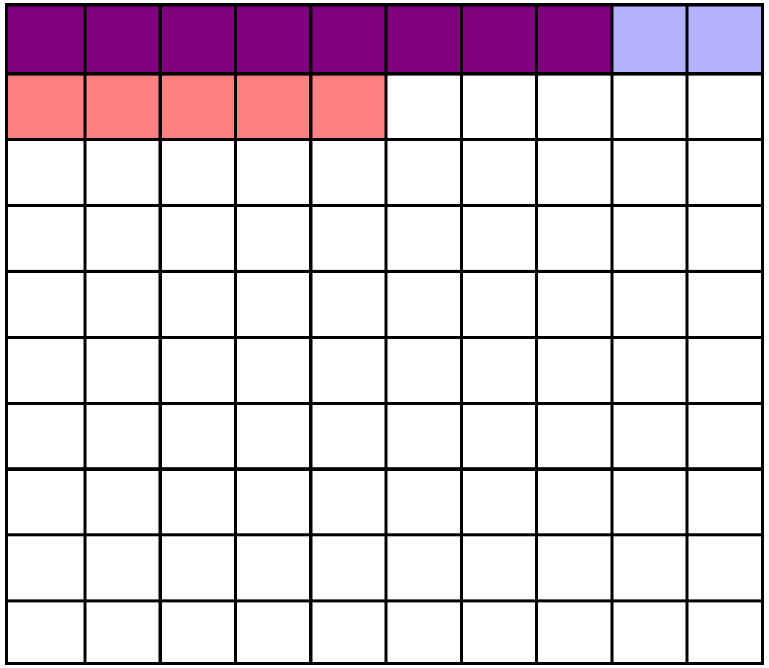
Looks like we have a false positive problem! Why?

We are only human:

- Apophenia (the tendency to see patterns in random data)
- Confirmation bias (the tendency to focus on evidence that is in line with our expectations or favoured explanation)
- Hindsight bias (the tendency to see an event as having been predictable only after it has occurred)

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., … Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 0021. https://doi.org/10.1038/s41562-016-0021

Why? Some flaws – e.g. difficulty in understanding how p-values work: base rate fallacy



Testing 100 drugs

Top 10 truly effective

Power of 80% -> I detect 8 of the 20 truly effective

p = .05 - > 5 of the 90 non-effective are false-positives

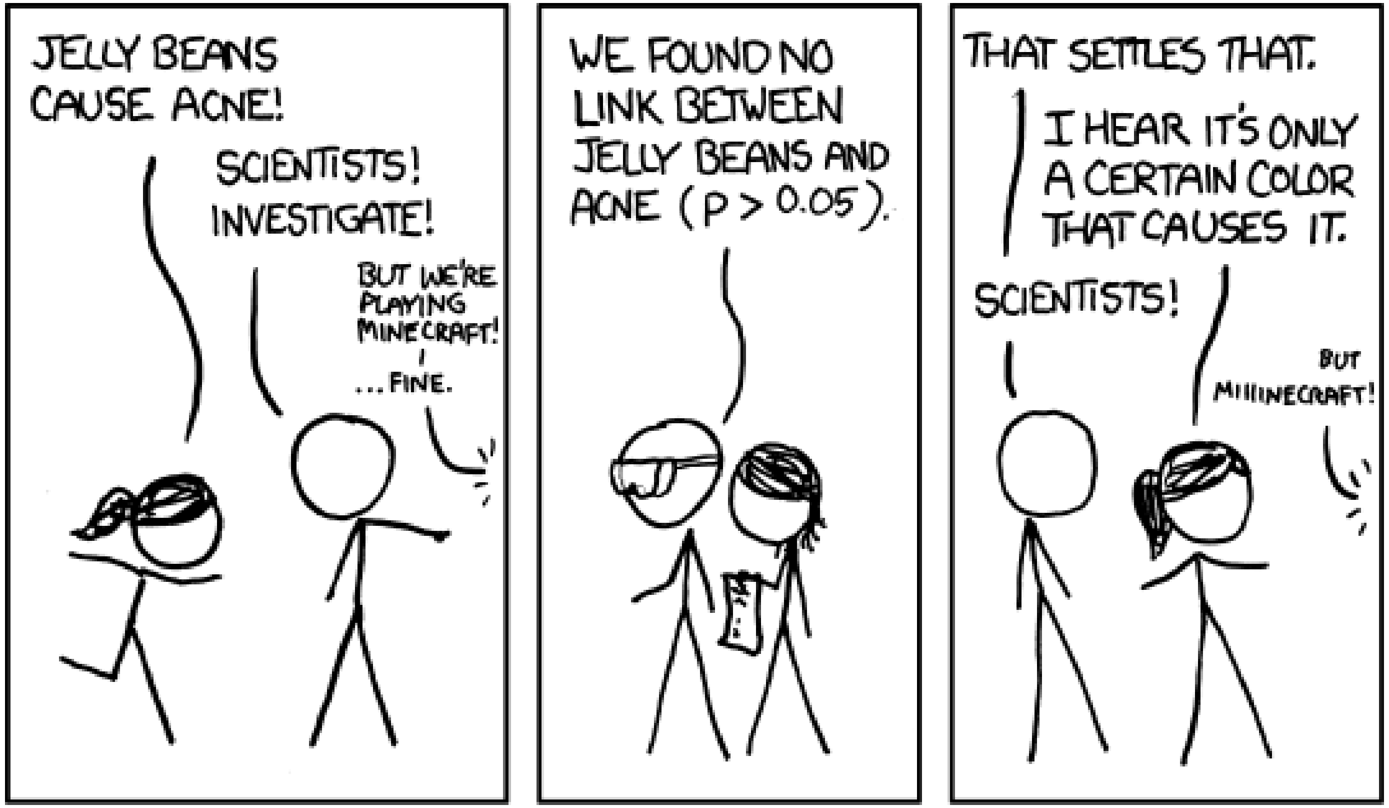
I “see” 5 + 8 as “effective” = 13

But I am getting 5/13 = 38% of false discovery! (not 5 %)

Reinhart, 2015

Difficulty in understanding how p-values work: e.g. multiple testing

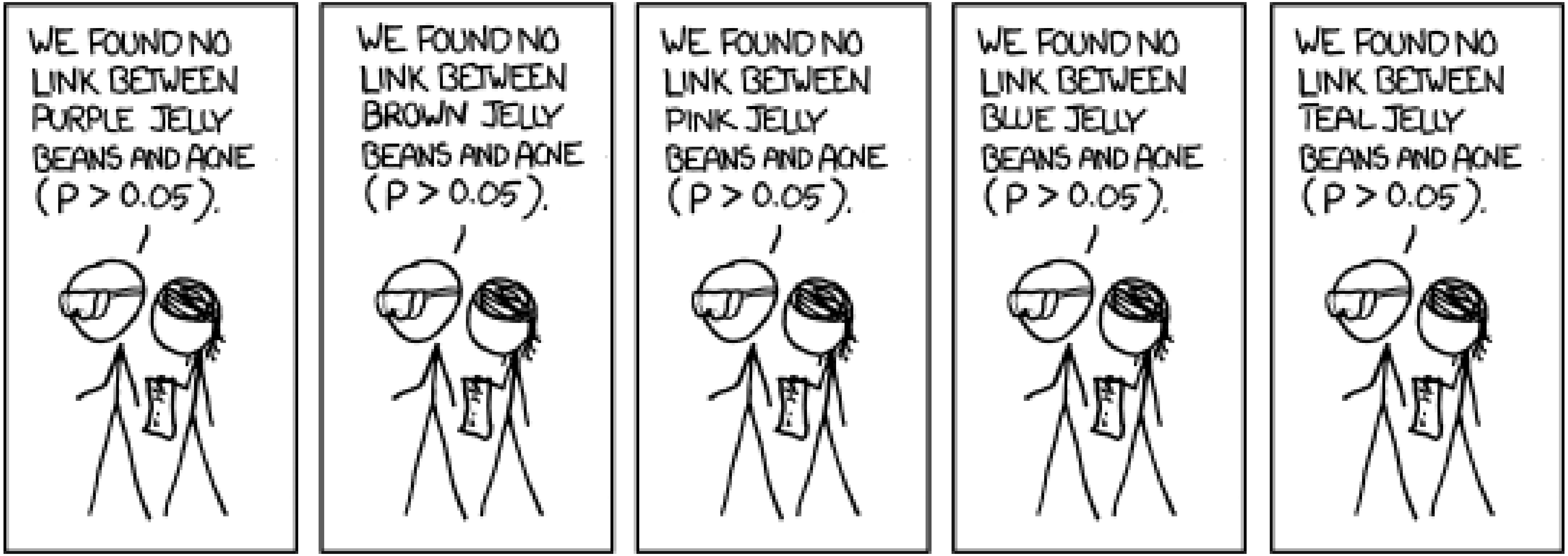
Reinhart, 2015


JELLY BEANS CAUSE ACNE!
SCIENTISTS! INVESTIGATE!
BUT WE'RE PLAYING MINECRAFT!
...FINE.
WE FOUND NO LINK BETWEEN JELLY BEANS AND ACNE (P > 0.05).
THAT SETTLES THAT.
I HEAR IT'S ONLY A CERTAIN COLOR THAT CAUSES IT.
SCIENTISTS!
BUT MIIIINECRAFT!

