

Division of Psychology and Language Sciences

Centre For Longitudinal Studies

MRC Unit for Lifelong Health and Ageing



UCL

Doing open science: transparency, integrity and inclusivity in research



Ritika Chokhani & Praveetha Patalay
Guest Lecture at Tata Institute of Social
Sciences, MA Applied Psychology
28th November 2025

Housekeeping

- Interactive – please raise hand or type questions into chat. If you have a question, it is likely that others have the same question!
- You can also interact through reactions if something resonates with you!
- Please keep yourself muted if not speaking, as it reduces background sounds for everyone
- Videos on, if possible for you, as it helps build engagement
- Take a moment to check if you have something to take notes (digital/paper)

Who are we?



Ritika Chokhani
PhD student in Mental Health
Science
University College London



Praveetha Patalay
Professor of Population Health
and Wellbeing
University College London

Tell us about you

In one or two words, type into the chat a topic you are interested in researching

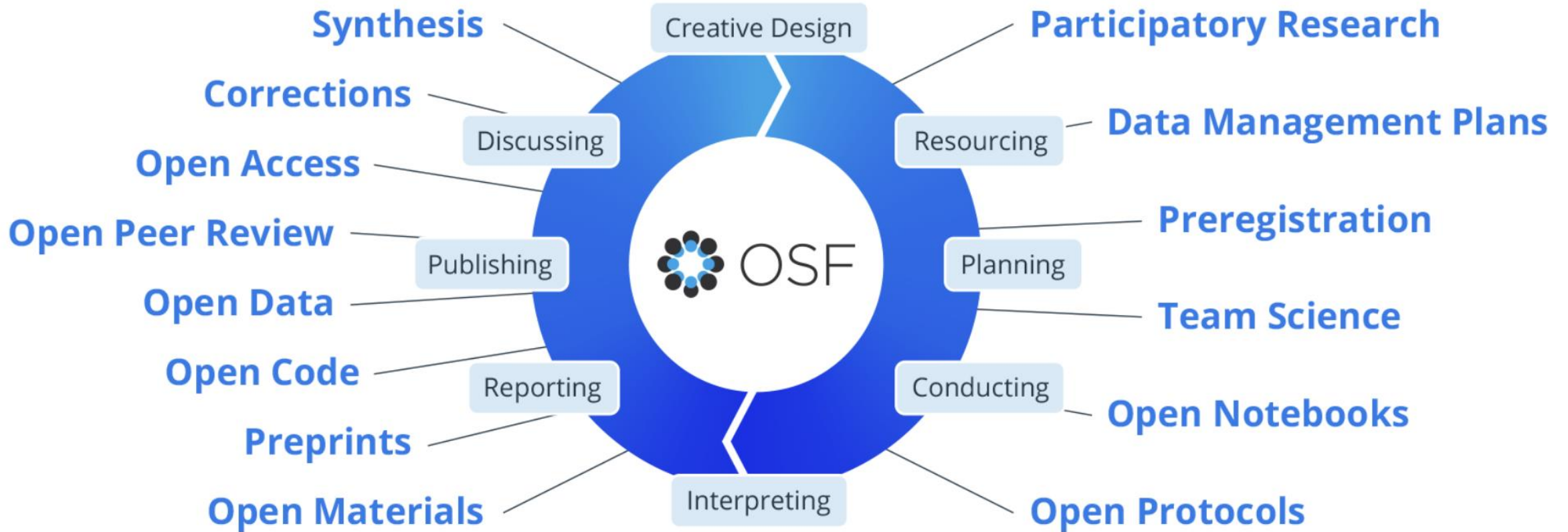
Today's plan

- What is open science?
- The need for open science
- Questionable Research Practices (QRPs)
- Communicating our science
- Inclusivity in science
- What can we do?

What is open science?

a global movement that aims to make scientific research and its outcomes freely accessible to everyone.

By fostering practices like data sharing and preregistration, open science not only accelerates scientific progress but also strengthens trust in research findings.



The need for open science

Why Most Published Research Findings Are Wrong

John P. A. Ioannidis

Published: August 30, 2005 • <https://doi.org/10.1371/journal.pmed.0020124>

SCIENCE

Daryl Bem Proved ESP Is Real

Which means science is broken.

