Framework for
Open and
Reproducible
Research Training

FORRT

Integrating principles of open and reproducible science into higher education and raising awareness of its pedagogical implications.

Summaries of Open and Reproducible Literature

Purpose

The purpose of these summaries is to reduce some of the burden on educators looking to incorporate open and reproducible research principles into their teaching. These crowdsourced and community-curated resource also aims to satisfy three of FORRT's Goals:

❖ Support scholars in their efforts to stay up-to-date on best practices regarding open and reproducible research;
❖ Facilitating conversations about the ethics and social impact of teaching substantive topics with due regard to scientific openness, epistemic uncertainty and the credibility revolution;
❖ Foster social justice through the democratization of scientific educational resources and its pedagogies.

Table of Contents

The ◈ symbol stands for non-peer-reviewed work.

Trust Your Science? Open Your Data and Code (Stodden, 2011)◈

Main Takeaways

- Computational results suffer from problems of errors in final published conclusions.
- Release scripts and data files, graphical user interface to allow independent replication and reproducible work.
- Repeatability is the sensitivity of results when underlying measurements are re-taken.
- Standards for code quality: More precise definitions of verification, validation, and error quantification in scientific computing.
- Research workflow involves changes made to data, including analysis, that affects data interpretation.
- To conclude, open data is a prerequisite for verifiable research.

Quotes

➢ "*Science has never been about open data per se, but openness is something hard fought and won in the context of reproducibility*" (p. 22).

Abstract

This is a view on the reproducibility of computational sciences by Victoria Stodden. It contains information on the Reproducibility, Replicability, and Repeatability of code created by the other sciences. Stodden also talks about the rising prominence of computational sciences as we are in the digital age and what that means for the future of science and collecting data.

APA Style Reference

Stodden, V. C. (2011). Trust your science? Open your data and code. https://doi.org/10.7916/D8CJ8Q0P

You may also be interested in

➔ Attitudes Toward Open Science and Public Data Sharing: A Survey Among Members of the German Psychological Society (Abele-Brehm et al., 2019)

➔ Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results (Wicherts et al., 2011)

➔ Open Data in Qualitative Research (Chauvette et al., 2019)

➔ CJEP Will Offer Open Science Badges (Pexman, 2017)

➔ Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency (Kidwell et al., 2016)

- ➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)
- ➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)
- ➔ Seven Easy Steps to Open Science: An Annotated Reading List (Crüwell et al., 2019)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)
- ➔ Using OSF to Share Data: A Step-by-Step Guide (Soderberg, 2018)

Publishing Research With Undergraduate Students via Replication Work: The Collaborative Replications and Education Project (CREP; Wagge et al., 2019)

Main Takeaways

- The Collaborative Replications and Education Project (CREP) allows undergraduates to participate in high-quality direct replication, using existing resources and providing structure for research projects.
- CREP samples seminal papers in 9 sub-disciplines published 3 years before the present year. Then alumni students rate papers based on time and level of interest.
- CREP teaches good scientific practices with direct replications using open science methods.
- CREP tells original authors of study selections and asks for materials and guidance for replication.
- The skills acquired from CREP can be applied to non-academic careers. For instance, teaching students the ability to evaluate scientific claims.
- CREP provides a forum and a community for replication results to be presented, the institutionalization of replications, thereby contributing to science.
- Students are invited to contribute to authorship, even if they do not involve lead authorship roles.
- CREP deems that most student projects are not adequately powered for publication, thus do not lead to publication.
- Working with CREP allows students to replicate/not replicate a seminal finding but also to provide them a publication.

Quotes

➢ "*CREP offers a supportive entry point for faculty…new to open science and large-scale collaboration…helps with fidelity and quality checks…eliminates need for instructors to vet every hypothesis and design for student research projects…not be experts in a topic…do not need to learn new programs…documentable experience blending teaching, scholarship, and close mentoring.*" (p. 4).

Abstract

The Collaborative Replications and Education Project (CREP; http://osf.io/wfc6u) is a framework for undergraduate students to participate in the production of high-quality direct replications. Staffed by volunteers (including the seven authors of this paper) and incorporated into coursework, CREP helps produce high-quality data using existing resources and provides structure for research projects from conceptualization to dissemination. Most notably, student research generated through CREP make an impact: data from these projects are available for meta-analyses, some of which are published with student authors.

APA Style Reference

Wagge, J. R., Brandt, M. J., Lazarevic, L. B., Legate, N., Christopherson, C., Wiggins, B., & Grahe, J. E. (2019). Publishing research with undergraduate students via replication work: The collaborative replications and education project. Frontiers in psychology, 10, 247. https://doi.org/10.3389/fpsyg.2019.00247

You may also be interested in

- ➔ Is science really facing a reproducibility crisis, and do we need it to? (Fanelli, 2018)
- ➔ Many hands make tight work(Silberzahn & Uhlmann, 2015)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)

Early co-authorship with top scientist predicts success in academic careers (Li et al., 2019)

Main Takeaways

- Academic impact is complex and linked to citation number. Junior researcher's output can be provided competitive advantage based on visibility.
- Present study asks whether a single event of interaction with 'top scientists' may alter junior researcher's future in academia.
- *Hypothesis*: the more co-authorship with 'top scientists', the more junior researchers have competitive advantage.
- *Method*: publication and citation data for four disciplines was indexed, since 1970 of selected journals for specific authors and institutions.
- The average prestige score of its authors' institution + Average prestige score of the researchers' papers = a paper's prestige score.
- *Results*: co-author with top scientists provide competitive advantage compared to peers of comparable early career profiles without top co-authors. Authors seem to suggest that students from less prestigious institutions would benefit junior researchers most.
- *Discussion*:
    - Authors seem to suggest 'top scientists' attract very best students-competitive advantage.
    - Authors seem to suggest a successful career is a result of interaction with a top scientist.
    - Authors seem to suggest that institutional prestige, productivity and a high citation count contributes to long-lasting academic impact for early-career researchers, especially those who are not among the best of the best.
    - Put simply, authors seem to suggest that being in the right place, at the right time, provides an early edge. To escape the prestige trap is to connect with top scientists in their field that benefits students from less prestigious institutions most.

Abstract

We examined the long-term impact of coauthorship with established, highly-cited scientists on the careers of junior researchers in four scientific disciplines. Here, using matched pair analysis, we find that junior researchers who coauthor work with top scientists enjoy a persistent competitive advantage throughout the rest of their careers, compared to peers with similar early career profiles but without top coauthors. Such early coauthorship predicts a higher probability of repeatedly coauthoring work with top-cited scientists, and, ultimately, a higher probability of becoming one. Junior researchers affiliated with less prestigious institutions show the most benefits from coauthorship with a top scientist. As a consequence, we argue that such institutions may hold vast amounts of untapped potential, which may be realised by improving access to top scientists.

APA Style Reference

Li, W., Aste, T., Caccioli, F., & Livan, G. (2019). Early coauthorship with top scientists predicts success in academic careers. *Nature communications, 10*(1), 1-9. https://doi.org/10.1038/s41467-019-13130-4 [ungated]

Is science really facing a reproducibility crisis, and do we need it to? (Fanelli, 2018)

Main Takeaways

- Science is in crisis due to unreliable findings, poor research quality and integrity, publication practices due to pressure to publish.
- Issues of flawed research and publication practice are higher than outright scientific misconduct.
- There are several differences between subfields: magnitude of true effect size, research bias, prior probability, true effects that are false positives and reproducibility of results and inferences.
- There is a strong decline in science such that strong initial findings were later contradicted by later studies.
- It is important to mention that published studies get longer, more complex, richer in data and null findings are placed in these long publications. This is done to remain accessible for researchers who are curious about these findings.
- We are going through a reproducibility crisis due to biased, fabricated, falsified, irreproducible, selective and underpowered findings.

Quotes

➢ *"Science always was and always will be a struggle to produce knowledge for the benefit of all of humanity against the cognitive and moral limitations of individual human beings, including the limitations of scientists themselves."* (p.2630)

➢ *"The second element of historical novelty is the rising power of information and communication technologies, which are transforming scientific practices in all fields… to make research more accurate, powerful, open, democratic, transparent, and self-correcting than ever before. At the same time, this technological revolution creates new expectations and new challenges that meta researchers are striving to address."* (p.2630)

Abstract

Efforts to improve the reproducibility and integrity of science are typically justified by a narrative of crisis, according to which most published results are unreliable due to growing problems with research and publication practices. This article provides an overview of recent evidence suggesting that this narrative is mistaken, and argues that a narrative of epochal changes and empowerment of scientists would be more accurate, inspiring, and compelling.