Cartoon from xkcd, by Randall Munroe. http://xkcd.com/882/

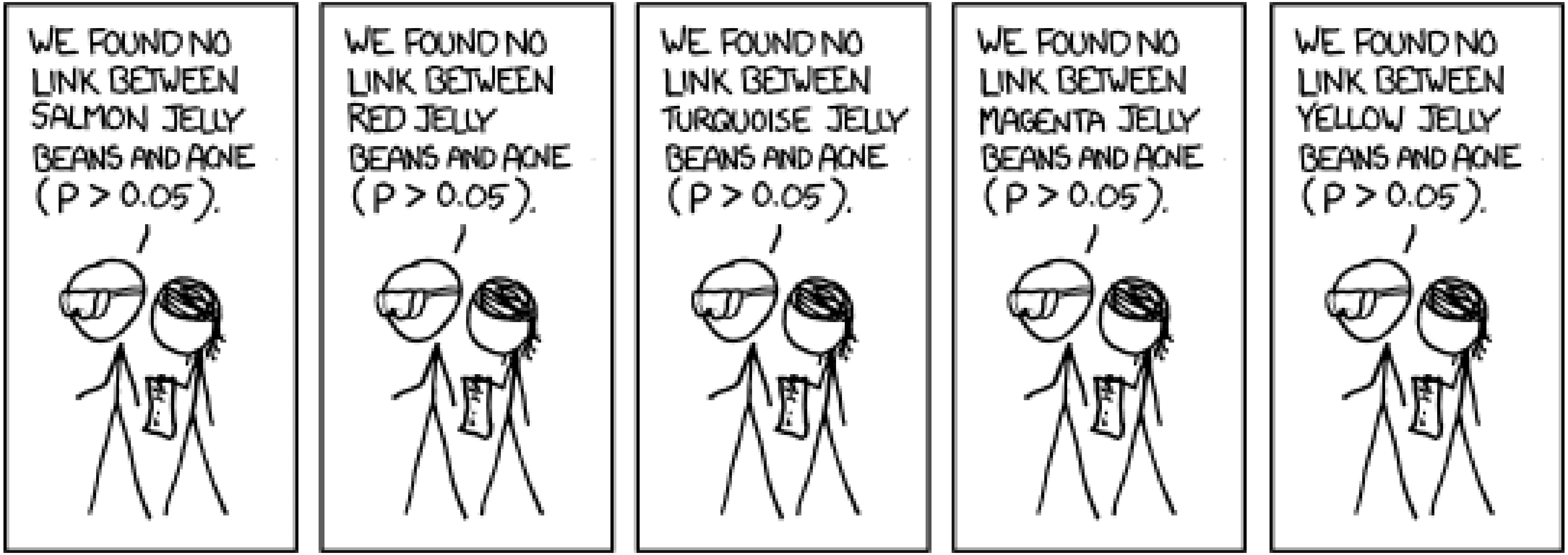
Difficulty in understanding how p-values work: e.g. multiple testing

Reinhart, 2015


WE FOUND NO LINK BETWEEN PURPLE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN BROWN JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN PINK JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN BLUE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN TEAL JELLY BEANS AND ACNE (P > 0.05).

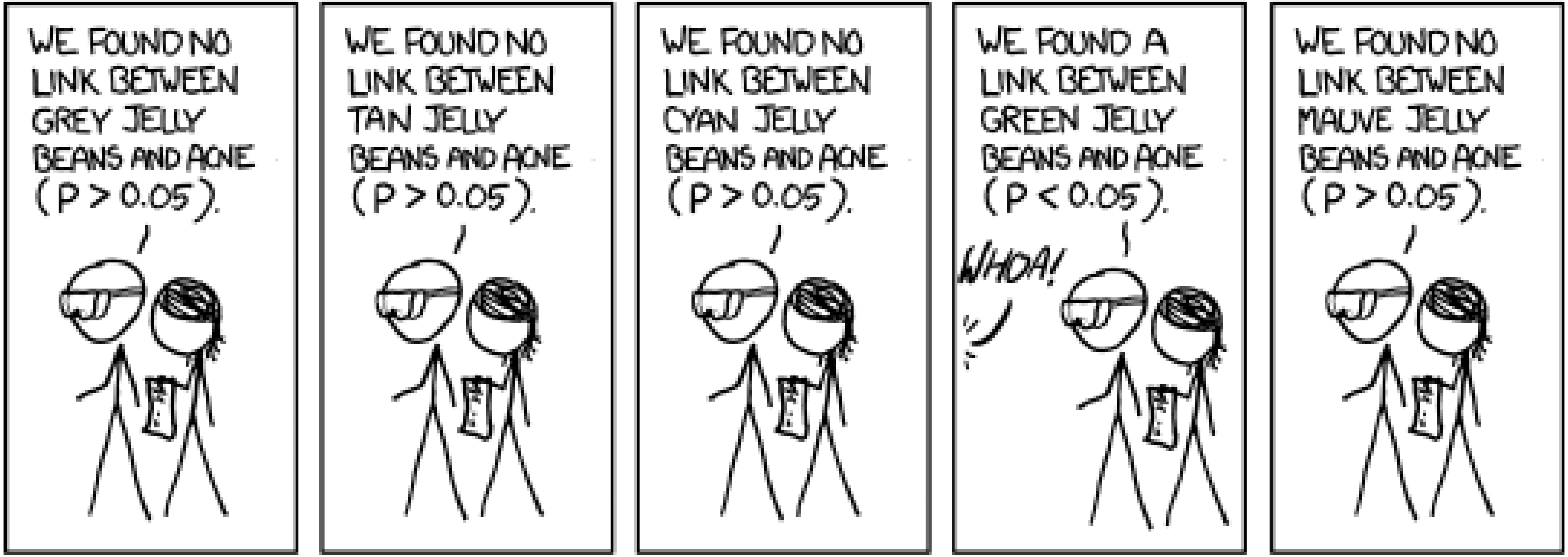
Difficulty in understanding how p-values work: e.g. multiple testing

Reinhart, 2015


WE FOUND NO LINK BETWEEN SALMON JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN RED JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN TURQUOISE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN MAGENTA JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN YELLOW JELLY BEANS AND ACNE (P > 0.05).

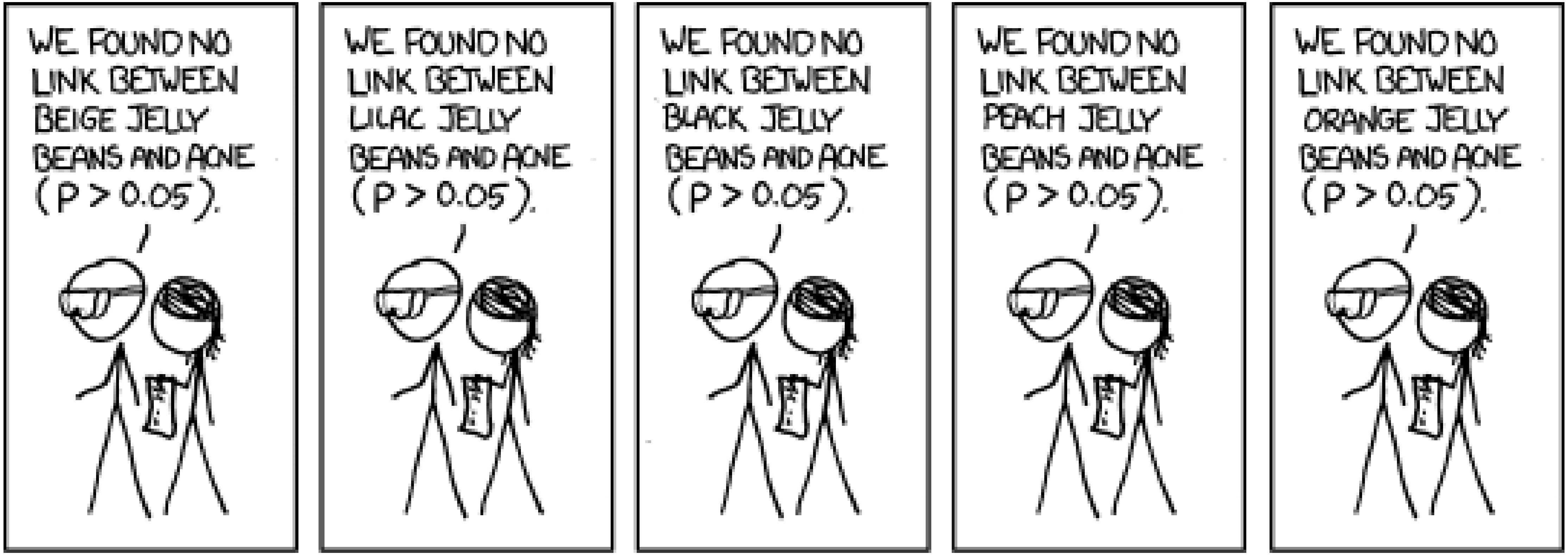
Difficulty in understanding how p-values work: e.g. multiple testing

Reinhart, 2015


WE FOUND NO LINK BETWEEN GREY JELLY BEANS AND ACNE ($P > 0.05$).
WE FOUND NO LINK BETWEEN TAN JELLY BEANS AND ACNE ($P > 0.05$).
WE FOUND NO LINK BETWEEN CYAN JELLY BEANS AND ACNE ($P > 0.05$).
WE FOUND A LINK BETWEEN GREEN JELLY BEANS AND ACNE ($P < 0.05$).
WHOA!
WE FOUND NO LINK BETWEEN MAUVE JELLY BEANS AND ACNE ($P > 0.05$).

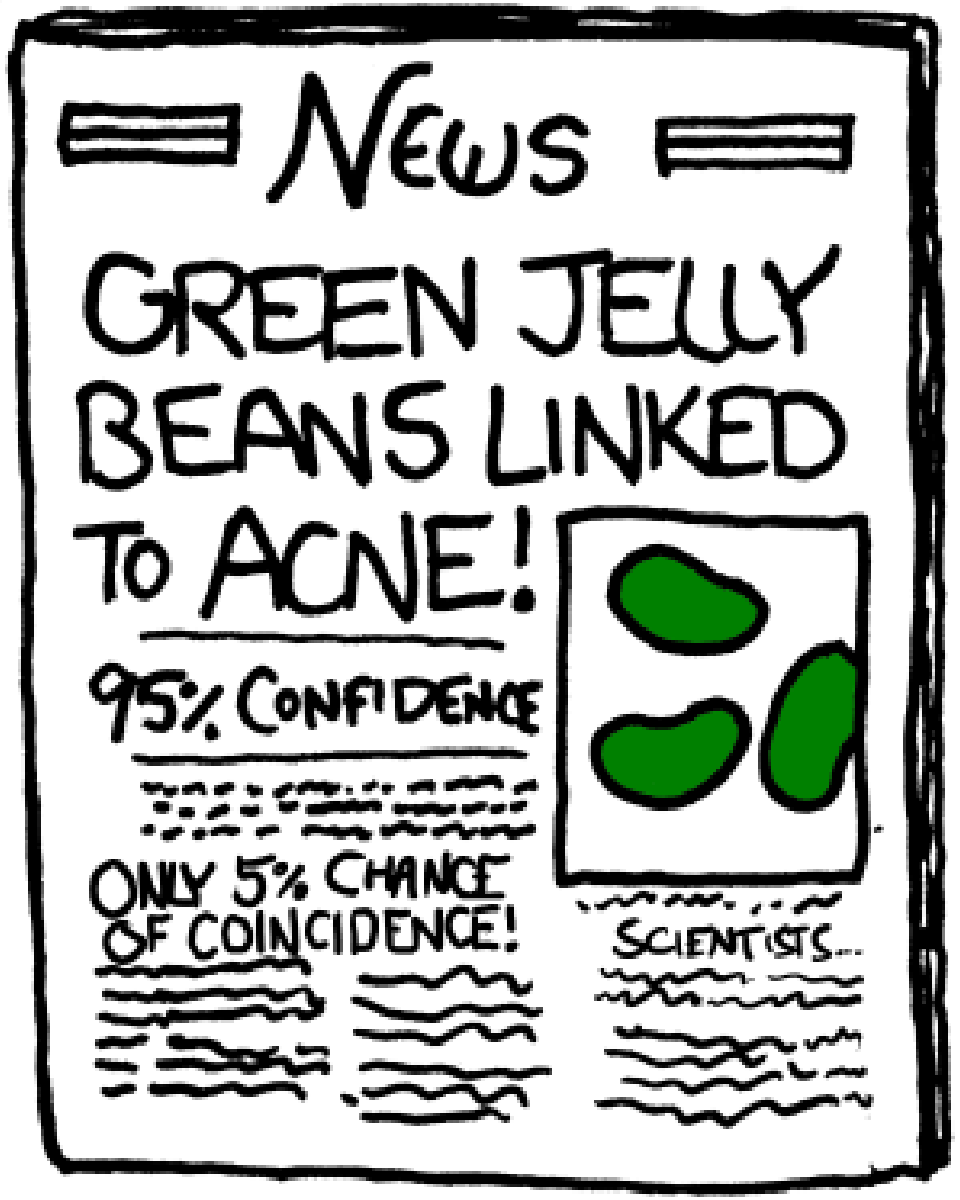
Difficulty in understanding how p-values work: e.g. multiple testing

Reinhart, 2015


WE FOUND NO LINK BETWEEN BEIGE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN LILAC JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN BLACK JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN PEACH JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN ORANGE JELLY BEANS AND ACNE (P > 0.05).