BY DANIEL ENGBER

JUNE 07, 2017 • 2:57 PM

It seemed obvious, at first, that Jade Wu was getting punked. In the fall of 2009, the Cornell University undergraduate had come across a posting for a job in the lab of one of the world's best-known social psychologists. A short while later, she found herself in a conference room, seated alongside several other undergraduate women. "Have you guys heard of extrasensory

RESEARCH ARTICLE

Estimating the reproducibility of psychological science

OPEN SCIENCE COLLABORATION [Authors Info & Affiliations](#)

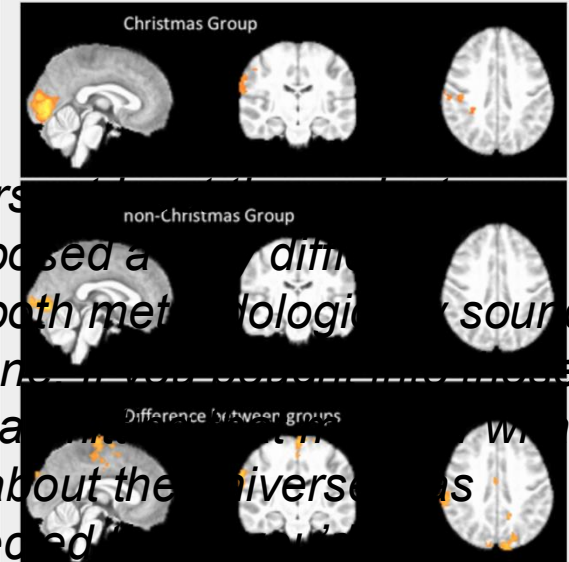
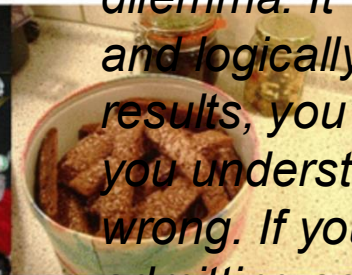
SCIENCE • 28 Aug 2015 • Vol 349, Issue 6251 • DOI: 10.1126/science.1249126

Research » Christmas 2015: All in the Mind

Evidence of a Christmas spirit network in the brain: functional MRI study

BMJ 2015 ; 351 doi: <https://doi.org/10.1136/bmj.h6266> (Published 11 December 2015)

Cite this as: BMJ 2015;351:h6266



for most observers, the paper posed a dilemma. It was both methodologically sound and logically insane. If you accepted the results, you'd be admitting something almost as momentous: that the standard methods of psychology cannot be trusted, and that much of what gets published in the field—and thus, much of what we think we understand about the mind—could be total bunk.

How citation distortions create unfounded authority: analysis of a citation network

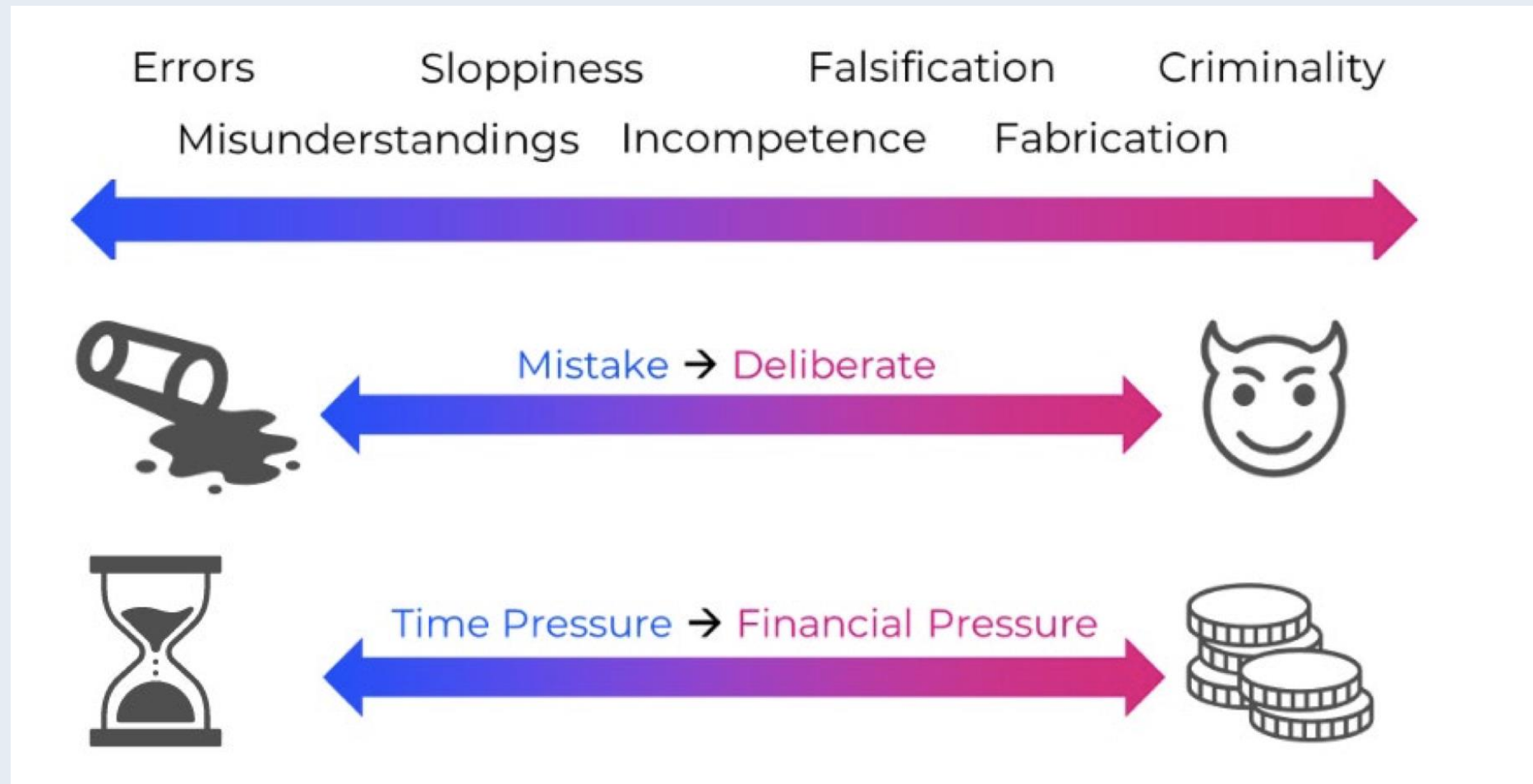
BMJ 2009 ; 339 doi: <https://doi.org/10.1136/bmj.b2680> (Published 21 July 2009)