APA Style Reference

Fanelli, D. (2018). Opinion: Is science really facing a reproducibility crisis, and do we need it to?. *Proceedings of the National Academy of Sciences, 115*(11), 2628-2631. https://doi.org/10.1073/pnas.1708272114

You may also be interested in

➔ Publishing Research With Undergraduate Students via Replication Work: The Collaborative Replications and Education Project (Wagge et al., 2019)

➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)

➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)

➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)

➔ On the persistence of low power in psychological science (Vankov et al., 2014)

Six principles for assessing scientists for hiring, promotion, and tenure (Naudet et al, 2018)

Main Takeaways

- Academic work is usually quantified by the quantity of publications. However, this is not a reliable measure.
- An alternative measure is impact factor: the average number of citations to research articles over the preceding two years. This is an imperfect measure that does not capture the ethos of an academic institution.
- Impact factor provides information about citation influence for a few papers but is less informative about an individual publication and the authors involved in the publication (cf. Goodhart's Law - a valid measurement becomes useless when it becomes an optimisation target).
- Promotions are based on questionable research practices that promote the quantity of publications, but reproducible research does not receive such support.
- The incentive structure in academia is problematic, as the high impact factor is taken to be similar to high societal impact. This is not the case!
- High impact factor leads to more funding, more citations and further funding (cf. Matthew's Effect), whereas the opposite is observed for papers with low impact factor. Papers with high societal impact seem to fit the papers with low impact factor.
- We need to provide a more inclusive evaluation scheme that allows researchers and research to focus more on open science practices.
- We need to consider societal and broader impact for promotions.

Abstract

The negative consequences of relying too heavily on metrics to assess research quality are well known, potentially fostering practices harmful to scientific research such as p-hacking, salami science, or selective reporting. The "flourish or perish" culture defined by these metrics in turn drives the system of career advancement in academia, a system that empirical evidence has shown to be problematic and which fails to adequately take societal and broader impact into account. To address this systemic problem,

APA Style Reference

Naudet, F., Ioannidis, J., Miedema, F., Cristea, I. A., Goodman, S. N., & Moher, D. (2018). Six principles for assessing scientists for hiring, promotion, and tenure. Impact of Social Sciences Blog. http://eprints.lse.ac.uk/90753/

You may also be interested in

➔ The Nine Circles of Scientific Hell (Neuroskeptic, 2012)

- ➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)
- ➔ A user's guide to inflated and manipulated impact factor (Ioannidis & Thombs, 2019)

Psychologists Are Open to Change, yet Wary of Rules (Fuchs et al., 2012)

Main Takeaways

- How research is conducted and reported by psychologists must change?
- Present study investigated whether psychologists support concrete changes to data collection, reporting and publication processes? If not, what are their reasons?
- *Method*: 1292 psychologists from 42 countries were surveyed to assess whether each of Simmons et al.'s (2011) requirements and guidelines should be followed as a measure of good practice and whether they should be placed as mandatory conditions for publication in psychological journals.
- *Results:* 98% of psychologists are open to change and agree at least one requirement should be placed as a condition for publication, especially authors must report all conditions, including failed manipulations.
- *Results*: Reasons for not including a condition was too rigorous, do not agree with the argument or it was not appropriate for all studies.
- Psychologists are open to change for reporting and conducting research and agree with guidelines. However, some requirements are rigid and questionable.

Quotes

➢ "*Researchers and editorial staff alike must also ensure that standards are enforceable so as to avoid punishing honest researchers. The psychological community should capitalize on the current openness to change in order to develop and implement appropriate changes and thus improve the quality of published psychological research.*" (p. 641).

Abstract

Psychologists must change the way they conduct and report their research—this notion has been the topic of much debate in recent years. One article recently published in Psychological Science proposing six requirements for researchers concerning data collection and reporting practices as well as four guidelines for reviewers aimed at improving the publication process has recently received much attention (Simmons, Nelson, & Simonsohn, 2011). We surveyed 1,292 psychologists to address two questions: Do psychologists support these concrete changes to data collection, reporting, and publication practices, and if not, what are their reasons? Respondents also indicated the percentage of print and online journal space that should be dedicated to novel studies and direct replications as well as the percentage of published psychological research that they believed would be confirmed if direct replications were conducted. We found that psychologists are generally open to change. Five requirements for researchers and three guidelines for reviewers were supported as standards of good practice, whereas one requirement was even supported as a publication condition. Psychologists appear to be

less in favor of mandatory conditions of publication than standards of good practice. We conclude that the proposal made by Simmons, Nelson & Simonsohn (2011) is a starting point for such standards.

APA Style Reference

Fuchs, H. M., Jenny, M., & Fiedler, S. (2012). Psychologists are open to change, yet wary of rules. *Perspectives on Psychological Science, 7*(6), 639-642. https://doi.org/10.1177/1745691612459521

You may also be interested in

Six principles for assessing scientists for hiring, promotion, and tenure (Naudet et al, 2018)

- ➔ The Nine Circles of Scientific Hell (Neuroskeptic, 2012)
- ➔ CJEP Will Offer Open Science Badges (Pexman, 2017)
- ➔ Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency (Kidwell et al., 2016)
- ➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)
- ➔ Quality Uncertainty Erodes Trust in Science (Vazire, 2017)
- ➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)
- ➔ Seven Easy Steps to Open Science: An Annotated Reading List (Crüwell et al., 2019)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)

Registered Reports: Realigning incentives in scientific publishing (Chambers et al., 2015)

Main Takeaways

- Registered allows peer review to focus on the quality and rigour of the experimental design instead of ground-breaking results. This should reduce questionable research practices such as selective reporting, post-hoc hypothesising and low statistical power.
- Registered reports are reviewed and revised prior to data collection.
- Cortex editorial sub-team triages submissions with one week to reject manuscripts, invite revision to meet necessary standards or send out for Stage 1 in-depth review.
- Stage 1 has 8-10 weeks to move from initial review to in-principle acceptance.
- Stage 2 review has 4 weeks to final editorial decision.
- Registered report is not a one-shot cure for reproducibility problems in science and poses no threat to exploratory analyses.

Abstract

This is a view on registered reports in Cortex by Chris Chambers and colleagues. It contains information on Registered Reports and the length of duration for submission and review. They discuss the editorial process and that a registered report is not a threat to exploratory research and is not a panacea to cure reproducibility problems.

APA Style Reference

Chambers, C. D., Dienes, Z., McIntosh, R. D., Rotshtein, P., & Willmes, K. (2015). Registered reports: realigning incentives in scientific publishing. *Cortex, 66*, A1-A2. https://doi.org/10.1016/j.cortex.2015.03.022 [ungated]

You may also be interested in

- ➔ Registered Reports: A new publishing initiative at Cortex (Chambers, 2013)
- ➔ Registered Reports: A step change in scientific publishing (Chambers, 2014)
- ➔ Registered reports : a method to increase the credibility of published results (Nosek & Lakens, 2014)
- ➔ Registered reports (Jamieson et al., 2019)
- ➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)
- ➔ Seven Easy Steps to Open Science: An Annotated Reading List (Crüwell et al., 2019)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)
- ➔ On the persistence of low power in psychological science (Vankov et al., 2014)

Registered Reports: A new publishing initiative at Cortex (Chambers, 2013)

Main Takeaways

- We value novel and eye-catching findings over genuine findings, thus increasing questionable research practices.
- Editorial decisions are one cause of questionable research practices, as they make decisions based on results.
- Science undergraduates are taught about data analysis and hypothesis generation before the data is collected, ensuring the observer is independent of observation.
- Cortex provides registered reports to allow null results and replications be encouraged.
- Registered reports are manuscripts submitted before the experiment begins. This includes the introduction, hypotheses, procedures, analysis pipeline, power analysis and pilot data, if possible.
- Following peer review, the article is rejected or accepted in principle for publication, irrespective of the obtained results.
- Authors have to submit a finalised manuscript for re-review, share raw data and laboratory log.
- Pending quality checks and a sensible interpretation of findings, the manuscript is, in essence, accepted.
- Registered reports are immune to publication bias and need authors to adhere to pre-approved methodology and analysis pipeline to prevent questionable research practices from being used.
- A priori power analysis is required and the criteria for a registered report is seen as providing the highest truth value.
- Registered reports does not exclude exploratory analyses but must be distinguished from the analyses that were planned, also not all modes of scientific investigation fits registered reports.

Abstract

This is an editorial by Chris Chambers who encouraged Registered Reports in Cortex as a viable initiative to reduce questionable research practices, its benefits, limitations and what information to include in a registered report.

APA Style Reference

Chambers, C. D. (2013). Registered reports: a new publishing initiative at Cortex. *Cortex, 49*(3), 609-610. https://doi.org/10.1016/j.cortex.2012.12.016 [ungated]

You may also be interested in

➔ Registered Reports: A step change in scientific publishing (Chambers, 2014)

- ➔ Registered Reports: Realigning incentives in scientific publishing (Chambers et al., 2015)
- ➔ Registered reports : a method to increase the credibility of published results (Nosek & Lakens, 2014)
- ➔ Registered reports (Jamieson et al., 2019)
- ➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)
- ➔ Seven Easy Steps to Open Science: An Annotated Reading List (Crüwell et al., 2019)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)
- ➔ On the persistence of low power in psychological science (Vankov et al., 2014)

Registered Reports: A step change in scientific publishing (Chambers, 2014)

Main Takeaways

- Registered reports foster clarity and replication prior to experiments being conducted.
- Readers feel more confident that work is replicable with initial study predictions and analysis plans were independently reviewed.
- Registered reports are a departure from peer review.
- Low power, high rate of cherry picking, post-hoc hypothesising, lack of data sharing, journal culture marked by publication bias and few replication studies contributes to reproducibility crisis.
- Allows us to publish positive, negative or null findings, thus producing a true picture of the literature.
- We will not suffer from publication bias, as manuscript is worthy of publication, editors and reviewers are driven by quality of methods, as opposed to results.
- Registered reports are not an innovation but closer to restoration-reinvention of publication and peer review mechanisms.
- Registered reports allow creativity, flexibility and reporting of unexpected findings.