Difficulty in understanding how p-values work: e.g. multiple testing

Reinhart, 2015


NEWS
GREEN JELLY BEANS LINKED TO ACNE!
95% CONFIDENCE
ONLY 5% CHANCE OF COINCIDENCE!
SCIENTISTS...

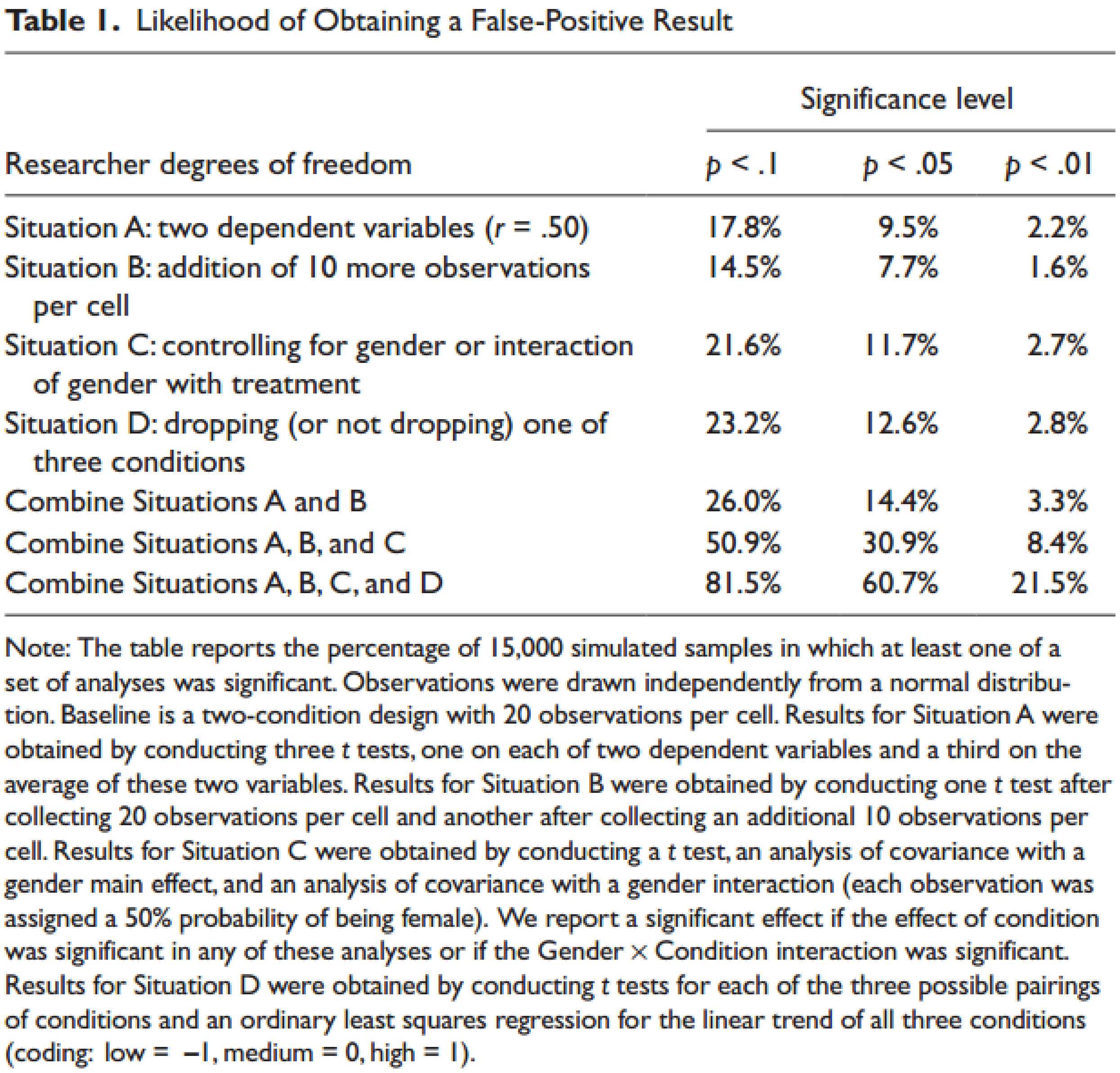


Table 1. Likelihood of Obtaining a False-Positive Result

	Significance level		
Researcher degrees of freedom	$p < .1$	$p < .05$	$p < .01$
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%

Note: The table reports the percentage of 15,000 simulated samples in which at least one of a set of analyses was significant. Observations were drawn independently from a normal distribution. Baseline is a two-condition design with 20 observations per cell. Results for Situation A were obtained by conducting three *t* tests, one on each of two dependent variables and a third on the average of these two variables. Results for Situation B were obtained by conducting one *t* test after collecting 20 observations per cell and another after collecting an additional 10 observations per cell. Results for Situation C were obtained by conducting a *t* test, an analysis of covariance with a gender main effect, and an analysis of covariance with a gender interaction (each observation was assigned a 50% probability of being female). We report a significant effect if the effect of condition was significant in any of these analyses or if the Gender × Condition interaction was significant. Results for Situation D were obtained by conducting *t* tests for each of the three possible pairings of conditions and an ordinary least squares regression for the linear trend of all three conditions (coding: low = −1, medium = 0, high = 1).

Simmons, 2011

Optional stopping

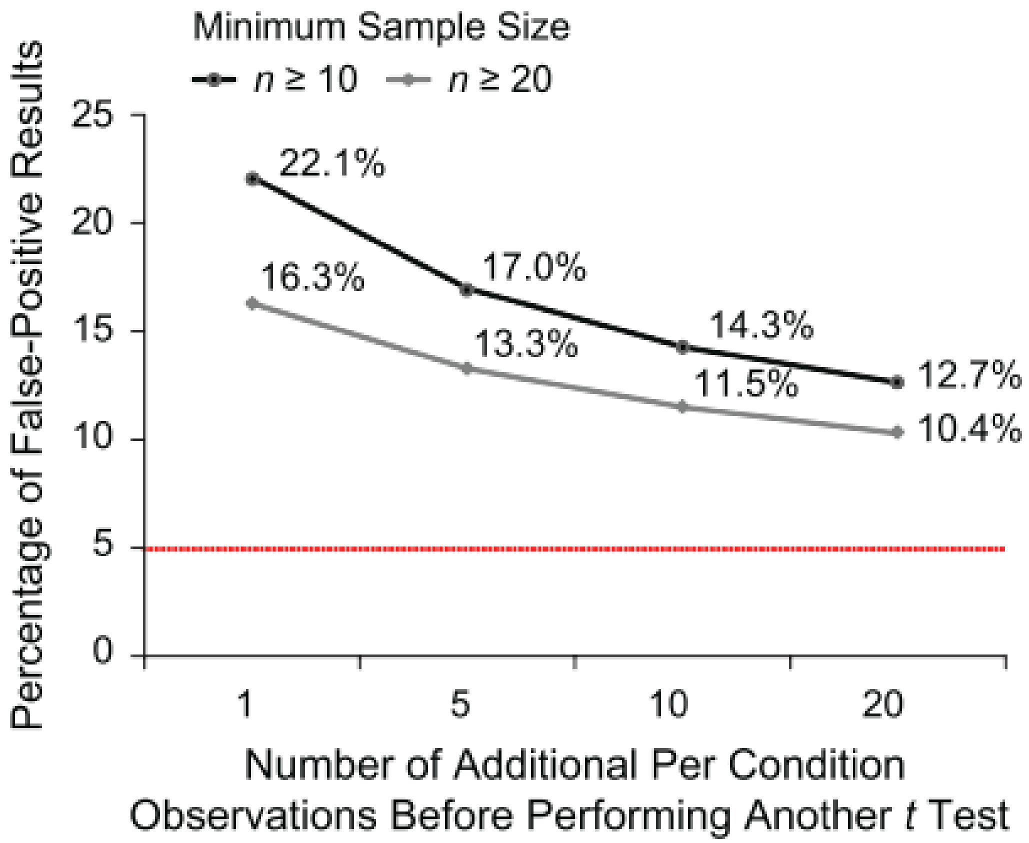
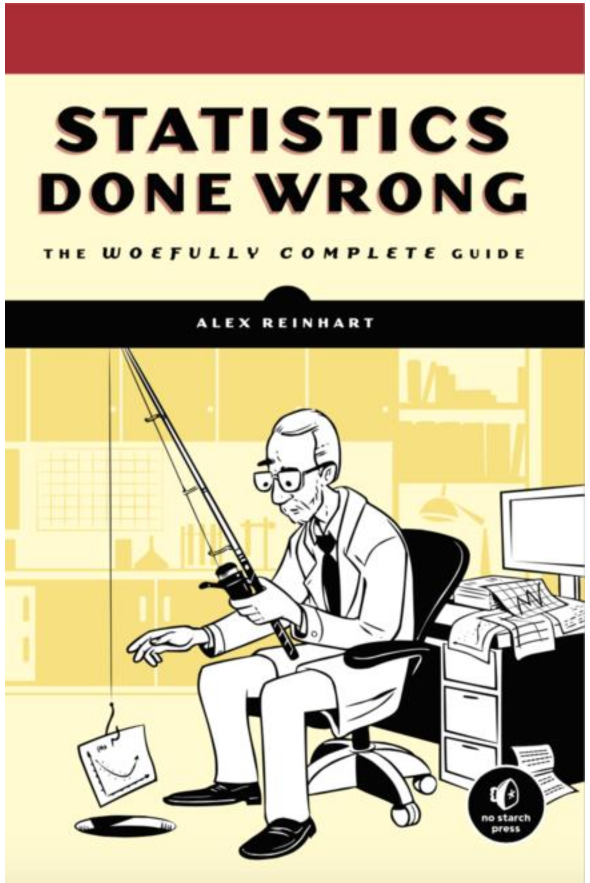


Fig. 1. Likelihood of obtaining a false-positive result when data collection ends upon obtaining significance ($p \leq .05$, highlighted by the dotted line). The figure depicts likelihoods for two minimum sample sizes, as a function of the frequency with which significance tests are performed.