Cite this as: BMJ 2009;339:b2680

How could this be possible?

Two broad reasons...

- Questionable Research Practices (QRPs)
- Fabrication, falsification and plagiarism



Zoom poll

Do you consider the following research practice acceptable or not acceptable?

- In a paper, not reporting all of the study's conditions and dependent variables due to word count
- Deciding to collect more data than planned because the results are not yet significant
- For a required sample of 150, a researcher was only able to collect data from 80 participants. They copy-paste the data from 80 participants twice to make it 160 participants to meet their sample size.
- To tell a clearer story in the paper, reporting an unexpected finding as having been predicted from the start
- Deciding to exclude some quotes from a qualitative study because they are discrepant from the main theme

QRP Example 1: Researcher flexibility or “p-hacking”

Researcher are human and have biases.

We want a particular result and can make decisions (in both quantitative & qualitative research) that lead to what we want.

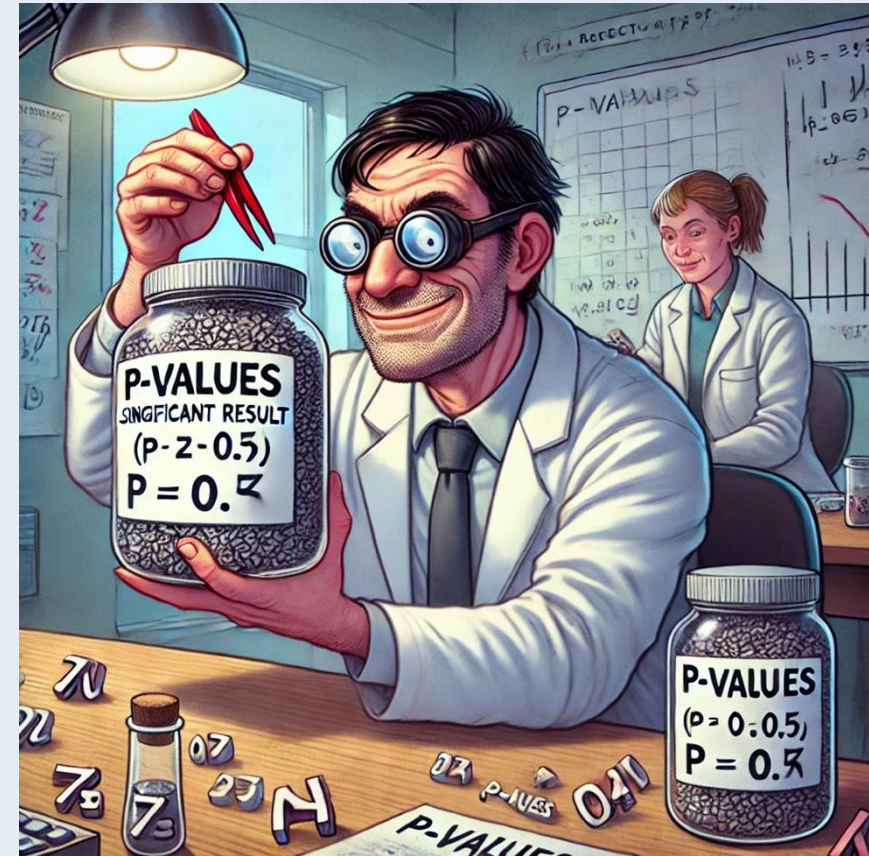
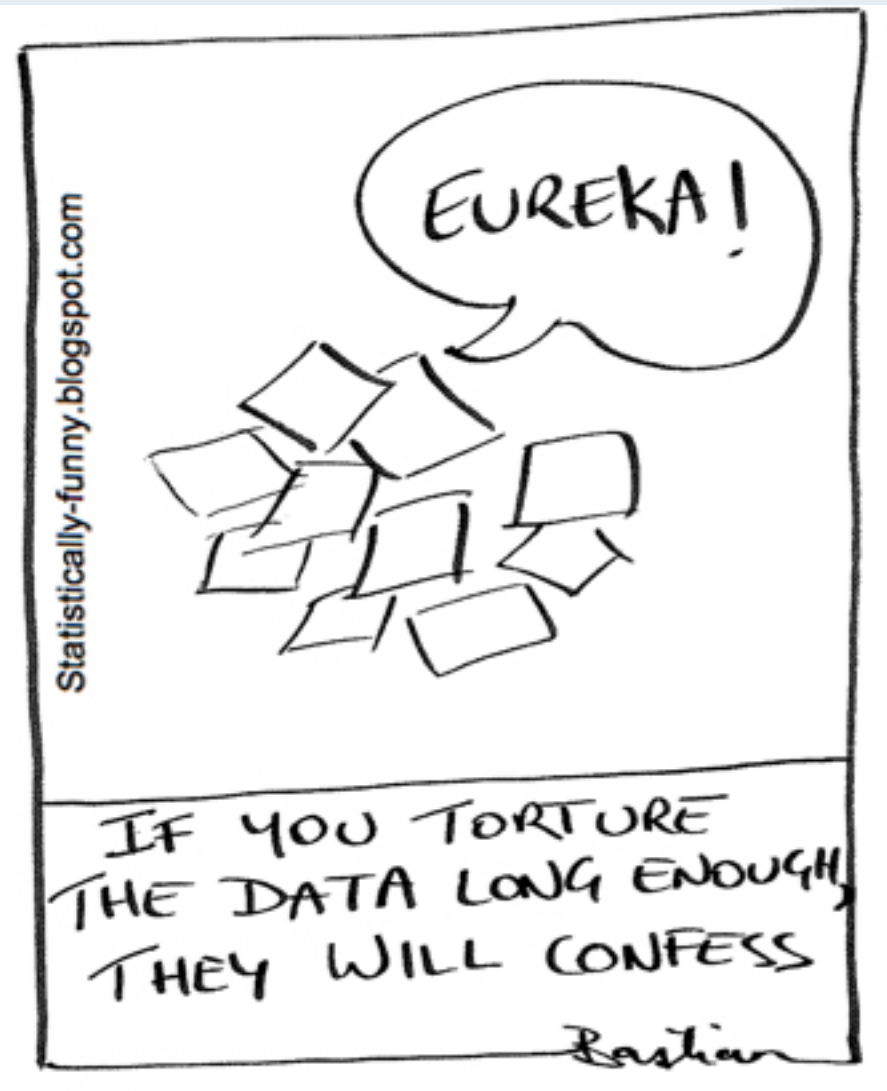
The Prevalence of Marginally Significant Results in Psychology Over Time

[Anton Olsson-Collentine](#)  , [Marcel A. L. M. van Assen](#), and [Chris H. J. Hartgerink](#)  [View all authors and affiliations](#)

[Volume 30, Issue 4](#) | <https://doi.org/10.1177/0956797619830326>

This researcher flexibility “is not the result of malicious intent, but rather the result of two factors:

- (a) ambiguity in how best to make these decisions**
- (b) researcher’s desire to find a statistically significant result”**



QRP Example 2: HARKing (Hypothesizing after results are known)

Initial hypothesis: Psychotherapy will lower the average symptoms of depression that people report.

Results: I find that the average symptoms on the Beck Depression Inventory do not show a change, but there is a change in the functional impairment scales I am using.

Changed hypothesis: I attempted to study whether psychotherapy has a positive effect on levels of functional impairment for people who are depressed.

QRP Example 3: Low power

- Imagine that childhood trauma increases the risk of depression
- You conduct a study with 50 people asking them questions about trauma & depression
- But you find there's no association between trauma & depression
- Is there really no association or was your power too low?

Why should I care about power?