Quotes

➢ "*Ultimately, it is up to all of us to determine the future of any reform, and if the community continues to support Registered Reports then that future looks promising. Each field that adopts this initiative will be helping to create a scientific literature that is free from publication bias, that celebrates transparency, that welcomes replication as well as novelty, and in which the reported science will be more reproducible.*" (p. 3)

Abstract

Professor Chris Chambers, Registered Reports Editor of the Elsevier journal Cortex and one of the concept's founders, on how the initiative combats publication bias.

APA Style Reference

Chambers, C. (2014). Registered reports: A step change in scientific publishing. *Reviewers' Update. November, 13, 2014.* https://www.elsevier.com/reviewers-update/story/innovation-in-publishing/registered-reports-a-step-change-in-scientific-publishing

You may also be interested in

➔ Registered Reports: A new publishing initiative at Cortex (Chambers, 2013)

- ➔ Registered Reports: Realigning incentives in scientific publishing (Chambers et al., 2015)
- ➔ Registered reports : a method to increase the credibility of published results (Nosek & Lakens, 2014)
- ➔ Registered reports (Jamieson et al., 2019)
- ➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)
- ➔ Seven Easy Steps to Open Science: An Annotated Reading List (Crüwell et al., 2019)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)
- ➔ On the persistence of low power in psychological science (Vankov et al., 2014)

Registered reports: a method to increase the credibility of published results (Nosek & Lakens, 2014)

Main Takeaways

- This editorial discusses the value of pre-registration and replication, as not all articles are published.
- Direct replication adds data that increases the precision of effect size estimate for meta-analytic research. No direct replication, means it is difficult to identify false positives.
- Conceptual replications are more popular than direct replications, as the original operationalised is placed as a phenomenon.
- Direct replication encourages generalisability of effects, providing evidence that indicates the effect was not due to sampling, procedural or contextual error.
- Direct replication produces negative results negative results, thus improving the identification of boundary conditions for real effects.
- The benefit of a registered report is that the feedback provided from peer review on design improves the methodology and can be resubmitted for review and acceptance or rejection based on feedback.
- Successful proposals can be high-powered, high quality and faithful replication designs. This can be all done before the research is conducted.
- Conflict of interest is reduced in order to ensure a fair test, allowing reviewers to focus on methodological quality of research.
- Replication can provide additional questions than answers. Effect sizes can be more genuine, as opposed to being exaggerated as a result of larger sample size.
- Registered reports enable exploratory and confirmatory analyses, but a distinction is required. However, more trust can be placed in confirmatory analyses, as it follows a plan and ensures interpretability of reported p value.

Quotes

"*No single replication provides the definitive word for or against the reality of an effect, just as no original study provides definitive evidence for it. Original and replication research each provides a piece of accumulating evidence for understanding an effect and the conditions necessary to obtain it. Following this special issue, Social Psychology will publish some commentaries and responses by original and replication authors of their reflections on the inferences from the accumulated data, and questions that could be addressed in follow-up research.*" (p. 139)

Abstract

Professor Daniel Laken and Professor Brian Nosek provide an editorial on how pre-registration and registered reports can be used for the journal of Social Psychology to increase credibility of individual results and findings.

APA Style Reference

Nosek, B. A., & Lakens, D. (2014). Registered reports : a method to increase the credibility of published results. Social Psychology, 45(3), 137-141. https://doi.org/10.1027/1864-9335/a000192

You may also be interested in

- ➔ Registered Reports: A new publishing initiative at Cortex (Chambers, 2013)
- ➔ Registered Reports: A step change in scientific publishing (Chambers, 2014)
- ➔ Registered Reports: Realigning incentives in scientific publishing (Chambers et al., 2015)
- ➔ Registered reports (Jamieson et al., 2019)
- ➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)
- ➔ Seven Easy Steps to Open Science: An Annotated Reading List (Crüwell et al., 2019)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)
- ➔ On the persistence of low power in psychological science (Vankov et al., 2014)

Registered reports (Jamieson et al., 2019)

Main Takeaways.

- Stage I article is submitted with introduction, methods, analyses and conclusions of a study before carrying out research.
- Stage I includes the article and a cover letter, confirming all support and approval is in place, timeline for completing this study, statement confirming authors share raw data, digital materials, analysis and statements confirming authors register Stage I article.
- Stage I article includes title page, abstract, introduction, methods and analysis plan.
- Sent to peer review to judge if it is of sufficient quality. Peer reviewers assess importance of research question, introduction, plausibility, quality of hypotheses and methodological quality and appropriateness of data analysis plan, validity of inferential conclusions based on data.
- Method includes justification of sample sizes compared to question, description of participants, problems investigated ,a priori justification, procedures to deduce inclusion and exclusion criteria and clear protocol of experimental procedures.
- Data analysis: how data is treated and justified including all pre-processing steps.
- If approved, authors submit a Stage II registered report. Stage II is accepted for publication, if no problems arise if Stage II is consistent with the approved Stage I proposal.
- Stage II provides a complete and final report of the approved Stage I article, which also includes raw data, digital materials and analyses. Stage II focuses on quality of data reported, soundness of conclusions drawn from data and consistency with arguments and reasoning.
- Are data sufficiently resolved to support conclusions? Does the data answer authors' proposed hypotheses? Does the introduction, analyses match Stage I submission? Any unregistered and post-hoc analyses justified, methodologically sound and informative? Are conclusions consistent with collected data.
- Editor can ask for revisions or reject Stage II articles.

Abstract

Professor Randall K. Jamieson provides an editorial on registered reports for the journal Canadian Journal of Psychology and how it works in this specific journal.

APA Style Reference

Jamieson, R. K., Bodner, G. E., Saint-Aubin, J., & Titone, D. (2019) Editorial: Registered reports. *Canadian Journal of Experimental Psychology, 73*, 3-4. http://dx.doi.org/10.1037/cep0000169 [ungated]

You may also be interested in

- ➔ Registered Reports: A new publishing initiative at Cortex (Chambers, 2013)
- ➔ Registered Reports: A step change in scientific publishing (Chambers, 2014)
- ➔ Registered Reports: Realigning incentives in scientific publishing (Chambers et al., 2015)
- ➔ Registered reports : a method to increase the credibility of published results (Nosek & Lakens, 2014)
- ➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)
- ➔ Seven Easy Steps to Open Science: An Annotated Reading List (Crüwell et al., 2019)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)
- ➔ On the persistence of low power in psychological science (Vankov et al., 2014)

Don't let transparency damage science (Lewandowsky & Bishop, 2016)

Main Takeaways.

- Scientific communities have launched initiatives to increase transparency, open critique and data sharing.
- Good researchers include all perspectives but their openness can be abused by opponents who aim to stall inconvenient research.
- Science is prone to attacks and research requires rigour but also transparency to help responses of scientists and their institutions to correct criticisms.
- Open data and scientists should not regard all requests for data as harassment.
- Researchers should explain why they cannot share their research. Confidentiality issues need to be considered, also researchers need control over how data is used if the participant agrees to the sharing of this data.
- Social media can be used to remove biased, incorrect or misleading information.
- Engagement with critics is a fundamental part of scientific practice, researchers may feel obliged to respond even to trolls but can ignore abusive or illogical critics that make the same points.
- Minor corrections and clarifications after publications should not be seen as a stigma against fellow researchers.
- Publications are living documents with corrigenda are unwelcome but are accepted as part of scientific progress.
- Self-censorship affects academic freedom and discussion. Publication retractions should be reserved for fraud or grave errors.Call of retraction is coming from people who do not like a paper's conclusion.
- Complaints may undervalue researchers for legal but contentious science. Harassed scientists feel alone. They should not tolerate harassment dependent on race or gender nor if it is based on controversial science.
- Training and support should be used to aid researchers in the ability to cope with harassment.

Abstract

Stephan Lewandowsky and Dorothy Bishop explain how the research community should protect its members from harassment, while encouraging the openness that has become essential to science.

APA Style Reference

Lewandowsky, S., & Bishop, D. (2016). Research integrity: Don't let transparency damage science. *Nature, 529*(7587), 459-461.http://dx.doi.org/10.1038/529459a

You may also be interested in

- ➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)
- ➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)
- ➔ Fallibility in Science: Responding to Errors in the Work of Oneself and Others (Bishop, 2018)

The Nine Circles of Scientific Hell (Neuroskeptic, 2012)

Main Takeaways.