Simmons, 2011

Difficulty in understanding how p-values work



https://www.statisticsdonewrong.com

Why many results in Psychology may be false

- p-hacking (fishing for significant results) or just misunderstanding
  - Using only p-values
  - Underpowered studies
  - Optional stopping
  - Selective removal of outliers
  - Selective reporting of results
- Publication bias of negative or complex findings
- HARKing: hypothesizing after results are known.

Zwaan, 2015

The researcher degrees of freedom: design

- Conducting explorative research without any hypothesis
- Studying a vague hypothesis that fails to specify the direction of the effect
- Creating multiple manipulated independent variables and conditions
- Measuring additional variables that can later be selected as covariates, independent variables, mediators, or moderators
- Measuring the same dependent variable in several alternative ways
- Measuring additional constructs that could potentially act as primary outcomes
- Measuring additional variables that enable later exclusion of participants from the analyses (e.g., awareness or manipulation checks)
- Failing to conduct a well-founded power analysis
- Failing to specify the sampling plan and allowing for running (multiple) small studies

Wicherts et al. (2016)

The researcher degrees of freedom: data collection

- Failing to randomly assign participants to conditions
- Insufficient blinding of participants and/or experimenters
- Correcting, coding, or discarding data during data collection in a non-blinded manner
- Determining the data collection stopping rule on the basis of desired results or intermediate significance testing Choosing

Wicherts et al. (2016)

The researcher degrees of freedom: analysis

- Choosing between different options of dealing with incomplete or missing data on ad hoc grounds
- Specifying pre-processing of data (e.g., cleaning, normalization, smoothing, motion correction) in an ad hoc manner
- Deciding how to deal with violations of statistical assumptions in an ad hoc manner
- Deciding on how to deal with outliers in an ad hoc manner
- Selecting the dependent variable out of several alternative measures of the same construct
- Trying out different ways to score the chosen primary dependent variable
- Selecting another construct as the primary outcome

Wicherts et al. (2016)

The researcher degrees of freedom: analysis

- Selecting independent variables out of a set of manipulated independent variables
- Operationalizing manipulated independent variables in different ways (e.g., by discarding or combining levels of factors)
- Choosing to include different measured variables as covariates, independent variables, mediators, or moderators
- Operationalizing non-manipulated independent variables in different ways
- Using alternative inclusion and exclusion criteria got selecting participants in analyses
- Choosing between different statistical models
- Choosing the estimation method, software package, and computation of SEs
- Choosing inference criteria (e.g., Bayes factors, alpha level, sidedness of the test, corrections for multiple testing)

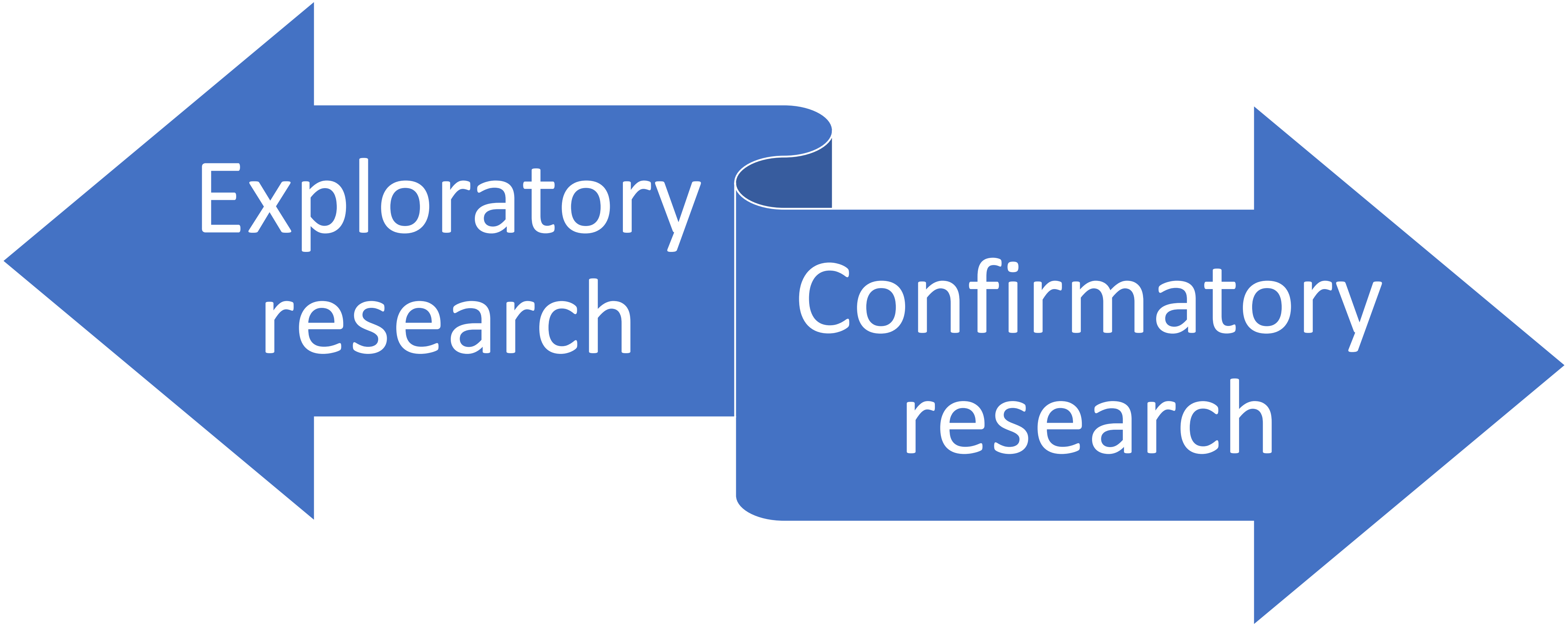
Wicherts et al. (2016)

The researcher degrees of freedom: reporting

- Failing to assure reproducibility (verifying the data collection and data analysis)
- Failing to enable replication (re-running of the study)
- Failing to mention, misrepresenting, or misidentifying the study preregistration
- Failing to report so-called “failed studies” that were originally deemed relevant to the research question
- Misreporting results and p-values
- Presenting exploratory analyses as confirmatory (HARKing)

Wicherts et al. (2016)

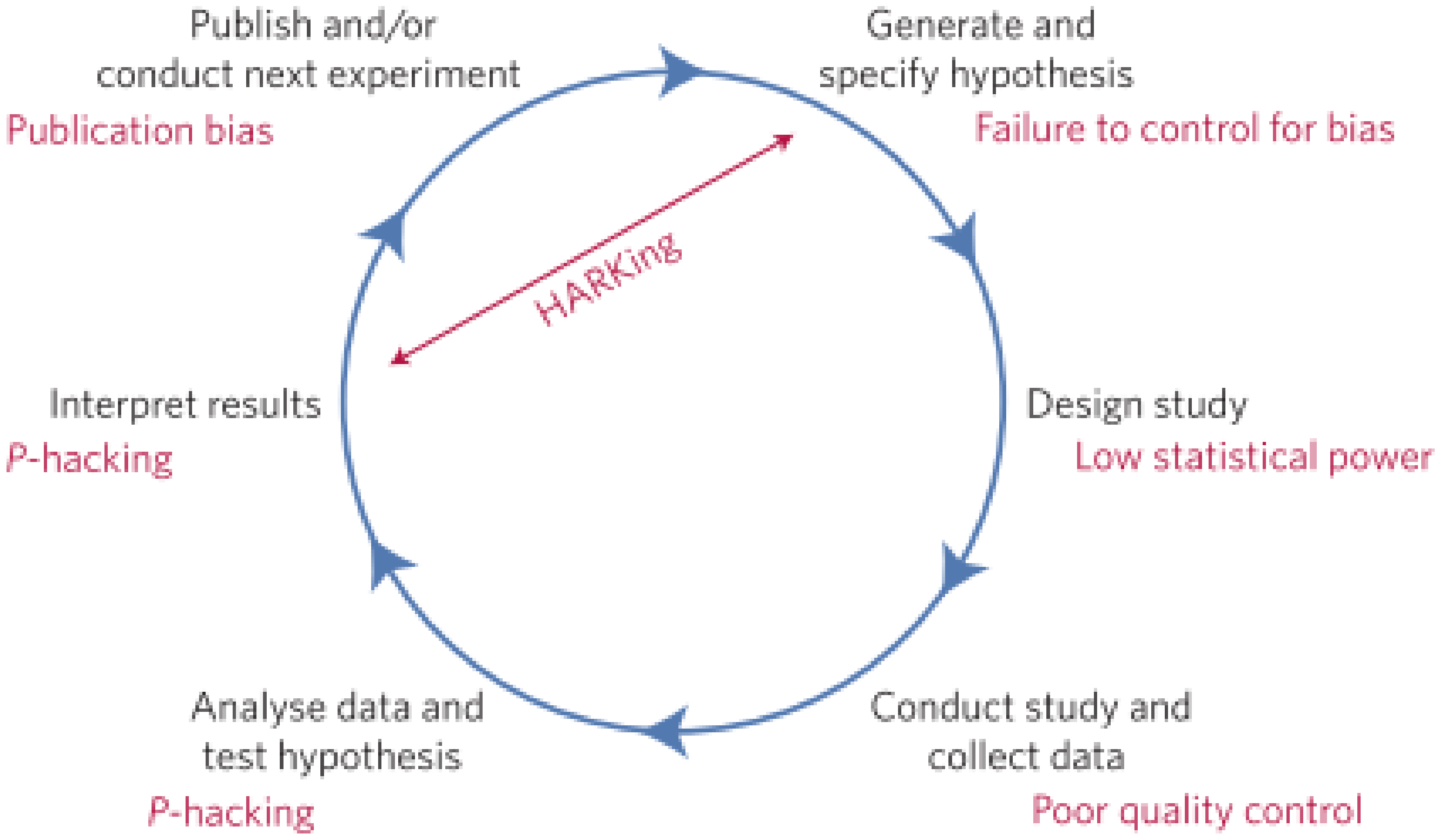
The researcher degrees of freedom issue



Exploratory research

Confirmatory research

“Small” mistakes that undermine our science



Publish and/or conduct next experiment

Publication bias

Generate and specify hypothesis

Failure to control for bias

HARKing

Design study

Low statistical power

Conduct study and collect data

Poor quality control

Analyse data and test hypothesis

P-hacking

Interpret results

P-hacking

Munafò et al. (2017)

Simple Solution to the Problem of False-Positive Publications: Authors

1. Authors must decide the rule for terminating data collection before data collection begins and report this rule in the article.
2. Authors must collect at least 20 observations per cell or else provide a compelling cost-of-data-collection justification.
3. Authors must list all variables collected in a study.
4. Authors must report all experimental conditions, including failed manipulations.
5. If observations are eliminated, authors must also report what the statistical results are if those observations are included.
6. If an analysis includes a covariate, authors must report the statistical results of the analysis without the covariate.

