



Open Science

0 perspectivă Germană

George Gunnesch-Luca

2023-12-04

[~]\$ WHOAMI? #ROOT

- 39.657 Studenți
- 837 Doctorate pe an
- 50 Abilitări
- 6570 Angajați
- Buget de 243 Milioane Euro



Un scurt istoric cu *iz* german

Publish or Perish, una din problemele sistemului

Încercări de remediere

Concluzie?

■ Ioannidis, 2005.

“Marea majoritate a rezultatelor de cercetare sunt false pentru majoritatea proiectelor de cercetare și pentru majoritatea domeniilor.”

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

Summary

There is increasing concern that most current published research findings are false. The probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and importantly, the ratio of true to no relationships among the relationships posited in each scientific field. In this framework, a research finding is less likely to be true when the studies conducted in a field are smaller, when effect sizes are smaller, when there is a greater number and lesser protection of tested relationships, where there is greater flexibility in designs, definitions, outcomes, and analytical models, when there is greater financial and other interest and prejudice and when more teams are involved in a scientific field in chase of statistical significance. Simulations show that for most study designs and settings, it is more likely for a research claim to be false than true. Moreover, for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias. This essay reviews the implications of these problems for the conduct and interpretation of research.

Published research findings are sometimes refuted by subsequent evidence, with ensuing confusion and disappointment. Retraction and controversy is seen across the range of research designs, from clinical trials and traditional epidemiological studies [1–3] to the most modern molecular research [4,5]. There is increasing concern that in modern research, false findings may be the majority or even the vast majority of published research claims [6–8]. However, this should not be surprising. It can be proven that most claimed research findings are false. Here I will examine the key

The Essay section contains opinion pieces on topics of broad interest to a general medical audience.

factors that influence this problem and some corollaries thereof.

Modeling the Framework for False Positive Findings

Several methodologies have pointed out [9–11] that the high rate of nonreplication (lack of confirmation) of research discoveries is a consequence of the convenience, yet ill-founded strategy of claiming conclusive research findings solely on the basis of a single study assessed by formal statistical significance, typically for a p -value less than 0.05. Research is not most appropriately represented and summarized by p -values, but, unfortunately, there is a widespread notion that medical research articles

It can be proven that most claimed research findings are false.

should be interpreted based only on p -values. Research findings are defined here as any relationship reaching formal statistical significance, e.g., effective interventions, informative predictors, risk factors, or associations. “Negative” research is also very useful. “Negative” is actually a misnomer, and the misinterpretation is widespread. However, here we will target relationships that investigators claim exist, rather than null findings.

As has been shown previously, the probability that a research finding is indeed true depends on the prior probability of it being true (before doing the study), the statistical power of the study, and the level of statistical significance [10,11]. Consider a 2×2 table in which research findings are compared against the gold standard of true relationships in a scientific field. In a research field both true and false hypotheses can be made about the presence of relationships. Let R be the ratio of the number of “true relationships” to “no relationships” among those tested in the field. R

is characteristic of the field and can vary a lot depending on whether the field targets highly likely relationships or searches for only one or a few true relationships among thousands and millions of hypotheses that may be postulated. Let us also consider, for computational simplicity, circumscribed fields where either there is only one true relationship (among many that can be hypothesized) or the power is similar to find any of the several existing true relationships. The pre-study probability of a relationship being true is $R/(R+1)$. The probability of a study finding a true relationship reflects the power $1 - \beta$ (one minus the Type II error rate). The probability of claiming a relationship when none truly exists reflects the Type I error rate, α . Assuming that relationships are being probed in the field, the expected values of the 2×2 table are given in Table 1. After a research finding has been claimed based on achieving formal statistical significance, the post-study probability that it is true is the positive predictive value, PPV. The PPV is also the complementary probability of what Wacholder et al. have called the false positive report probability [10]. According to the 2×2 table, one gets $PPV = (1 - \beta)R / (R - \beta R + \alpha)$. A research finding is thus

Creative Commons Attribution License (CC BY): This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright: © 2005 John P. A. Ioannidis. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abbreviations: PPV (positive predictive value)

John P. A. Ioannidis is in the Department of Hygiene and Epidemiology, University of Ioannina, School of Medicine, Ioannina, Greece, and Institute for Clinical Research and Health Policy Studies, Department of Medicine, Tufts-New England Medical Center, Tufts University School of Medicine, Boston, Massachusetts, United States of America. E-mail: ioannidis@hsph.harvard.edu

Competing Interests: The author has declared that no competing interests exist.