- There are nine circles of hell and Neuroskeptic explains which level of hell relates to questionable research practice.
- The first circle: Limbo- not a place of punishment but regret as scientists ignore or encourage scientists by awarding them grants and promotions.
- The second circle: overselling- the scientist exaggerates the importance of their work in order to attain grants or write better papers.
- The third circle: post-hoc storytelling- The scientist fires arrows at random, if a finding is noticed, a demon will explain at length or ramble that it aimed for this precise finding.
- The fourth circle: p-value fishing- obtain the result desired by ensuring that the *p* value is less than .05.
- The fifth circle: creative use of outliers- those who clean their results and exclude any data point without clear, explicit and a priori justification.
- The sixth circle: plagiarism- presenting another individual's work as their own work.
- The seventh circle: non-publication of data-each desk with file drawer stuffed with articles but drawers are locked.
- The eighth circle: partial publication of data - choose which group to chase at random-group is matched for age, gender, height and weight.
- The ninth circle: inventing data- data is made up.

Abstract

In the spirit of Dante Alighieri's Inferno, this paper takes a humorous look at the fate that awaits scientists who sin against best practice.

APA Style Reference

Neuroskeptic. (2012). The nine circles of scientific hell. *Perspectives on Psychological Science, 7*(6), 643-644. https://doi.org/10.1177/1745691612459519

You may also be interested in

- ➔ Psychologists Are Open to Change, yet Wary of Rules (Fuchs et al., 2012)
- ➔ Six principles for assessing scientists for hiring, promotion, and tenure (Naudet et al, 2018)
- ➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)
- ➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)

Attitudes Toward Open Science and Public Data Sharing: A Survey Among Members of the German Psychological Society (Abele-Brehm et al., 2019)

***Main Takeaways*.**

- Open science is the idea that scientific knowledge of all kinds should be openly shared: open access to published research, open methodology and open data. Public data sharing is the only topic not discussed.
- We should make data accessible for re-analyses in a secure, reliable and competently managed repository. A positive attitude towards open science but reservation whether data sharing will benefit young researchers' careers.
- The present study investigated the attitude towards open science in general and public data sharing, as attitudes not only contribute to an individual's research practice but also undergraduate, and postgraduate students, post-doctoral students, colleagues and the wider scientific community.
- *Method*: 337 people were given scales and open-ended questions with 14 items measured attitudes toward open science and toward public data sharing (e.g. what are the long-term consequences if a researcher shares raw data as part of a publication?).
- *Method*: Attitudes towards open science were separated into hopes and fears.
- *Results*: More hopes were related to open science and data sharing attitudes than fears. Both hopes and fears were highest among early-career researchers and lowest among professors. Positive attitudes toward data sharing is reduced by cost/benefit consideration.
- Attitudes towards open science and public data sharing was positive but fears that sharing data may have negative consequences for an individual's career, specifically if not all researchers participate or research parasites profit from data sharing if incentives remain unchanged.
- Professors exhibited least positive attitudes concerning consequences of open science and cost-benefit ratio of data sharing. In addition, they express less fear and hope linked to public data sharing than pre- and post-doctoral researchers.

Quotes

"This is, of course, true, but the idea of OS is transparency, and the question whether transparency and a higher commitment to data sharing and OS practices will eventually decrease QRPs and, thus, increase the robustness and replicability of psychological effects remains to be determined empirically." (p.259).

Abstract

Central values of science are, among others, transparency, verifiability, replicability, and openness. The currently very prominent Open Science (OS) movement supports these values. Among its most important principles are open methodology (comprehensive and useful documentation of methods and materials used), open access to published research

output, and open data (making collected data available for re-analyses). We here present a survey conducted among members of the German Psychological Society (N = 337), in which we applied a mixed-methods approach (quantitative and qualitative data) to assess attitudes toward OS in general and toward data sharing more specifically. Attitudes toward OS were distinguished into positive expectations ("hopes") and negative expectations ("fears"). These were uncorrelated. There were generally more hopes associated with OS and data sharing than fears. Both hopes and fears were highest among early career researchers and lowest among professors. The analysis of the open answers revealed that generally positive attitudes toward data sharing (especially sharing of data related to a published article) are somewhat diminished by cost/benefit considerations. The results are discussed with respect to individual researchers' behavior and with respect to structural changes in the research system.

APA Style Reference

Abele-Brehm, A. E., Gollwitzer, M., Steinberg, U., & Schönbrodt, F. D. (2019). Attitudes toward open science and public data sharing. *Social Psychology, 50*, 252-260. https://doi.org/10.1027/1864-9335/a000384

You may also be interested in

- ➔ Trust Your Science? Open Your Data and Code (Stodden, 2011)
- ➔ Open Data in Qualitative Research (Chauvette et al., 2019)
- ➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)
- ➔ Using OSF to Share Data: A Step-by-Step Guide (Soderberg, 2018)

Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results (Wicherts et al., 2011)

Main Takeaways.

- The American Psychological Association asks authors to sign a contract that data is available for individuals who wish to re-analyse the data to verify claims put forth in the paper.
- There has been no published research to assess this scenario in reality. The present study examined the willingness to share data for re-analysis linked to strength of evidence and quality of reporting of statistical results.
- *Method*: Wicherts et al. contacted corresponding authors of 141 papers published in the second half of 2004 in one of the four high-ranking journals published by the American Psychological Association and determined whether the effects of outliers contributed to statistical outliers.
- *Method*: They included authors from journal of personality and social psychology and journal of experimental psychology: learning, memory and cognition, as authors are more willing to share data than other journals.
- *Method*: They included tests results that were complete and reported as significant effects.
- *Results*: Reluctance to share was linked with weaker evidence and higher prevalence of apparent errors to report results. An unwillingness to share data was linked to reporting errors that affected statistical significance.
- The authors seem to suggest that a reluctance to share data was linked to more errors in reporting of results and with weaker evidence.The unwillingness to share data was more pronounced when errors concerned significance.
- Statistically rigorous researchers archive data better and are more attentive to statistical power than less statistically rigorous researchers.

Quotes

"Best practices in conducting analyses and reporting statistical results involve, for instance, that all co-authors hold copies of the data, and that at least two of the authors independently run all the analyses (as we did in this study). Such double-checks and the possibility for others to independently verify results later should go a long way in dealing with human factors in the conduct of statistical analyses and the reporting of results" (pp.6-7).

Abstract

The widespread reluctance to share published research data is often hypothesized to be due to the authors' fear that reanalysis may expose errors in their work or may produce conclusions that contradict their own. However, these hypotheses have not previously been studied systematically. We related the reluctance to share research data for reanalysis to 1148 statistically significant results reported in 49 papers published in two

major psychology journals. We found the reluctance to share data to be associated with weaker evidence (against the null hypothesis of no effect) and a higher prevalence of apparent errors in the reporting of statistical results. The unwillingness to share data was particularly clear when reporting errors had a bearing on statistical significance.Our findings on the basis of psychological papers suggest that statistical results are particularly hard to verify when reanalysis is more likely to lead to contrasting conclusions. This highlights the importance of establishing mandatory data archiving policies.

APA Style Reference

Wicherts, J. M., Bakker, M., & Molenaar, D. (2011). Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results. *PloS one, 6*(11), e26828. https://doi.org/10.1371/journal.pone.0026828

You may also be interested in

- ➔ Trust Your Science? Open Your Data and Code (Stodden, 2011)
- ➔ Attitudes Toward Open Science and Public Data Sharing: A Survey Among Members of the German Psychological Society (Abele-Brehm et al., 2019)
- ➔ Open Data in Qualitative Research (Chauvette et al., 2019)
- ➔ CJEP Will Offer Open Science Badges (Pexman, 2017)
- ➔ Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency (Kidwell et al., 2016)
- ➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)
- ➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)
- ➔ Seven Easy Steps to Open Science: An Annotated Reading List (Crüwell et al., 2019)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)
- ➔ Using OSF to Share Data: A Step-by-Step Guide (Soderberg, 2018)

Constraints on Generality (COG): A Proposed Addition to All Empirical Papers (Simons et al., 2017)

Main Takeaways.

- When a paper identifies a target population and specifies constraints on generality (COG) of findings, researchers conduct direct replications that sample from the target population, leading to more appropriate tests of reliability of the original claim.
- A COG statement indicates why the sample and target population is representative, justifying why subjects, materials and procedures are representative of broader populations.
- A COG statement does not limit the claim but leads the reader to correctly infer these findings limit to the groups of populations being tested such as undergraduate students.
- A COG statement inspires follow-up studies building on results by testing generality populations not originally tested.
- A COG statement encourages reviewers and editors more receptive to next-step studies to test constraints identified.
- A COG statement should be included in all papers, so editors support manuscripts with well-justified constraint on generality statements explicitly ground claims of generality.
- Editors can evaluate whether claims are sufficiently important to justify publication.
- A COG statement incentivises cumulative follow-up research, leading to greater reliability, influence and increased citations.
- This COG statement values rigor, honesty, accuracy and supports the conclusion justified by evidence and theory, allowing readers to understand the limits of generalisability.
- If science was more cumulative and self-correcting, broad generalisation might be justifiable.
- A COG statement describes known or anticipated limits on finding and not mediation by unknown factors. It asks how our sample is representative of a broader population.