- If a study has 20% power, it means that if 100 similar studies are done, only 20 of those will discover the true effect; the other 80 will falsely show a null effect
- Recently, Button and colleagues showed that the average power of neuroscience studies is probably around 20%
- Low power matters because:
 - You (and your participants) might spend a lot of time and effort & find a null effect (when actually there is a significant relationship)
 - You have a higher chance of finding a significant false effect (“winner’s curse”)
 - If an underpowered study finds a null or significant effect that is false, you are adding something into the literature that’s incorrect

Fabrication, falsification & plagiarism

RETRACTED: Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end

Lisa L. Shu, Nina Mazar , Francesca Gino, , and Max H. Bazerman [Authors Info & Affiliations](#)

Edited* by Daniel Kahneman, Princeton University, Princeton, NJ, and approved July 23, 2012 (received for review August 27, 2012)

109 (38) 15197-15200 | <https://doi.org/10.1073/pnas.1209746109>

THIS ARTICLE HAS BEEN RETRACTED + VIEW RELATED CONTENT +

	A	B	C	D	E	F	G	H	I
1	P#	Cond	Stude	Major	CS3	Male	Age	#B	\$B
47	35	1	1	Journalism	3	1	19	12	12
48	37	1	1	Economics	4	0	21	9	9
49	40	1	1	Political Science	5	1	29	15	15
50	42	1	1	Political Science	3	0	20	7	7
51	46	1	1	Political Science	4	0	21	12	12
52	49	1	1	English	4	1	21	9	9
53	49	1	1	English	4	1	21	7	7
54	55	1	1	Biology	4	1	21	12	12
55	58	1	1	Environmental Sciences	3	0	20	10	10
56	61	1	1	Nursing	3	0	20	15	15
57	63	1	0	NA		0	22	12	12
58	68	1	1	Business	3	1	20	16	16
59	70	1	1	Chemistry	4	0	21	11	11
60	73	1	1	Chemistry	5	0	20	16	16
61	76	1	1	Chemistry	2	1	19	15	15
62	80	1	1	Nursing	4	0	21	15	15
63	82	1	1	Economics	4	1	21	9	9
64	85	1	1	Psychology	4	0	20	5	5
65	88	1	1	Chemistry	3	0	20	13	13
66	95	1	1	Math Education	3	1	22	13	13
67	51	1	0	NA	0	0	52	4	4
68	12	1	1	Psychology	3	0	20	13	13
69	101	1	0	Business	3	1	20	6	6
70	7	2	0	Political Science	5	1	22	17	17
71	91	2	1	Japanese	2	1	20	17	17
72	52	2	0	NA	5	0	22	8	8
73	5	2	1	Biology/Psychology	2	0	18	16	16
74	8	2	1	Communication Studies	4	0	22	15	15
75	13	2	1	Chemistry	4	0	20	18	18
76	17	2	1	Communications	4	0	21	14	14

Plagiarism

AI is complicating plagiarism. How should scientists respond?

The explosive uptake of generative artificial intelligence in writing is raising difficult questions about when use of the technology should be allowed.

Why wouldn't I use ChatGPT to write a paper?

It's a matter of self-respect.

I believe research and writing are valuable activities.

Writing helps me make sense of the world.

I don't want to contribute to a harmful industry.

I don't want to betray others' trust.

I can write better than ChatGPT.

When might I use ChatGPT to write a paper?

I don't have enough time before the deadline.

I feel I can't use my own words well or the "original says it so well"

I'm not sure how to reference or cite

Left side points from
Miriam Posner, Bluesky

Research fraud & paper mills

Usual terminology	Tortured phrase
Breast cancer	Bosom peril
p value	p esteem
deep learning	profound learning
Mean squared error	Mean squared blunder
Random forest	Irregular timberland
Big data	Huge information
Fatty acids	Blubbery acids
Smart city	Shrewd city
Thermal stress	Warm anxiety

Q&A + Break for 10 min

Communicating our science

Scientific publishing and dissemination

Communicating our findings

- Communicating scientific findings is important
- Many ways of communicating e.g. blogs, podcasts, videos, infographics
- One way is through a peer-reviewed journal article
 - Taken as index for many promotions & jobs in academia
 - The process of peer-review is (theoretically) supposed to be very rigorous