DOI: 10.1371/journal.pmed.0020124

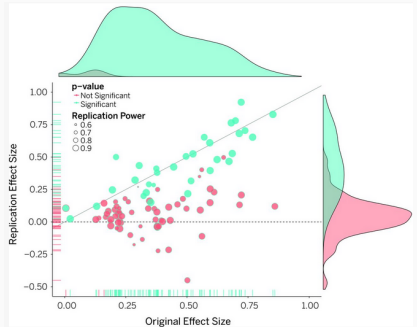
Rata mică de replicabilitate nu este o funcție a unui domeniu specific...

- Cercetare fundamentală oncologică (46%; Errington et al. 2021)
- Fizică cuantică (% neclar. Ex: fermionii Majorana; Frolov, 2021)
- Economie (*își merită soarta!*: <37%; Herbert et al. 2023)

Psihologia: deja ultra-cunoscut

- ManyLabs, Nosek, 2015

Replicarea a 100 de studii din psihologie; Doar 36% dintre ele cu succes.



Original study effect size versus replication effect size (correlation coefficients).

Diagonal line represents replication effect size equal to original effect size. Dotted line represents replication effect size of 0. Points below the dotted line were effects in the opposite direction of the original. Density plots are separated by significant (blue) and nonsignificant (red) effects.

Fiske și terorismul metodologic...

- **Congresul DGPs, Leipzig, 20 Septembrie 2016;** Keynote de Brian Nosek: "Addressing the Reproducibility of Psychological Science"
- **Susan Fiske, atac preventiv...**

Scrisoarea pentru APS Observer, a fost "scursă" înainte de termen pe Dropbox. Fiske acuză pei cei care au criticat public metodele de cercetare psihologică de intimidare publică și ceea ce ea numește "terorism metodologic".



Andrew D Wilson
@PsychScientists

Susan Fiske: 'Mob Rule or Wisdom of Crowds?' (Hint: she doesn't think we're wise) dropbox.com/s/9zubn9fyitx...

addressed by APS President Susan Fiske (2016): "New forms of media are making it easier and easier for us to react to, and comment on, research within our community. Although live (during) comment and criticism can often push an argument or research program forward in a good direction, they can also derail, and perhaps even threaten, the process. I invited guest columnist Susan Fiske, a former APS president, to think about the impact that the new media are having on us only on our science, but also on our scientists."

Susan T. Fiske
APS Past President

Our field has always encouraged - required, really - peer critiques. But the new media (e.g., blogs, Twitter, Facebook posts) are encouraging uncensored, unfettered trash talk. In the most extreme examples, online vigilantes are attacking individuals, their research programs, and their careers. Self-appointed data police are volunteering critiques of each personal finding and inflicting frequency that they resemble a denial-of-service attack that crashes a website by sheer volume of traffic.

Only what's crashing are people. These unmoderated attacks create collateral damage to targets' careers and well-being, with no accountability for the bullies. Our colleagues at all career stages are leaving the field because of the sheer adversarial viciousness. I have heard from graduate students opting out of academia, assistant professors afraid to come up for tenure, and career people wondering how to protect their jobs, and senior faculty retiring early, all because of methodological terrorism. I am not naming names because all human error tactics are already damaging our field. Instead, I am describing a dangerous minority trend that has an outward impact and a chilling effect on scientific discourse. I am not a primary target, but my goal is to give voice to others too sensible to object publicly.

To be sure, constructive critics have a role, with their rebuttals and letters-to-the-editor subject to editorial oversight and peer review for tone, substance, legitimacy. Some moderated social-media groups monitor individual posts to ensure they are appropriate, always, of course, of course, of course, to write a personal letter to the editor, that's their business. If they request the original data, scientific norms demand delivery within reasonable constraints. And these venues respect the target.

often in private with a chance to improve (peer review), or at least in moderate, scholastic (curated comments and rebuttals). These venues offer constraints, education, open discussion, and quality control. These constructive efforts draw on the volunteer talent of many, in the greater good and respecting the individual investigator.

In contrast, the self-appointed destructive critic's role now includes public shaming and blaming, often implying dishonesty on the part of the target and other unnamed users on unchecked assumptions. Targets often seem to be chosen for scientifically irrelevant reasons: their contrary opinions, professional prominence, or career path vulnerability.