Abstract

Psychological scientists draw inferences about populations based on samples—of people, situations, and stimuli—from those populations. Yet, few papers identify their target populations, and even fewer justify how or why the tested samples are representative of broader populations. A cumulative science depends on accurately characterizing the generality of findings, but current publishing standards do not require authors to constrain their inferences, leaving readers to assume the broadest possible generalizations. We propose that the discussion section of all primary research articles specify Constraints on Generality (i.e., a "COG" statement) that identify and justify target

populations for the reported findings. Explicitly defining the target populations will help other researchers to sample from the same populations when conducting a direct replication, and it could encourage follow-up studies that test the boundary conditions of the original finding. Universal adoption of COG statements would change publishing incentives to favor a more cumulative science.

APA Style Reference

Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science, 12*(6), 1123-1128. https://doi.org/10.1177/1745691617708630 [ungated]

You may also be interested in

➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)

Open Data in Qualitative Research (Chauvette et al., 2019)

Main Takeaways.

- This article argues that as a result of epistemological, methodological, legal and ethical issues, not all qualitative data is appropriate for open access.
- Open data allows researchers to test or refute new theories by validating research findings.
- Although data is becoming more available, we need to consider hotly debated issues concerning open data and that not all data is created equally, especially qualitative research.
- Qualitative research is not equally useful when decontextualized and requires contextualisation. Secondary analyses occur in teams or between collaborators when insider knowledge is shared.
- Qualitative research design is not beneficial to secondary analysis. Researchers become part of the research and may bias the data. Also, preconceptions should not be removed from the analyses.
- Personal knowledge is important for phenomenological research.
- Open data is not captured in transcripts and participants may conduct research to become active contributors to the research process. Field notes are written by researchers.
- Blanket consent form is used that discusses their data is kept indefinitely and reused by anyone. Confidentiality and anonymity becomes an issue for participants with open data.
- This becomes more problematic with small sample sizes, nature of questions, disclosure of information about sensitive issues that may be harmful to the individual and researcher.

Quotes

"Requirements for data access must consider the uniqueness and context of the data in each qualitative study. Consideration should be given to policies that grant the original research team adequate opportunities for involvement in publication of secondary analyses, perhaps with the rights to authorship to future publications if circumstances warrant. Alternatively, opportunities to comment on the new analysis and interpretation, considering the investigators' understanding of the unique context of the study, would provide some additional accountability" (p.4).

Abstract

There is a growing movement for research data to be accessed, used, and shared by multiple stakeholders for various purposes. The changing technological landscape makes it possible to digitally store data, creating opportunity to both share and reuse data anywhere in the world for later use. This movement is growing rapidly and becoming widely accepted as publicly funded agencies are mandating that researchers open their research data for sharing and reuse. While there are numerous advantages to use of open

data, such as facilitating accountability and transparency, not all data are created equally. Accordingly, reusing data in qualitative research present some epistemological, methodological, legal, and ethical issues that must be addressed in the movement toward open data. We examine some of these challenges and make a case that some qualitative research data should not be reused in secondary analysis.

APA Style Reference

Chauvette, A., Schick-Makaroff, K., & Molzahn, A. E. (2019). Open data in qualitative research. *International Journal of Qualitative Methods, 18,* 1609406918823863. https://doi.org/10.1177/1609406918823863

You may also be interested in

- ➔ Trust Your Science? Open Your Data and Code (Stodden, 2011)
- ➔ Attitudes Toward Open Science and Public Data Sharing: A Survey Among Members of the German Psychological Society (Abele-Brehm et al., 2019)
- ➔ Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results (Wicherts et al., 2011)
- ➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)

How scientists can stop fooling themselves (Bishop, 2020b)

Main Takeaways.

- Lab scientists should not be allowed to handle dangerous substances without safety training, researchers should not be allowed to be near a p value or similar measure of probability until we demonstrate they understand what it means.
- We ignore contradicting views when confronted with new data and preconceived notions may make us see a structure not there.
- People under-estimate how noisy small samples can be and conduct studies that lack the necessary power to detect an effect.
- More variables investigated, the more likely a significant value is significant.
- Basic statistical training is insufficient or counterproductive, providing misplaced confidence.
- Students discover how easy it is to find false results that are significant via simulation data. Students learn with simulation that small sample sizes are useless to show a moderate difference.
- Researchers need to build lifelong habits to avoid being led astray by specific confirmation bias.
- It is easy to forget papers that counter our own instacts, albeit the papers had no flaws.It enables us to understand the blind spots and how to avoid them.

Abstract

Sampling simulated data can reveal common ways in which our cognitive biases mislead us.

APA Style Reference

Bishop, D. (2020). How scientists can stop fooling themselves over statistics. *Nature, 584*(7819), 9. https://doi.org/10.1038/d41586-020-02275-8

You may also be interested in

➔ The Statistical Crisis in Science (Gelman & Loken, 2014)
➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)
➔ False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant (Simmons et al., 2011)
➔ Publication Prejudices: An Experimental Study of Confirmatory Bias in the Peer Review System (Mahoney, 1977)
➔ Quality Uncertainty Erodes Trust in Science (Vazire, 2017)
➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)
➔ Seven Easy Steps to Open Science: An Annotated Reading List (Crüwell et al., 2019)

- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)
- ➔ A consensus-based transparency checklist (Aczel et al., 2020)
- ➔ Tell it like it is (Anon, 2020)
- ➔ Is pre-registration worthwhile? (Szollosi et al., 2020)
- ➔ The life of p: "Just significant" results are on the rise (Leggett et al., 2013)

CJEP Will Offer Open Science Badges (Pexman, 2017)

Main Takeaways.

- Open data- the data is digitally shareable and made publicly available to reproduce results. Information necessary for replications must be included.
- Open materials: all materials necessary to reproduce reported results digitally shareable with descriptions of non-digital materials necessary for replication.
- Pre-registration: provide planned sample size, motivated research questions or hypotheses, outcome and predictor variables, including controls, co-variates and independent variables. This is provided prior to the data being collected.
- Pre-register + analysis: design a pre-register study with an analysis plan for research and results are recorded to plan.

Quote

"Indeed, in most cases, authors who wish to apply for badges will do so only after the editorial decision has been made. I understand that there are many reasons why it may not be possible to share data or materials, or to preregister a study, and so I certainly do not expect all authors to apply for badges. Nonetheless, I hope that many authors will devote the time required to make their data, materials, or research plans publicly available; these efforts are an important step toward improving our science." (p.1).

Abstract

This is a view on open science badges in Canadian Journal of Psychology by Professor Penny Pexman. It contains information about the badges, how to apply for them and that it is not a mandatory requirement.

APA Style Reference

Pexman, P. M. (2017). CJEP will offer open science badges. *Canadian Journal of Experimental Psychology= Revue Canadienne de Psychologie Experimentale, 71*(1), 1-1. https://doi.org/10.1037/cep0000128

You may also be interested in

- ➔ Trust Your Science? Open Your Data and Code (Stodden, 2011)
- ➔ Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results (Wicherts et al., 2011)
- ➔ Psychologists Are Open to Change, yet Wary of Rules (Fuchs et al., 2012)
- ➔ Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency (Kidwell et al., 2016)
- ➔ Quality Uncertainty Erodes Trust in Science (Vazire, 2017)

➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)

Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency (Kidwell et al., 2016)

Main Takeaways.

- Incentives in academia focus on publications and grants. More common are ineffective policies to encourage or require sharing on request.
- Researchers are not likely to share data and materials unless there are incentives. Badges is one means to signal and incentivise desirable behaviours.
- Badges acknowledge open practice signals-journal values transparency, authors met transparency standards for research and immediate signal of accessible data, materials, or pre-registration to readers.
- The present study investigated the influence of adopting badges by comparing data and material sharing rates before badges (i.e. 2012-2013) and after adoption (2014-May 2015) in Psychological Science and across the same time period in comparison journals from the same discipline.
- *Method*: 2478 articles based on experiment or observations measure data accessibility and research materials. Variables included were open data or open material badge, availability statement of data and material.
- *Method*: Whether data or materials are available at publicly accessible location. Correct data/material-if data or materials could be retrieved, whether it was linked to what was reported.
- *Results*: There was an increase of reporting open data after badges were introduced. However reporting openness does not guarantee openness. When badges are earned, available data is provided, correct, usable and complete than when it was not earned.
- *Results*: Open materials increased but not to the same extent.
- Psychological science adopts badges, report sharing rates increases 10-fold to 40%. Without badges- small percentage of reported sharing is a gross exaggeration of sharing.
- Sharing data was larger when a badge was earned than when it was not earned.
- Effects on sharing research materials were similar sharing data but weaker with badges producing only three times more sharing.

Quote

"However, actual evidence suggests that this very simple intervention is sufficient to overcome some barriers to sharing data and materials. Badges signal a valued behavior, and the specifications for earning the badges offer simple guides for enacting that behavior. Moreover, the mere fact that the journal engages authors with the possibility of promoting transparency by earning a badge may spur authors to act on their scientific values. Whatever the mechanism, the present results suggest that offering badges can increase sharing by up to an order of magnitude or more. With high return coupled with comparatively little cost, risk, or bureaucratic requirements, what's not to like?" (p.13).