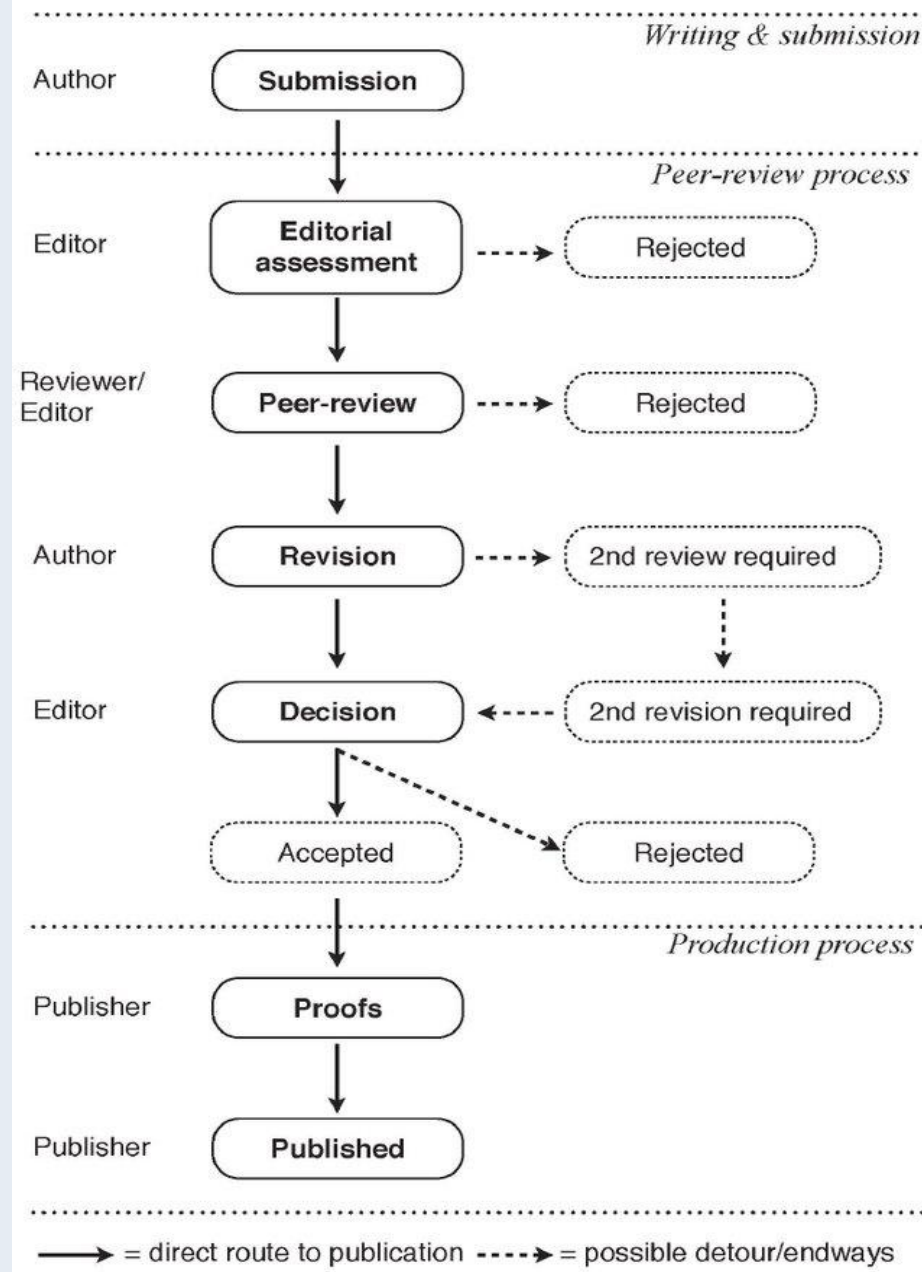


Figure from Tress et al. 2014

Journal publication considerations

- Open Access vs subscription journals: <https://www.jisc.ac.uk/open-policy-finder>
- Usually have to pay an article processing charge for open access which can be reduced or waived for authors from low & middle income countries

Problem 1. Incentive to publish only significant results

- Journals want to/are perceived to want to publish positive results (treatment works!)
- People are more interested in reading positive results

This leads to two outcomes:

- Negative results don't get published, leading to an overinflation of positive results
- Researchers are more incentivised to get positive results (and more risk of p-hacking etc)

*A null finding (finding that what you hypothesized was NOT true)
is as good a scientific result as a significant finding.*



SPECIAL ARTICLE



Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy

Authors: Erick H. Turner, M.D., Annette M. Matthews, M.D., Eftihia Linardatos, B.S., Robert A. Tell, L.C.S.W., and Robert Rosenthal, Ph.D. [Author Info & Affiliations](#)

Published January 17, 2008 | N Engl J Med 2008;358:252-260

DOI: 10.1056/NEJMsa065779 | VOL. 358 NO. 3 | Copyright © 2008

You can publish negative results!

The impact of treating parental anxiety on children's mental health: An empty systematic review

[Laura Chapman](#), Rosie Hutson, Abby Dunn, Maddy Brown, Ella Savill, Sam Cartwright-Hatton

[Bristol Medical School \(PHS\)](#)

Research output: Contribution to journal > Article (Academic Journal) > peer-review

16
Citations
(Scopus)

196
Downloads
(Pure)



Overview



Fingerprint

Problem 2. Predatory journals and fallible peer review practices

Science

[Current Issue](#)

[First release papers](#)

[Archive](#)

[About](#) ▼

[Submit manuscript](#)

[HOME](#) > [SCIENCE](#) > [VOL. 342, NO. 6154](#) > [WHO'S AFRAID OF PEER REVIEW?](#)

 | [SPECIAL ISSUE NEWS](#)



Who's Afraid of Peer Review?