The destructive critics are ignoring ethical rules of conduct because they circumvent constructive peer review. They attack the person, not just the work; they attack subtly, without quality controls; they have sent their unsolicited, unwelcome attack a tenure-review committee and public-speaking sponsors; they have implicated targets' family members and advisors. But all self-appointed critics behave maliciously, and some do no more than others. One hopes that these critics aim to improve the field, not harm people. But the fact is that these vigilante critiques are harming people. They are a far cry from peer-reviewed critiques, which serve science without destroying lives.

Ultimately, science is a community, and we are in it together. We agree to abide by scientific standards, ethical norms, and mutual respect. We treat but verify, aim sincere improves in the process. Psychological science has achieved much through collaboration, but also through responding to constructive adversaries who ask their critiques respectfully. The key word here is constructive.

APS Past President Susan T. Fiske is Eugene Higgins Professor of Psychology and Professor of Public Affairs at Princeton University. Her research focuses on bias, stereotyping, prejudice, and discrimination are influenced by cooperation, competition, power, and social relationships. She is a recipient of the 2017 APS Anne McKee Cattell Fellow Award.

1:00 AM · Sep 20, 2016

Urmările scandalului pe termen scurt

- Impact inițial emoțional major în Germania (*DGPs* fiind și locul unde poți detona bomba)
- O tentă ușor revanșardă și aplauze...
- Dar pe termen lung?
- Care sunt cauzele?

Iată câteva, deloc exhaustiv:

- Practici științifice îndoielnice (HARKing, selective reporting, șamd)
- Studii caracterizate de putere mică
- Metode nepotrivite (NHST)
- Incapacitatea de a implementa reforme în structura educațională (Bayes)
- Dar și *Zero Accountability* Ca dovadă:
 - câte reviste au fost desființate?
 - și câte publică acum replicări esențiale pentru revizuire de teorii?
 - **Journal of Consumer Research** (revista de top în domeniul comportamentului consumatorului) are **acum** o rată de replicare de 3,4%. Care este sensul existenței acestei reviste?

Ca idee, câte tone la hectar produce Germania?

Universitate	WoS articles
Heidelberg Univ.	72,556
LMU Univ.	67,525
Charité Berlin Univ.	63,949
Technical Munich Univ	63,641
Bonn Univ.	54,671
Nuremberg Univ.	53,289
Karlsruhe Univ.	51,266
Hamburg Univ.	48,880
Freiburg Univ.	47,574
Technical Dresden Univ.	47,137

2010 - 2017

Cine e Publish or Perish?

Băi, cine mai e, mă, și Caniculă ăsta, că am văzut la televizor că face victime multe, omoară oameni?" -

Claudiu Răducanu



Deținători de doctorat în Germania

- 862.000 de germani dețineau grade de doctorat (2019)
- 1,2% din populație
- 28.153 de finalizări reușite ale doctoratelor (2021)
- Tendință ușor descrescătoare începând din 2016

Cine lucrează în universități?

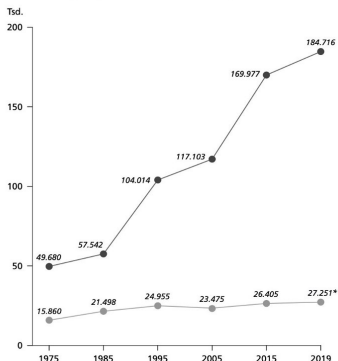
- Aproximativ 190.000 de oameni lucrează în universități
- Activi în predare, cercetare
- 28.000 de profesori în 2020

Bun, dar cine produce toate erorile astea de tip I?

Evoluția personalului academic: Germania (1976-2019)

- Deteriorarea condițiilor de muncă în ultimii 20 de ani.
- Contracte pe termen limitat sunt normă nu și excepție (*practica ilegală pe piața liberă în Germania*)
- 6 ani pt Dr./ 6 ani postDoc (cu excepții pentru Proiecte de cercetare) conform *Wissenschaftszeitvertragsgesetz* (WissZeitVG)
- Precariat științific (**dual labor market**)

Abbildung 2 Akademisches Personal und Professuren an Universitäten und gleichgestellten Hochschulen im Vergleich, 1975–2019



— Dozentinnen und Dozenten, Assistentinnen und Assistenten, wissenschaftliche und künstlerische Mitarbeiterinnen und Mitarbeiter
— Professorinnen und Professoren