Simmons, 2011

Be transparent

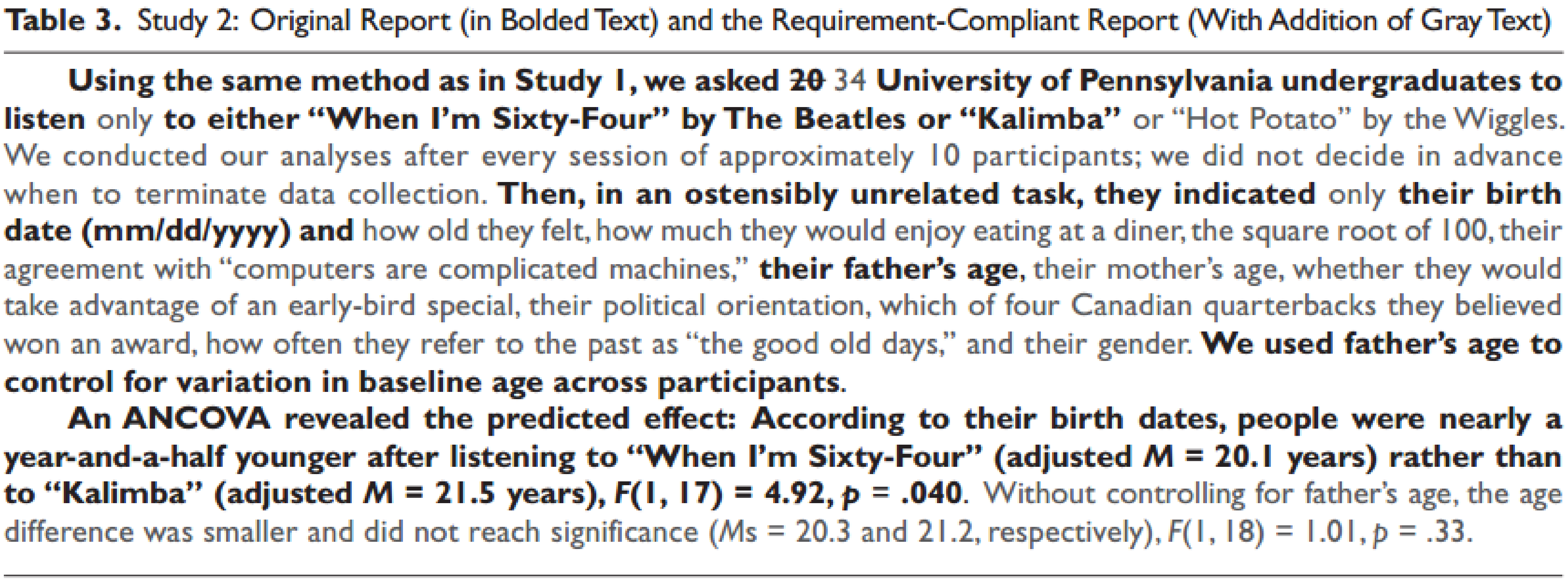


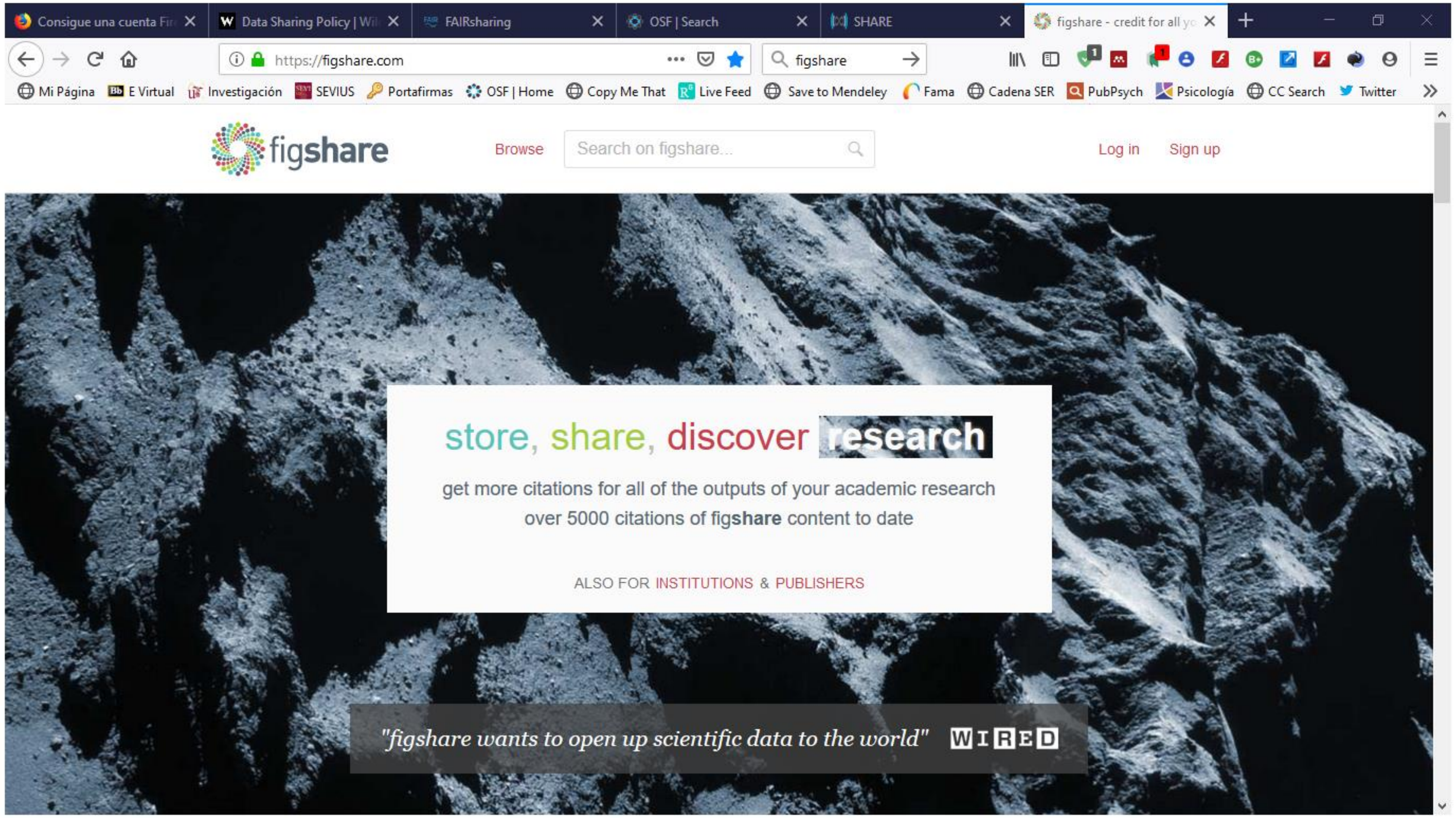
Table 3. Study 2: Original Report (in Bolded Text) and the Requirement-Compliant Report (With Addition of Gray Text)

Using the same method as in Study 1, we asked ~~20~~ 34 **University of Pennsylvania undergraduates to listen** only **to either "When I'm Sixty-Four" by The Beatles or "Kalimba"** or "Hot Potato" by the Wiggles. We conducted our analyses after every session of approximately 10 participants; we did not decide in advance when to terminate data collection. **Then, in an ostensibly unrelated task, they indicated** only **their birth date (mm/dd/yyyy) and** how old they felt, how much they would enjoy eating at a diner, the square root of 100, their agreement with "computers are complicated machines," **their father's age**, their mother's age, whether they would take advantage of an early-bird special, their political orientation, which of four Canadian quarterbacks they believed won an award, how often they refer to the past as "the good old days," and their gender. **We used father's age to control for variation in baseline age across participants**.

An ANCOVA revealed the predicted effect: According to their birth dates, people were nearly a year-and-a-half younger after listening to "When I'm Sixty-Four" (adjusted *M* = 20.1 years) rather than to "Kalimba" (adjusted *M* = 21.5 years), *F*(1, 17) = 4.92, *p* = .040. Without controlling for father's age, the age difference was smaller and did not reach significance (*M*s = 20.3 and 21.2, respectively), *F*(1, 18) = 1.01, *p* = .33.

Simmons, 201

More Solutions: data sharing



store, share, discover research

get more citations for all of the outputs of your academic research
over 5000 citations of **figshare** content to date