[JOHN BOHANNON](#) [Authors Info & Affiliations](#)

SCIENCE • 4 Oct 2013 • Vol 342, Issue 6154 • pp. 60-65 • [DOI: 10.1126/science.342.6154.60](#)

 29,951  556





Predatory Journals

- <https://beallslist.net/>
- Accepting articles with little or no peer-review
- Hiding information about Article Processing Charges
- Misrepresenting members of the journal's editorial board
- Other violations of copyright or scholarly ethics

Post-publication peer review

- <https://pubpeer.com/>

Inclusivity in science

Inclusivity: diverse opinions and stakeholders

- This could include:
- People with lived experience of the condition/issue you are studying
- People who work with the relevant groups (e.g. clinicians, teachers, etc)

Inclusivity: Samples and populations

- Is the sample representative of the target population?
- Understudied groups and populations
- E.g. in Psychology most published findings are based on studies in Western, High-income countries
- Even within countries, often participants recruited (especially in convenience samples), are not representative of the target group of interest; for instance, people from certain social backgrounds, castes, religions might be under-represented.

So what can we do?

1. Prioritizing quality over quantity

- Hard to resist pressures: pressure from supervisor; deadlines (e.g., thesis submission)
- When you are in a research career, there are incentive structures in place to motivate you to increase publications
- Things are always going to take more time than you anticipated

BUT

- Taking the time to think about research question and research design is important
- Quality is ultimately better from a career perspective as well (e.g. some students may feel pressured they have to get a "publication" but publishing something of quality is better than publishing something in a predatory journal)

2. Collaborate

- Even as a student, you can team up with another student to collect data.
 - E.g. you are interested in measuring “emotional intelligence” and your peer is interested in measuring “depression”. You could use the same sample and ask them to fill out one questionnaire on each.
 - Collaborate on systematic reviews / scoping reviews for screening
 - Peer-led teaching e.g., each person brings a question they have about their analysis
- In an organization, can consider asking to be part of another ongoing research project.
 - Advantages for data collection: larger sample size, more possibilities
 - Advantages for analysis: learning, analysis checks

3. Share what you can

- Data sharing. Can share on Open Science Framework (OSF) or on specific repositories.
- Code & coding sharing (e.g. how you coded qualitative transcripts)
- There are guidelines on best practice (e.g. FAIR)

4. Transparency in reporting

Article | [Open access](#) | Published: 20 October 2020

RETRACTED ARTICLE: A mechanistic model of the neural entropy increase elicited by psychedelic drugs

[Rubén Herzog](#) , [Pedro A. M. Mediano](#), [Fernando E. Rosas](#), [Robin Carhart-Harris](#), [Yonatan Sanz Perl](#), [Enzo Tagliazucchi](#) & [Rodrigo Cofre](#)

[Scientific Reports](#) **10**, Article number: 17725 (2020) | [Cite this article](#)

74k Accesses | **133** Altmetric | [Metrics](#)

 This article was [retracted](#) on 15 September 2022

 This article has been [updated](#)

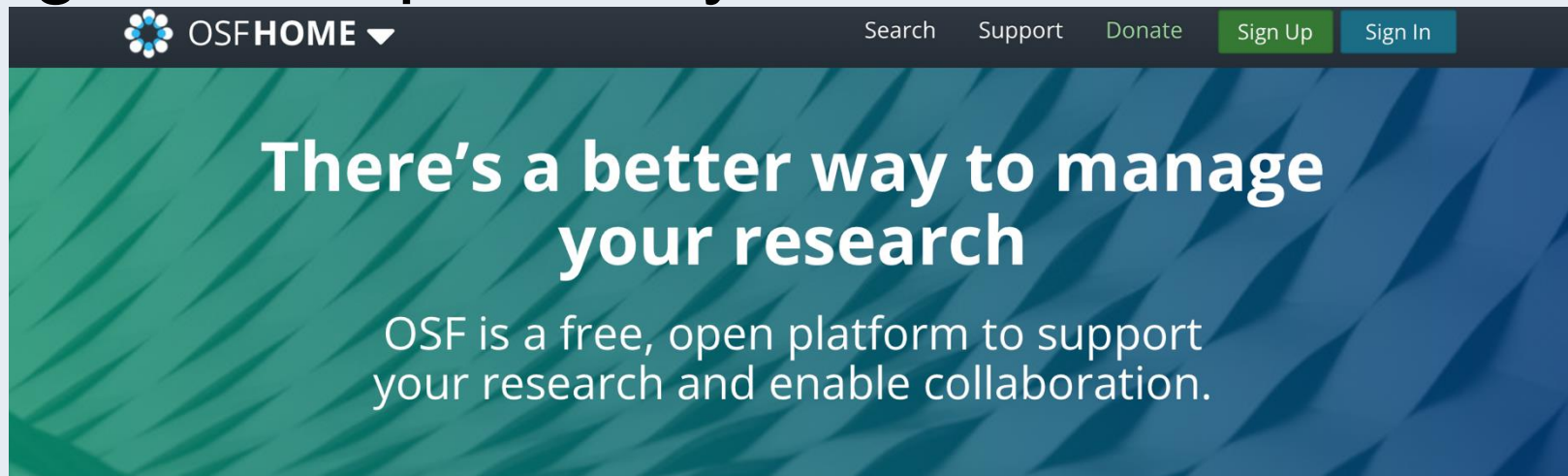
4. Transparency in reporting cont.