* davon 1.212 Juniorprofessorinnen und -professoren ohne Tenure Track

Source: Wissenschaftler auf Zeit (Max Planck Institut 2020)

https://www.mpifg.de/pu/mpifg_dp/2020/dp20-15.pdf

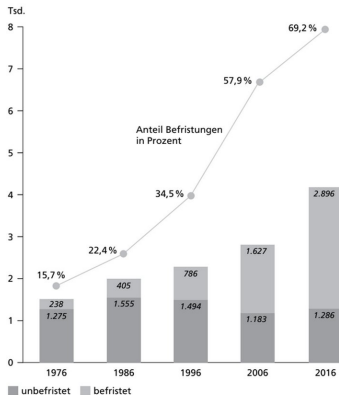
Publish or Perish 4 Real - pentru cei nefamiliarizați

- Sistemul de catedră
 - Libertate academică limitată: Personalul tânăr fara libertate de a-și urmări propriile interese de cercetare.
 - Dependență: Progresul în carieră este adesea strâns legat de profesorul care deține catedra.

Bonus

- Reforma din **Martie 2023**: Scurtarea perioadei de post-Doc la 3 ani! yay!

Abbildung 1 Entwicklung des wissenschaftlichen Personals in der MPG, 1976–2016



Source: Wissenschaftler auf Zeit (Max Planck Institut 2020)
https://www.mpg.de/pu/mpi-fg_dp/2020/dp20-15.pdf

Unele încercări de remediere...

În acest context, practicile Open Science devin o problemă pentru că standardele sunt vechi:

- Deciziile de angajare la universități sunt influențate de cantitatea de publicații și prestigiul jurnalului.
- Multe universități se bazează pe valori prag ale indexului h și numărul de articole în jurnale “cu impact mare”.
- Percepția predominantă: numărul de publicații și prestigiul jurnalului sunt factori-cheie pentru succesul profesional.
- Focalizarea pe cantitate și rezultate perfecte duce la scăderea valorii de adevăr a cunoștințelor noastre comune.

Open Science încetinește procesul

1. Efectuarea studiilor cu o putere statistică ridicată, pregătirea datelor și încercarea de a publica rezultate realiste care nu sunt modificate pentru a ajunge la o perfecțiune (nerealistă)
2. Cercetătorii care se angajează în aceste bune practici de cercetare probabil vor avea un număr mai mic de publicații, iar dacă acesta este criteriul principal de selecție, vor avea un dezavantaj pe o piață competitivă a locurilor de muncă pentru poziții titularizate.

Schimbarea standardelor de angajare.

- LMU München (2016): prima universitate care schimbă standardele de angajare.
- Trier, Köln în 2017; O bună parte din ele până în 2023.
- FAU din 2022

(listă completa aici)

Unele încercări de remediere...

Open Science Policies

Organizație	Document
Germany	Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities
The Federal Ministry for Education and Research	Open Access in Germany
Leibniz Association	Open Access Guidelines
Research Foundation (German Research Foundation)	Guidelines for the Use of Funds
Helmholtz Association	Open Access Publications Popularization
Fraunhofer	Open Access Publications Popularization
German Research Foundation	German Research Foundation Document
Max Planck Society	Principles for the Handling of Research Data

Open Science Policies

- Idealizări uneori, dar cu putere simbolică
- FAU, semnatară și cu propria politică suplimentară: *“When it comes to the accessibility of data, FAU follows the principle: ‘as open as possible, but as closed as necessary.’”*
- Se axează foarte mult pe open access și nu pe open input sau process, dar asta este o problema europeana si nu specific germană (Moradi, Abdi, 2023):
 - De exemplu: Open Peer Review?

Unele încercări de remediere...

Open Access, DEAL

- Sume mari investite;
- Project DEAL

Wiley: - PAR (publish and read) pentru publicarea în jurnalele hibride:
2750 de euro per articol

- APC (article processing charge) pentru publicarea în jurnalele de acces deschis de tip aur: reducere de 20%

Springer Nature:

- Taxele PAR/publish and read pentru articolele de cercetare: 2.750 de euro
- Taxele PAR/publish and read pentru articolele non-research: 917 de euro
- Discount de 75% la prețul de listă pentru edițiile print adiționale ale revistelor respective
- Prețurile aici: (Open Access in Germania)

Concluzie: Mai e mult până departe...

- Open reviews...
- Data-sharing este încă o raritate
- Publish or Perish nu va dispărea rapid
- Și multe altele