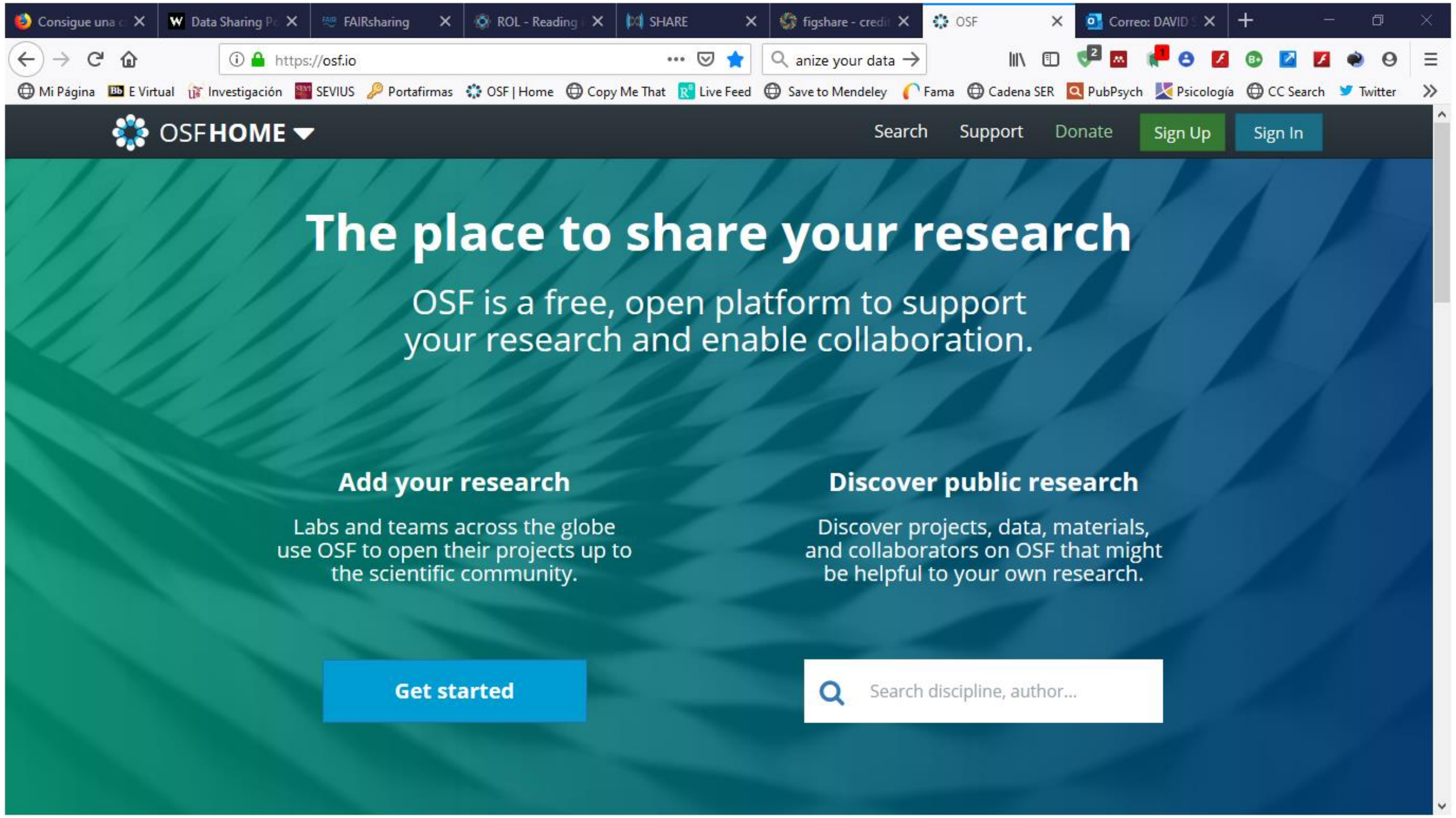
ALSO FOR INSTITUTIONS & PUBLISHERS

"figshare wants to open up scientific data to the world" WIRED

Some rules for data sharing

- Anticipate how your data will be used
- Keep raw data raw
- Store data in open formats
- Data structured for analysis (tidy data)
- Data should be uniquely identified
- Link relevant metadata
- Adopt proper privacy protocols
- Systematic backup scheme (2 onsite and 1 offsite)
- Analyse your capacity needs

Hart, E. M., Barmby, P., LeBauer, D., Michonneau, F., Mount, S., Mulrooney, P., … Hollister, J. W. (2016). Ten Simple Rules for Digital Data Storage. *PLoS Computational Biology*, *12*(10), 1-12. https://doi.org/10.1371/journal.pcbi.1005097



The place to share your research

OSF is a free, open platform to support your research and enable collaboration.

Add your research

Labs and teams across the globe use OSF to open their projects up to the scientific community.

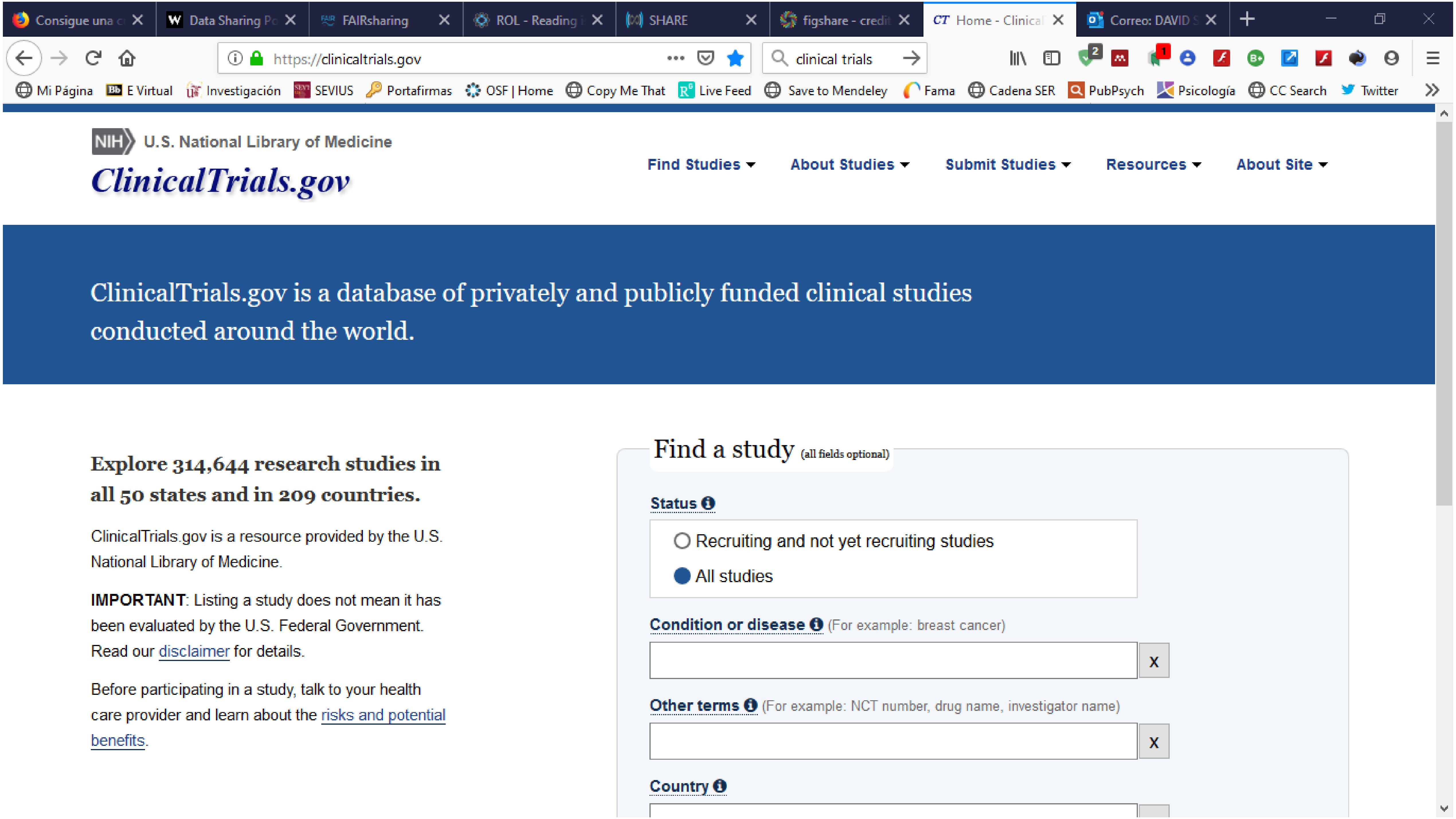
Get started

Discover public research

Discover projects, data, materials, and collaborators on OSF that might be helpful to your own research.

Search discipline, author...

Pre-registration



ClinicalTrials.gov is a database of privately and publicly funded clinical studies conducted around the world.

Explore 314,644 research studies in all 50 states and in 209 countries.

ClinicalTrials.gov is a resource provided by the U.S. National Library of Medicine.

IMPORTANT: Listing a study does not mean it has been evaluated by the U.S. Federal Government. Read our disclaimer for details.

Before participating in a study, talk to your health care provider and learn about the risks and potential benefits.

Find a study (all fields optional)

Status

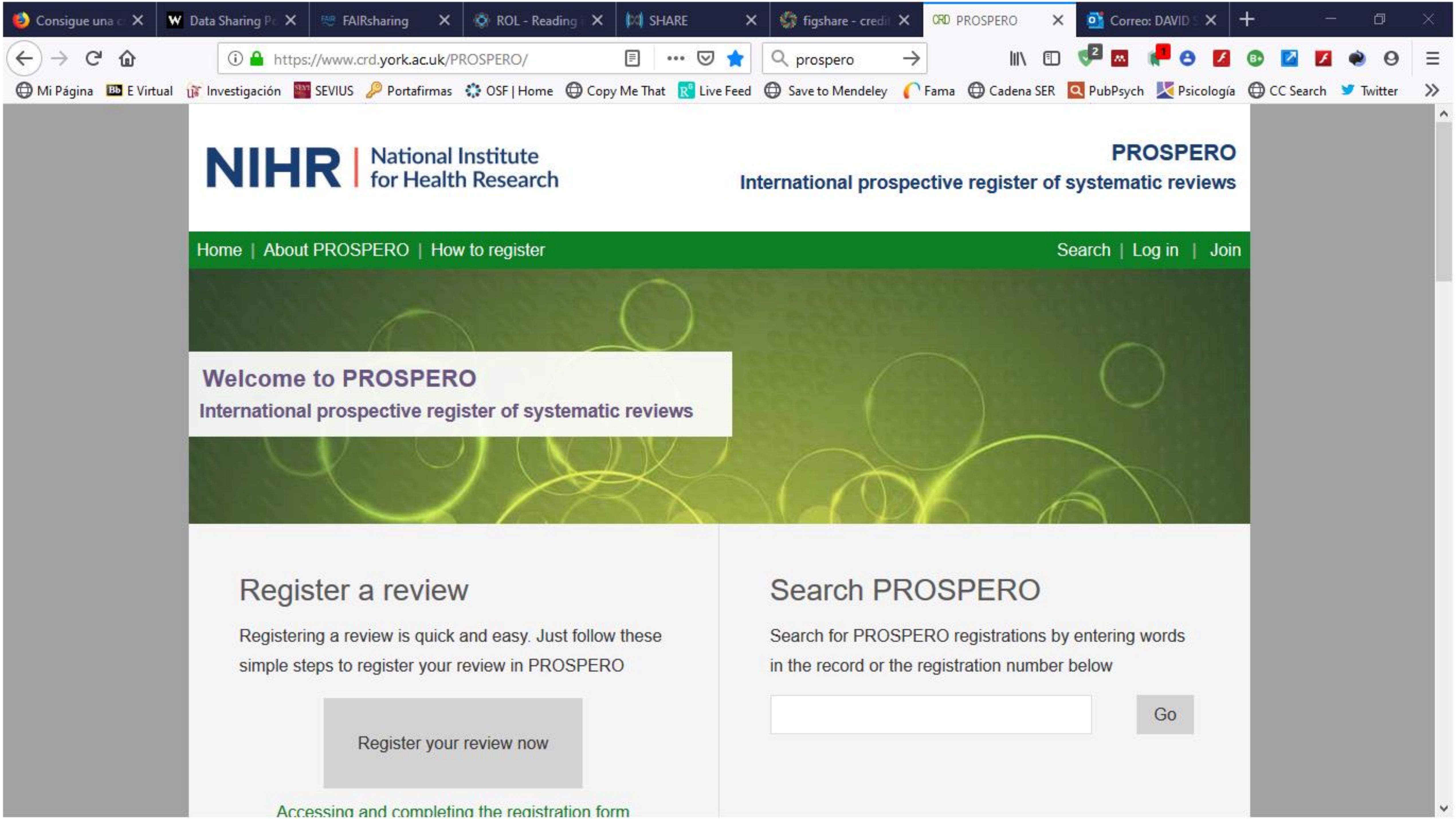
○ Recruiting and not yet recruiting studies