- Who is the target population that your sample actually provides information on?(e.g. if you sampled people in your network based in Mumbai, can you generalise to all Indians?)
- How the sampling approach might bias your findings?
 - Add this to your discussion and limitations

5. Preregistration

Say what you are going to do before you do it: research question, hypotheses, designs, methods

One of the advantages is it forces you to really clarify your research question, design and method e.g., articulate the reasons for choosing the sample size you have chosen



Qualitative preregistration

OSF heading	Subsections	Possible modifications/extensions
Study information		
	Title	
	Authorship	
	Research aim	Please specify the overall aim of the research.
	Research questions	Research questions (subject to modifications at <i>n</i> moments). <i>Typical changes in exact phrasing of research questions may occur during the process; after the first instance(s) of data collection, etc.</i>
	Hypotheses	Theoretical expectations If you have any expectations (at the start of the study), please write them here.
	Use of literature	Please specify the role of theory in your research design. Elaborate on how you used literature to formulate your research question and how you expect the theory to guide your data collection and data analysis (for example elaborate on your sensitizing concepts).
	Use of literature rationale	Please elaborate if your research is conducted from a certain theoretical perspective.
Design Plan		
	Tradition	Please specify the type of tradition you work in: - grounded theory - phenomenology - narrative approach - ethnography - text-based approach (discourse analysis, conversation analysis) - generic - other
	Study type	Specify your study type (select multiple if appropriate): - case study - evaluation research - intervention research - participatory research - other

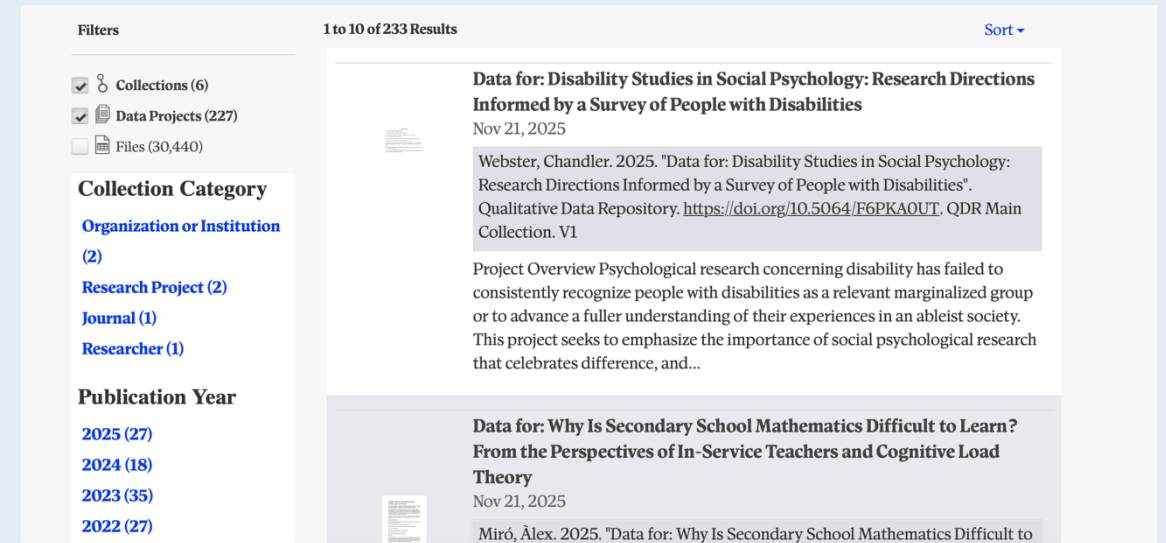
Using secondary instead of primary data in research

Data repositories

- UK Data Service:
<https://datacatalogue.ukdataservice.ac.uk>



- Qualitative Data Repository:
<https://data.qdr.syr.edu>



Indian data

- Young Lives study on UK Data service
- National Family Health Survey Data
- Government data initiatives: <https://www.data.gov.in/>
- <https://ndap.niti.gov.in/help?tab=getting-started>
- Other initiatives: <https://www.devdatalab.org/#research>

Thank you for listening

For questions or comments, contact:

ritika.chokhani.14@ucl.ac.uk or p.patalay@ucl.ac.uk