Abstract

Beginning January 2014, Psychological Science gave authors the opportunity to signal open data and materials if they qualified for badges that accompanied published articles. Before badges, less than 3% of Psychological Science articles reported open data. After badges, 23% reported open data, with an accelerating trend; 39% reported open data in the first half of 2015, an increase of more than an order of magnitude from baseline. There was no change over time in the low rates of data sharing among comparison journals. Moreover, reporting openness does not guarantee openness. When badges were earned, reportedly available data were more likely to be actually available, correct, usable, and complete than when badges were not earned. Open materials also increased to a weaker degree, and there was more variability among comparison journals. Badges are simple, effective signals to promote open practices and improve preservation of data and materials by using independent repositories.

APA Style Reference

Kidwell, M. C., Lazarević, L. B., Baranski, E., Hardwicke, T. E., Piechowski, S., Falkenberg, L. S., ... & Errington, T. M. (2016). Badges to acknowledge open practices: A simple, low-cost, effective method for increasing transparency. *PLoS biology, 14*(5), e1002456. https://doi.org/10.1371/journal.pbio.1002456

You may also be interested in

- ➔ Trust Your Science? Open Your Data and Code (Stodden, 2011)
- ➔ Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results (Wicherts et al., 2011)
- ➔ Psychologists Are Open to Change, yet Wary of Rules (Fuchs et al., 2012)
- ➔ CJEP Will Offer Open Science Badges (Pexman, 2017)
- ➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)

Signalling the trustworthiness of science should not be a substitute for direct action against research misconduct (Kornfeld & Titus, 2020)

Main Takeaways.

- Truth is undermined by misconduct, fraud, failure to replicate and rise in the number of retractions and the public media.
- Fraudulent behaviour does not decrease trust in science.
- Fraudulent behaviour is a result of the fraudulent scientist, not untrustworthy science.
- Reports indicate failure to publish will prevent academic appointment, tenure and ensuring funding of laboratories as main concerns.
- Educating the public about high standards of science and scientists will not reduce outrage concerning fraudulent research.

Quote

"When then will these leaders of the scientific community finally direct their talents and energy to the culprit per se, research misconduct, and its perpetrators" (p.41).

Abstract

This is a response to the paper by Jamieson et al. (2019) on signalling trustworthiness in science. It contains information that the trust in science from the public and scientific community contributes to misconduct and fraudulent behaviour.

APA Style Reference

Kornfeld, D. S., & Titus, S. L. (2020). Signaling the trustworthiness of science should not be a substitute for direct action against research misconduct. *Proceedings of the National Academy of Sciences of the United States of America, 117*(1), 41. https://doi.org/10.1073/pnas.1917490116

You may also be interested in

- ➔ Reply to Kornfeld and Titus: No distraction from misconduct (Jamieson et al., 2020)
- ➔ Stop ignoring misconduct (Kornfeld & Titus, 2016)
- ➔ Fallibility in Science: Responding to Errors in the Work of Oneself and Others (Bishop, 2018)

Reply to Kornfeld and Titus: No distraction from misconduct (Jamieson et al., 2020)

***Main Takeaways*.**

- Funders should make research ethics a condition of support.
- Institutions should provide education to investigate misconduct fairly, rapidly and transparently, while protecting whistle-blowers.
- Scientists and outlets that publish their work need to provide methods (e.g. statistical checks, plagiarism checks, badges, checklists) used to honour science's integrity-protecting norms but when and how they have completed this task.
- These methods should uncover and increase awareness of biases that undermine the ability to fairly interpret their findings.
- These indicators of trustworthiness indicate that the honor of scientific integrity is protected and institutions can protect its integrity but signal how to protect itself.

Abstract

This is a response to the commentary by Kornfeld and Titus (2020). It contains information about the importance of research ethics for funders, how institutions should protect whistleblowers and provide education to prevent misconduct and how scientists and outlets can provide evidence they honour scientific integrity.

APA Style Reference

Jamieson, K. H., McNutt, M., Kiermer, V., & Sever, R. (2020). Reply to Kornfeld and Titus: No distraction from misconduct. *Proceedings of the National Academy of Sciences of the United States of America, 117*(1), 42. https://doi.org/10.1073/pnas.1918001116

You may also be interested in

➔ Signalling the trustworthiness of science should not be a substitute for direct action against research misconduct (Kornfeld & Titus, 2020)

➔ Stop ignoring misconduct (Kornfeld & Titus, 2016)

➔ Fallibility in Science: Responding to Errors in the Work of Oneself and Others (Bishop, 2018)

Stop ignoring misconduct (Kornfeld & Titus, 2016)

Main Takeaways.

- History of science shows irreproducibility is not a product of our times. These problems result from inadequate research practices and fraud. Current initiatives to improve science ignores fraudulent behaviour.
- Reducing irreproducibility is a wasted opportunity, if dishonesty is not being given much attention. Scientific leaders are trying to reduce questionable research practices but choose to ignore, not confront, the issue.
- These ethical issues and practices occurred long before people entered science.
- We need to consider reasons for misconduct: some are perfectionists and unable to cope with failure.
- Funders should craft policies to ensure mentors are advisers, teachers and role models, while limiting the number of trainees per mentor by discipline.
- Established scientists are less likely to commit misconduct if they were more concerned about being detected and punished.
- Whistle-blowers need to come forward and be protected. One method is to provide research integrity officers in the university who will protect them from retaliation.
- Research funds should be given only when current certification is provided by the institution. Those that fail to establish and execute these policies to ensure integrity will be made accountable when misconduct occurs.

Quote

“We believe that these system-wide interventions are essential to have an impact on the irreproducibility produced by research misconduct.” (p.30).

Abstract

This is an editorial by Kornfeld and Titus (2016) who discusses that misconduct needs to be taken seriously and discussed. It contains solutions to resolve matters concerning research integrity for both the scientist and research institute.

APA Style Reference

Kornfeld, D. S., & Titus, S. L. (2016). Stop ignoring misconduct. *Nature, 537*(7618), 29-30.https://doi.org/10.1038/537029a

You may also be interested in

➔ Signalling the trustworthiness of science should not be a substitute for direct action against research misconduct (Kornfeld & Titus, 2020)

➔ Reply to Kornfeld and Titus: No distraction from misconduct (Jamieson et al., 2020)

- ➔ Fallibility in Science: Responding to Errors in the Work of Oneself and Others (Bishop, 2018)

The Statistical Crisis in Science (Gelman & Loken, 2014)

Main Takeaways.

- Scientists argued that *p* values are seen as the perceived result of random variation. The value of *p* is a measure of the extent the dataset provides evidence against the null hypothesis.
- It is appropriate to look at data and create rules for data exclusion, coding and analysis to lead to statistical significance. This error is risky in small effect sizes, small sample sizes, large measurement errors and large variability.
- Statistically significant p values cannot be taken at face value even if linked to comparison consistent with existing theory.
- Paper is not published in a high-impact journal without a significant $p < .05$ result.
- There is a garden of forking paths. Put simply, you make multiple routes and determine this route leads to a significant result but the choices to reach this decision are done implicitly.
- It is good scientific practice to refine one's research hypotheses in light of the data. However, we need to be aware of data dredging, using both confidence intervals and p values to avoid getting fooled by noise.
- There is an issue of multiple comparisons that emerge as different choices about combining variables, inclusion and exclusion of cases, transformations of variables, tests for interactions in absence of main effects and other steps could occur with different data.
- Pre-registration is practical but cannot be a general solution. Researchers should be made aware of choices involved in data analysis.
- We can perform two experiments: exploratory and confirmatory with its own pre-registered protocol.
- We should move toward an analysis of all data instead of focusing on a single comparison or a small set of comparison.

Abstract

Data-dependent analysis— a "garden of forking paths" — explains why many statistically significant comparisons don't hold up.

APA Style Reference

Gelman, A., & Loken, E. (2014). The statistical crisis in science: data-dependent analysis--a" garden of forking paths"--explains why many statistically significant comparisons don't hold up. *American scientist, 102*(6), 460-466. [ungated]

You may also be interested in

➔ How scientists can stop fooling themselves (Bishop, 2020b)

- ➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)
- ➔ False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant (Simmons et al., 2011)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)
- ➔ The life of p: "Just significant" results are on the rise (Leggett et al., 2013)

Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)

Main Takeaways.

- To get an h index for promotion, hiring and funding, there is a publish or perish culture.
- The h index depends on the number of publications, citations and productivity. They are cited due to faults in methodology or lack of replicability? Does this mean citations are a good measure? No! (cf. Goodhart's Law).
- Academia provides short-term contracts to exploit without wasting resources.
- Publication aims for newsworthy results, leading to false positives and less integration with the literature, leading to only positive findings to be produced with unethical behaviours.
- Questionable research practices are seen as unethically as it distorts data to support the researchers' hypotheses and scientists are unwilling to self-correct.
- Scientists need to be open about their results. Many scientists subscribe to norm of communality.
- There is data sharing but not many people share their data. Scientists are assumed to self-regulate, but this assumption is erroneous.
- Incentives need to change and focus on quality, reproducibility, data sharing and impact on society.
- Pre-registration can help with publication biases and questionable research practice. Study should be published irrespective of findings.
- Fraud could occur but workload will increase, evaluation of methodology, data collection to evaluate adherence to pre-registration plan.
- Pre-registration could backfire, as editors may require revisions to protocols, study is complete and changes may be impossible.
- Scientists prioritise their own research over scientific inquiry or credibility.