● All studies

Condition or disease (For example: breast cancer)

Other terms (For example: NCT number, drug name, investigator name)

Country



PROSPERO

International prospective register of systematic reviews

Welcome to PROSPERO

International prospective register of systematic reviews

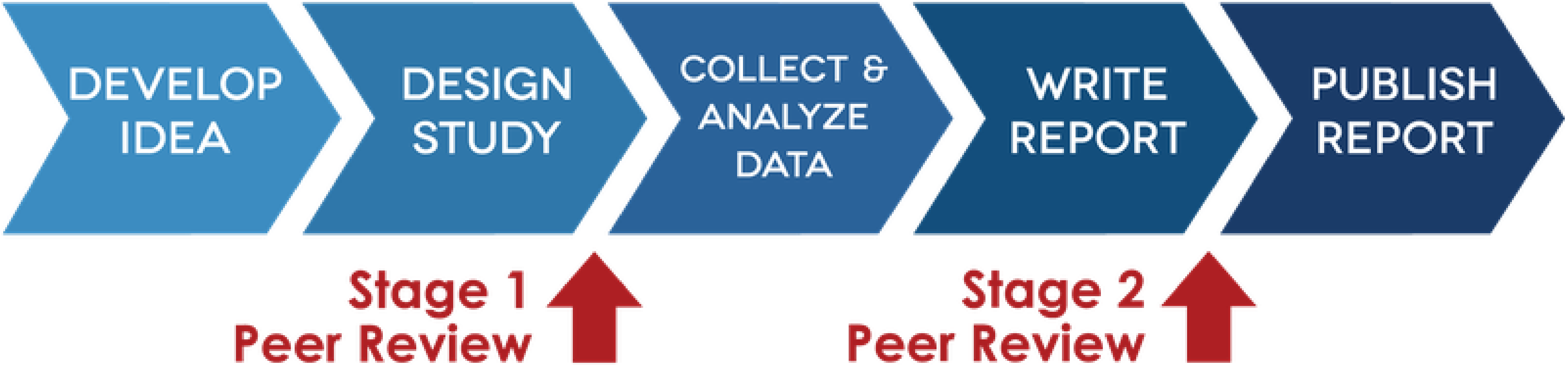
Register a review

Registering a review is quick and easy. Just follow these simple steps to register your review in PROSPERO

Accessing and completing the registration form

Search PROSPERO

Search for PROSPERO registrations by entering words in the record or the registration number below



DEVELOP IDEA

DESIGN STUDY

COLLECT & ANALYZE DATA

WRITE REPORT

PUBLISH REPORT

Stage 1 Peer Review

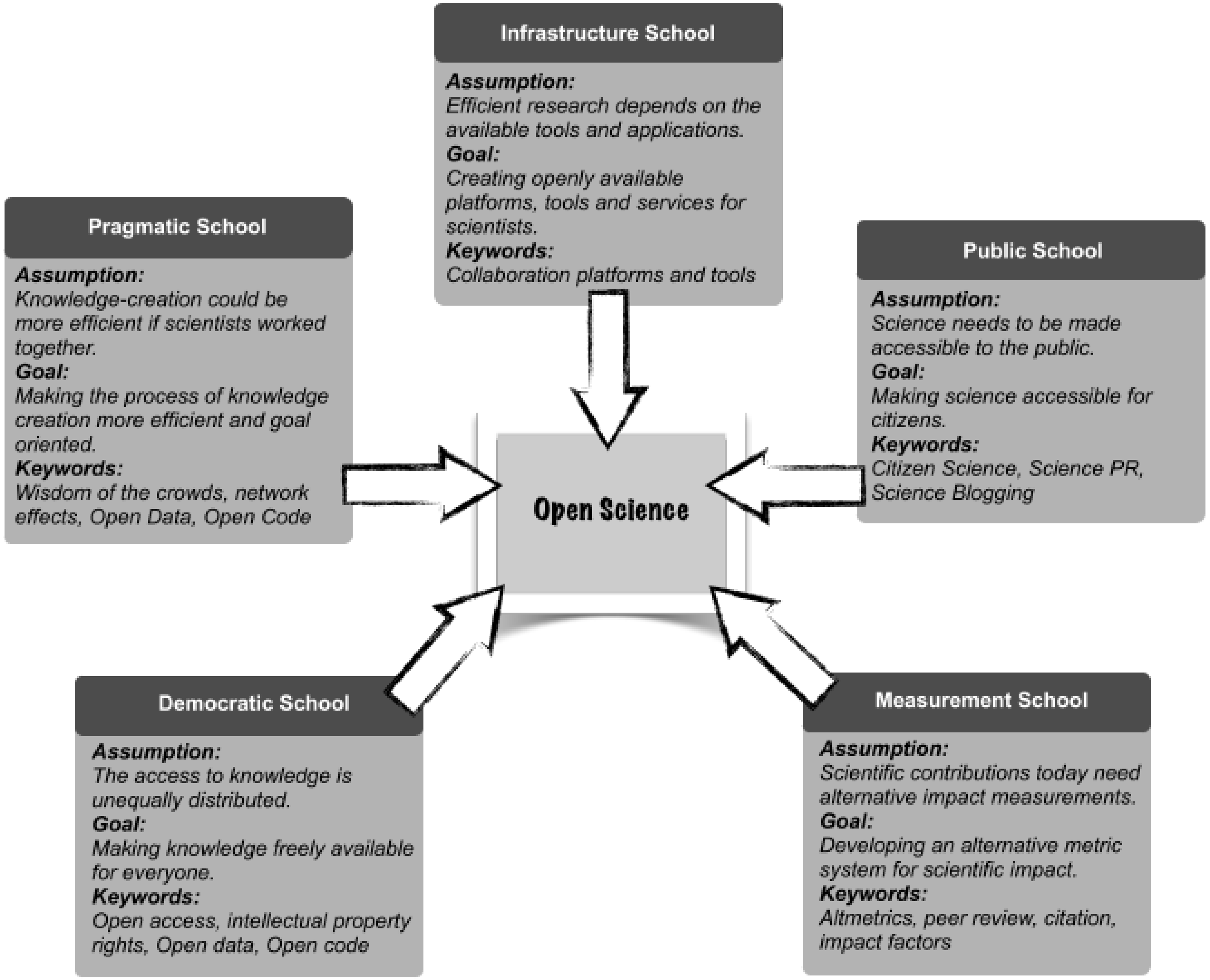
Stage 2 Peer Review

204 journals https://cos.io/rr/

Registered reports

https://www.fosteropenscience.eu/content/what-open-science-introduction

Availability



Infrastructure School

Assumption:
Efficient research depends on the available tools and applications.
Goal:
Creating openly available platforms, tools and services for scientists.
Keywords:
Collaboration platforms and tools

Pragmatic School

Assumption:
Knowledge-creation could be more efficient if scientists worked together.
Goal:
Making the process of knowledge creation more efficient and goal oriented.
Keywords:
Wisdom of the crowds, network effects, Open Data, Open Code

Public School

Assumption:
Science needs to be made accessible to the public.
Goal:
Making science accessible for citizens.
Keywords:
Citizen Science, Science PR, Science Blogging

Open Science

Democratic School

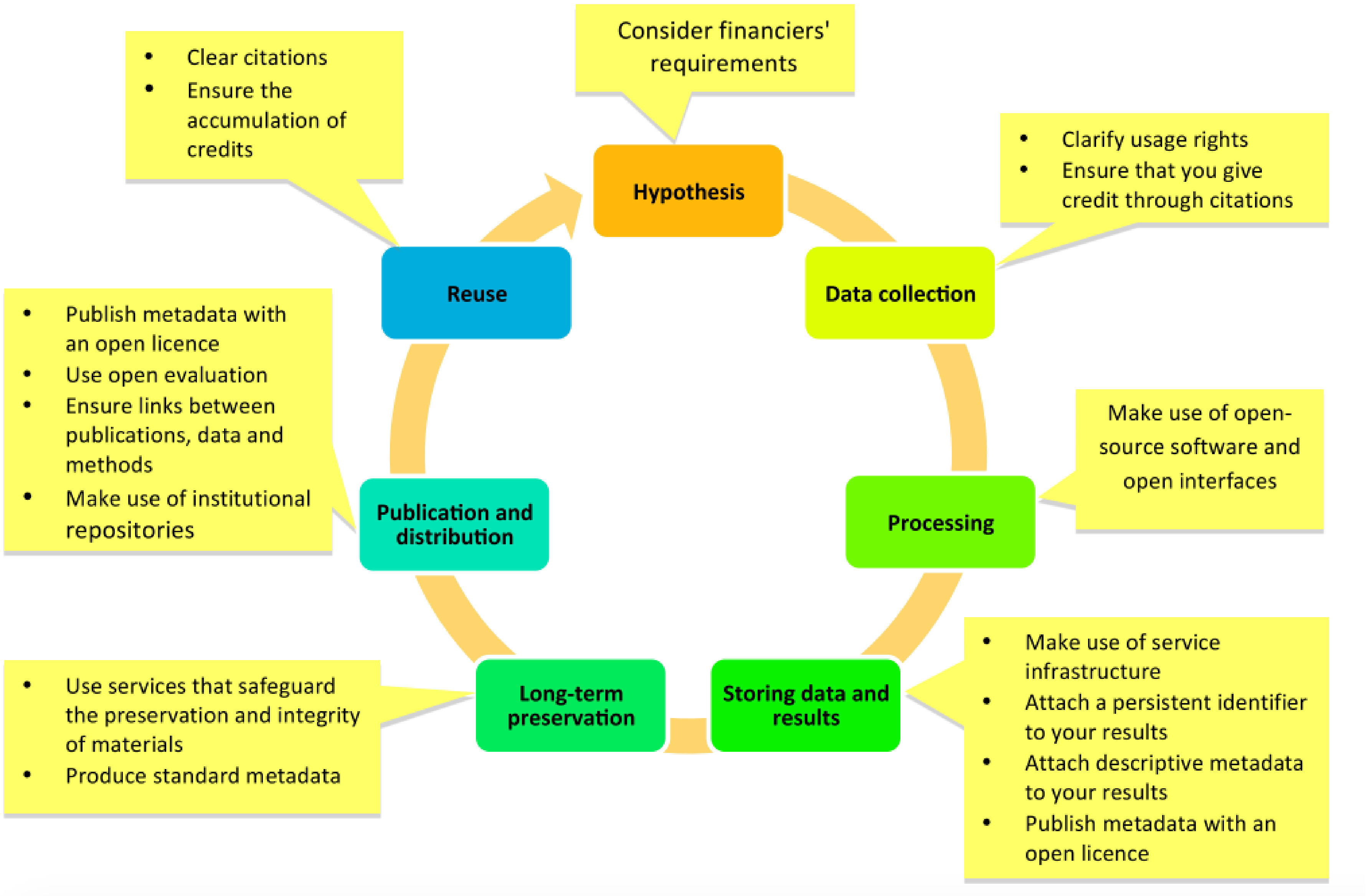
Assumption:
The access to knowledge is unequally distributed.
Goal:
Making knowledge freely available for everyone.
Keywords:
Open access, intellectual property rights, Open data, Open code