Quotes

"*The success of science is often attributed to its objectivity: surely science is an impartial, transparent, and dispassionate method for obtaining the truth? In fact, there is growing concern that several aspects of typical scientific practice conflict with these principles and that the integrity of the scientific enterprise has been deeply compromised.*" (p.1)

"The first step is to recognise that science is fundamentally a human endeavour, and thus subject to the limitations and biases that underlie human behaviour. Can we design a scientific ecosystem that acknowledges scientists are only human?" (p.2)

"*Yet in a scientific ecosystem that rewards researchers for their productivity more than for their methodological rigor, a young investigator who is fully devoted to the truth* *cannot afford to be passionate about their reputation, and a young investigator passionate about their reputation cannot afford to be fully devoted to the truth. It is time to rehabilitate the scientific ecosystem, and the first step is to acknowledge that scientists are only human.*" (p.9).

Abstract

It is becoming increasingly clear that science has sailed into troubled waters. Recent revelations about cases of serious research fraud and widespread 'questionable research practices' have initiated a period of critical self-reflection in the scientific community and there is growing concern that several common research practices fall far short of the principles of robust scientific inquiry. At a recent symposium, 'Improving Scientific Practice: Dealing with the Human Factors' held at The University of Amsterdam, the notion of the objective, infallible, and dispassionate scientist was firmly challenged. The symposium was guided by the acknowledgement that scientists are only human, and thus subject to the desires, needs, biases, and limitations inherent to the human condition. In this article, five post-graduate students from University College London describe the issues addressed at the symposium and evaluate proposed solutions to the scientific integrity crisis.

APA Style Reference

Hardwicke, T E et al 2014 Only Human: Scientists, Systems, and Suspect Statistics. *Opticon1826, 16* (25), 1-12, http://dx.doi.org/10.5334/opt.ch

You may also be interested in

- ➔ Trust Your Science? Open Your Data and Code (Stodden, 2011)
- ➔ Is science really facing a reproducibility crisis, and do we need it to? (Fanelli, 2018)
- ➔ Six principles for assessing scientists for hiring, promotion, and tenure (Naudet et al, 2018)
- ➔ The Nine Circles of Scientific Hell (Neuroskeptic, 2012)
- ➔ Don't let transparency damage science (Lewandowsky & Bishop, 2016)
- ➔ Signalling the trustworthiness of science should not be a substitute for direct action against research misconduct (Kornfeld & Titus, 2020)
- ➔ Attitudes Toward Open Science and Public Data Sharing: A Survey Among Members of the German Psychological Society (Abele-Brehm et al., 2019)
- ➔ Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results (Wicherts et al., 2011)
- ➔ Open Data in Qualitative Research (Chauvette et al., 2019)
- ➔ How scientists can stop fooling themselves (Bishop, 2020b)

➔ Reply to Kornfeld and Titus: No distraction from misconduct (Jamieson et al., 2020)

➔ Stop ignoring misconduct (Kornfeld & Titus, 2016)

➔ The Statistical Crisis in Science (Gelman & Loken, 2014)

➔ False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant (Simmons et al., 2011)

➔ Publication Prejudices: An Experimental Study of Confirmatory Bias in the Peer Review System (Mahoney, 1977)

➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)

➔ Seven Easy Steps to Open Science: An Annotated Reading List (Crüwell et al., 2019)

➔ Fallibility in Science: Responding to Errors in the Work of Oneself and Others (Bishop, 2018)

➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)

➔ The life of p: “Just significant” results are on the rise (Leggett et al., 2013)

False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant (Simmons et al., 2011)

Main Takeaways.

- Hypotheses are made, data is collected and data is compared to the hypothesis. However, errors will frequently occur in the form of false positives. Null results result from several causes such as a failure to replicate previous research.
- High-impact journals do not discuss the lack of reproducibility and researchers have little incentive to replicate.
- False positives waste resources and inspire investment in research programs and ineffective policy changes.
- Data exclusion can contribute to false positive findings (e.g. why are some observations excluded, why are conditions combined and which condition is being compared, which is the control variable and should measures be combined, transformed or both?)
- We need to consider how outliers are treated. What is too fast and what is too slow (e.g. 1.5/2.5/3SD outlier removal or 1000-5000ms).
- Authors should decide a stopping rule before data collection begins. Power calculations should be reported or recruit as many students as possible prior to the end of the semester.
- Authors should provide cost-of-data-collection justification. Smaller samples reflect interim data and flexible termination rules.
- Authors should provide an exhaustive list on all variables collected in the study.
- Authors should report all experimental conditions, especially the failed manipulations.
- If observations are eliminated, statistical results need to be reported and any eliminations of the data need to be justified.
- If a covariate is included, authors should include analyses with and without the covariate- this is to ensure the results are due to the covariate instead of random assignment.
- Reviewers need to ensure authors follow the requirements and exclude alternative explanations to make sure the findings do not result from chance alone.
- Imperfections in the findings be tolerated by reviewers.
- Reviewers should require authors to show findings not due to an arbitrary analytical decision.
- If justification or data collection is not compelling, reviewers should conduct an exact replication.
- Authors should not be too selective.
- Bayesian approach increases researcher degrees of freedom and offers new analyses and flexibly try out on data.

Abstract

In this article, we accomplish two things. First, we show that despite empirical psychologists' nominal endorsement of a low rate of false-positive findings (≤ .05), flexibility in data collection, analysis, and reporting dramatically increases actual false-positive rates. In many cases, a researcher is more likely to falsely find evidence that an effect exists than to correctly find evidence that it does not. We present computer simulations and a pair of actual experiments that demonstrate how unacceptably easy it is to accumulate (and report) statistically significant evidence for a false hypothesis. Second, we suggest a simple, low-cost, and straightforwardly effective disclosure-based solution to this problem. The solution involves six concrete requirements for authors and four guidelines for reviewers, all of which impose a minimal burden on the publication process.

APA Style Reference

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological science, 22*(11), 1359-1366. https://doi.org/10.1177/0956797611417632 [ungated]

You may also be interested in

- ➔ How scientists can stop fooling themselves (Bishop, 2020b)
- ➔ The Statistical Crisis in Science (Gelman & Loken, 2014)
- ➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)
- ➔ A 21 Word Solution (Simmons et al., 2012)
- ➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)
- ➔ The life of p: "Just significant" results are on the rise (Leggett et al., 2013)

Publication Prejudices: An Experimental Study of Confirmatory Bias in the Peer Review System (Mahoney, 1977)

Main Takeaways.

- Confirmatory bias is humans seek out experiences to confirm beliefs. Cognitive bias is more prevalent in scientific publication. A piece of research is threatened by human decision making (i.e. the journal editor and reviewer).
- The present study investigated confirmation bias as problems of current review practices.
- To what extent do editors and referees weigh various components in evaluation?
- Although the ideal publication review system might focus on methodological quality and relevance over data outcome and interpretation, writing styles and conclusion also affect editorial decisions.
- What contributes to this review system and how can we reduce confirmatory bias?
- *Method*: five groups of referees read manuscripts that had data consistent or inconsistent with the reviewer's theoretical perspective.
- *Method:* Reviewers had to evaluate manuscript based on relevance and methodology.
- *Method*: two final groups of reviewers received mixed findings, supporting one perspective of the reviewer and the second was contradictory to the reviewer's perspective.
- *Results*: There was poor inter-rater reliability. Reviewers were more likely to show confirmation bias, thus were more supportive of manuscripts in favour of their theoretical perspective and strongly against manuscripts that contradict their perspective.
- Referees should be asked to evaluate relevance and methodology of an experiment without seeing its results or interpretations (cf. registered reports).
- Referees show little agreement on topics-train them to produce better and unprejudiced consensus. There will be perfect agreement if the same ideological or methodological biases are shared.
- Peer review is seen as an objective measure but ironically is very subjective in nature to biases. We need to investigation peer review and publication policies in detail to assess the transmission of scientific knowledge.

Abstract

Confirmatory bias is the tendency to emphasize and believe experiences which support one's views and to ignore or discredit those which do not. The effects of this tendency have been repeatedly documented in clinical research. However, its ramifications for the behavior of scientists have yet to be adequately explored. For example, although publication is a critical element in determining the contribution and impact of scientific findings~ little research attention has been devoted to the variables operative in journal review policies. In the present study, 75 journal reviewers were asked to referee

manuscripts which described identical experimental procedures but which reported positive, negative, mixed, or no results. In addition to showing poor interrater agreement, reviewers were strongly biased against manuscripts which reported results contrary to their theoretical perspective. The implications of these findings for epistemology and the peer review system are briefly addressed.