Measurement School

Assumption:
Scientific contributions today need alternative impact measurements.
Goal:
Developing an alternative metric system for scientific impact.
Keywords:
Altmetrics, peer review, citation, impact factors

opening science

Figure 3. Five Open Science schools of thought (Fecher and Friesike, 2014)



Hypothesis

Consider financiers' requirements

Data collection

- Clarify usage rights
- Ensure that you give credit through citations

Processing

Make use of open-source software and open interfaces

Storing data and results

- Make use of service infrastructure
- Attach a persistent identifier to your results
- Attach descriptive metadata to your results
- Publish metadata with an open licence

Long-term preservation

- Use services that safeguard the preservation and integrity of materials
- Produce standard metadata

Publication and distribution

- Publish metadata with an open licence
- Use open evaluation
- Ensure links between publications, data and methods
- Make use of institutional repositories

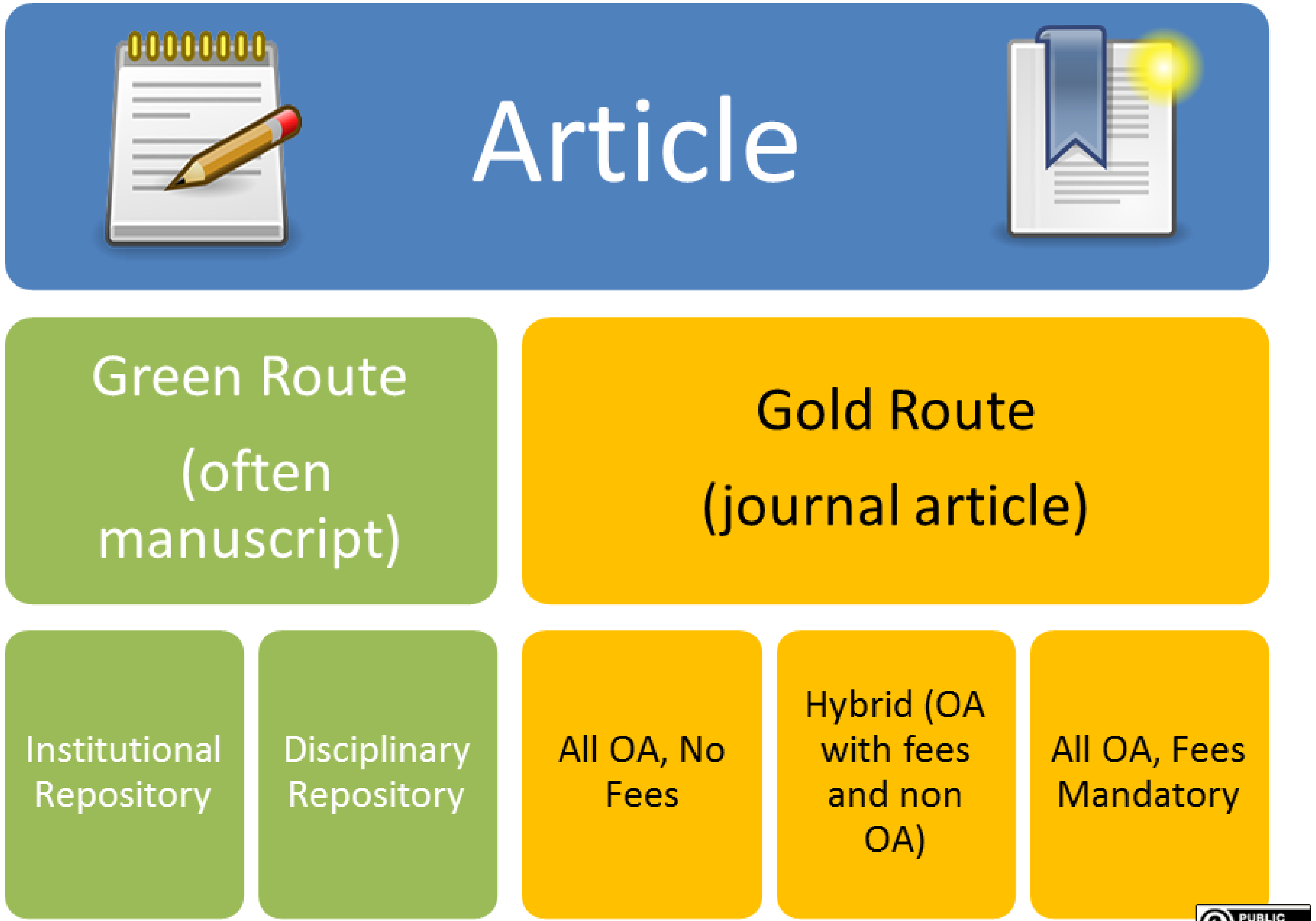
Reuse

- Clear citations
- Ensure the accumulation of credits

*Figure 1. Promoting openness at different stages of the research process (Open Science and Research Initiative, 2014)**

https://www.fosteropenscience.eu

Open Access publishing



Article

Green Route

(often manuscript)

Gold Route

(journal article)

Institutional Repository

Disciplinary Repository

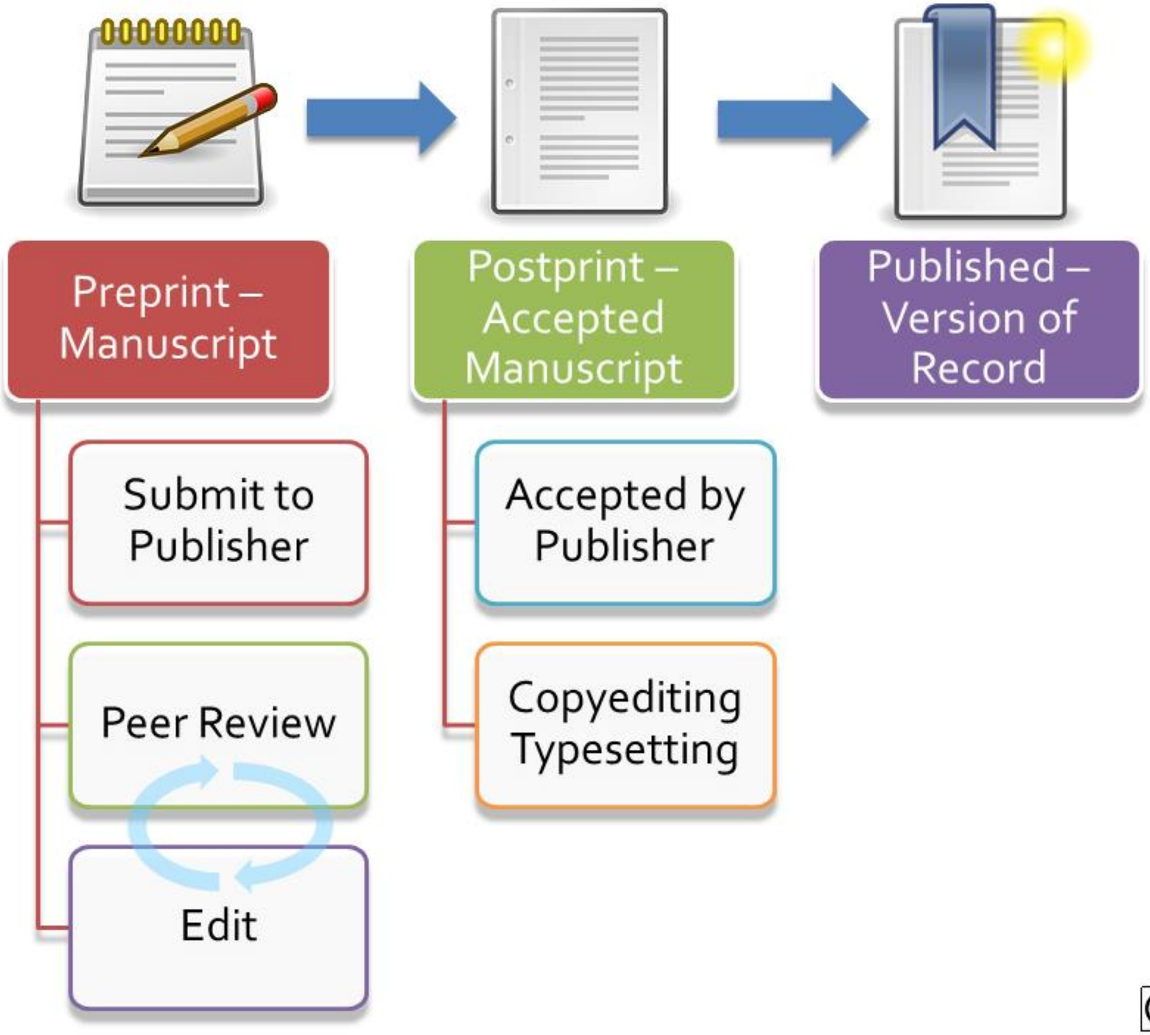
All OA, No Fees

Hybrid (OA with fees and non OA)

All OA, Fees Mandatory

https://canterbury.libguides.com/scholarly/OA

Open Access publishing: the Green Route



Preprint – Manuscript

Submit to Publisher

Peer Review

Edit

Postprint – Accepted Manuscript

Accepted by Publisher

Copyediting Typesetting

Published – Version of Record

https://canterbury.libguides.com/scholarly/OA

A potential workflow

Not yet accepted

- Pre-print for discussion on repository

Accepted

- Pre-print on repository
- Post-print on repository (potential embargo)
- Post-print on personal page
- Disseminate via social media (ResearchGate and similar)

Thanks!