APA Style Reference

Mahoney, M. J. (1977). Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cognitive therapy and research, 1*(2), 161-175. https://doi.org/10.1007/BF01173636

You may also be interested in

- ➔ How scientists can stop fooling themselves (Bishop, 2020b)
- ➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)
- ➔ Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised trial (van Rooyen et al., 1999)
- ➔ The Peer Reviewers' Openness Initiative: incentivising open research practices through peer review (Morey et al., 2016)

Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised trial (van Rooyen et al., 1999)

Main Takeaways.

- The British Medical Journal wants to improve peer review. There was no evidence that investigated whether anonymous peer review is better than other forms of peer review.
- Open peer review (i.e. reviewer signing their review) was argued to produce more effort into their reviews, producing better reviews and increasing credibility and accountability.
- The present article aimed to confirm whether the reviews in open review was the same as traditional review.
- *Method*: when both reviews were received, manuscript was passed to a responsible editor and second editor randomly chosen from 12 other editors to measure review quality.
- *Method*: The corresponding author of each manuscript sent anonymous copies of two reviews and was told a decision on the manuscript and a review quality instrument was used to measure the quality of the review.
- *Method*: The quality of the review measure had 7 items and was based on the means of two editor scores and the corresponding author's score. The time taken to write the review and reviewer's recommendation concerning publication: minor or major revision or rejection.
- *Results*: Twelve percent of reviewers were more likely to decline to review if they were identified than if they were anonymous.
- *Results*: There was no difference in quality of reviews, difference in recommendation of reviewers or time taken to review the papers for anonymous and identified reviewers.
- *Results*: The editors' quality scores for reviews was higher than that of the authors. Most authors support open peer review.
- There was no difference in quality of and time taken to produce the review.
- Authors rate reviews that recommend publications higher than those who recommend rejection.
- Editors were not affected by the reviewer's opinion of the merit of a paper when assessing the quality of the review.

Abstract

To examine the effect on peer review of asking reviewers to have their identity revealed to the authors of the paper. Randomised trial. Consecutive eligible papers were sent to two reviewers who were randomised to have their identity revealed to the authors or to remain anonymous. Editors and authors were blind to the intervention. The quality of the reviews was independently rated by two editors and the corresponding author using a validated instrument. Additional outcomes were the time taken to complete the review and the recommendation regarding publication. A questionnaire survey was undertaken of the authors of a cohort of manuscripts submitted for publication to find out their views on open peer review. Two editors' assessments were obtained for 113 out of 125 manuscripts, and the corresponding author's assessment was obtained for 105. Reviewers randomised to be asked to be identified were 12% (95% confidence interval 0.2% to 24%) more likely to decline to review than reviewers randomised to remain anonymous (35% v 23%). There was no significant difference in quality (scored on a scale of 1 to 5) between anonymous reviewers (3.06 (SD 0.72)) and identified reviewers (3.09 (0.68)) ($P = 0.68$, 95% confidence interval for difference - 0.19 to 0.12), and no significant difference in the recommendation regarding publication or time taken to review the paper. The editors' quality score for reviews (3.05 (SD 0.70)) was significantly higher than that of authors (2.90 (0.87)) ($P < 0.005$, 95%confidence interval for difference - 0.26 to - 0.03). Most authors were in favour of open peer review. Asking reviewers to consent to being identified to the author had no important effect on the quality of the review, the recommendation regarding publication, or the time taken to review, but it significantly increased the likelihood of reviewers declining to review.

APA Style Reference

Van Rooyen, S., Godlee, F., Evans, S., Black, N., & Smith, R. (1999). Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised trial. *Bmj, 318*(7175), 23-27. https://doi.org/10.1136/bmj.318.7175.23

You may also be interested in

- ➔ Publication Prejudices: An Experimental Study of Confirmatory Bias in the Peer Review System (Mahoney, 1977)
- ➔ The Peer Reviewers' Openness Initiative: incentivising open research practices through peer review (Morey et al., 2016)

The Peer Reviewers' Openness Initiative: incentivising open research practices through peer review (Morey et al., 2016)

Main Takeaways.

- Openness and transparency is crucial to science. However, technology limits open science.
- Open data, open materials, open code and better replication studies have accelerated scientific progress.
- Scientific articles allow collaboration to learn what is true instead of the findings based on the analysis.
- Openness is an ethical obligation that provides further advantages, is being seen as a policy change and granting agencies.
- Learning openness is not difficult and implementing them will delay publication even if only for a few days.
- Open practices should be considered by reviewers, as it increases scientific quality.
- The relationship between reviewers and authors are important for the process. A missing figure or statistical results requires the author to be contacted for clarification.
- Authors provide a link of the data and materials and if they do not provide, they must justify this reason. If there is no real reason (e.g. legal, ethical or impracticality), reviewers should provide a short review for a lack of openness and failure to justify.
- Documents with details on how to interpret any files or code and how to run software should be made available.
- Joining this initiative will ease review load, if these papers do not meet this requirement, they are rejected.
- Open research is not a matter of policy, but a matter of scientific value and quality of product.
- Open practices are not standardised and are driven by practice. Authors that lack training in open practices and scientists need to learn new skills and knowledge.
- Senior researchers can help students curate data and research materials. Once the student leaves, they can allow people to use the materials.
- Open data allows the reviewer the option to check the analysis.
- Initiative is targeted at reviewers, not action editors. All science needs to be open and researchers who value open research practices should join the Initiative to help promote open research.

Abstract

Openness is one of the central values of science. Open scientific practices such as sharing data, materials and analysis scripts alongside published articles have many benefits, including easier replication and extension studies, increased availability of data for

theory-building and metaanalysis, and increased possibility of review and collaboration even after a paper has been published. Although modern information technology makes sharing easier than ever before, uptake of open practices had been slow. We suggest this might be in part due to a social dilemma arising from misaligned incentives and propose a specific, concrete mechanism—reviewers withholding comprehensive review—to achieve the goal of creating the expectation of open practices as a matter of scientific principle.

APA Style Reference

Morey, R. D., Chambers, C. D., Etchells, P. J., Harris, C. R., Hoekstra, R., Lakens, D., ... & Vanpaemel, W. (2016). The Peer Reviewers' Openness Initiative: incentivizing open research practices through peer review. *Royal Society Open Science, 3*(1), 150547. https://doi.org/10.1098/rsos.150547

You may also be interested in

- ➔ Publication Prejudices: An Experimental Study of Confirmatory Bias in the Peer Review System (Mahoney, 1977)
- ➔ Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised trial (van Rooyen et al., 1999)

A 21 Word Solution (Simmons et al., 2012)◈

Main Takeaways.

- Scientific journals should require authors to disclose data collection and data analysis.
- False positives need to be scrutinised.
- Say what your sample size was in advance, be transparent and disclose information that you did not drop any variables or conditions.
- We cannot trust our colleagues to run and report studies properly, if some people believe it is okay to drop conditions and variables and others do not believe this is good scientific practice.
- Many forms of p-hacking is also encouraged. We should ask if this is a 1 or 2 dependent variable study.
- Disclosure does not reduce p-hacking and does not reduce probability of false positives. We can maintain red tape by substituting some less vital aspects of style requirement from the American Psychological Association guide with those who propose it.
- Papers should include this proposed 21 words to improve its credibility.

Quote

"*We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study*." (p.1)

"*We hope that editors will emulate the pragmatic politicians of the 1900s and implement disclosure requirements in our journals before a perfect solution with no detractors is arrived at. In the meantime, those of us who realize transparency is a necessary condition for evidence to be scientific can start adding 21 words to our papers.*" (p.4)

Abstract

This is an editorial by Simmons et al. (2012) who discusses how to improve the credibility of psychological research by using 21 words.

APA Style Reference

Simmons, Joseph P. and Nelson, Leif D. and Simonsohn, Uri, A 21 Word Solution (October 14, 2012). Available at SSRN: https://ssrn.com/abstract=2160588 or http://dx.doi.org/10.2139/ssrn.2160588

You may also be interested in

➔ False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant (Simmons et al., 2011)

- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)
- ➔ The life of p: “Just significant” results are on the rise (Leggett et al., 2013)

Quality Uncertainty Erodes Trust in Science (Vazire, 2017)

Main Takeaways.

- Quality of products and willingness to pay is determined by the evaluation of the consumer.
- Consumers of new scientific findings evaluate the strength of findings and place stock in them.
- Inability to produce an informed evaluation of the quality leads to lower willingness to place stock in any product, leading to a lack of trust in the market.
- This uncertainty in quality affects our confidence in the findings and research based on these results.
- Greater transparency increases quality certainty and restores trust in science.
- Researchers cannot discriminate between lemons and high-quality findings. Quality certainty is determined by the amount of valuable information provided to the consumer of the manuscript (e.g. Raw data, original design and analysis plan). The more information is hidden, the larger quality uncertainty.
- Shoddy findings do not stand up and it is too late-high-quality researchers have been removed.
- We cannot depend on a few experts to evaluate the claims and ask people to naturally trust these experts. We need to make a judgment of the quality of the manuscript.
- Transparency provides buyers the information required to detect errors in these articles.
- Increased transparency should allow us to increase the chance of catching misconduct and fraudulent behaviour, unless they are willing to manipulate background information.
- Journals focus on maximising citation impacts instead of reliable, reproducible and robust science. They want to give us shiny and low-quality items. The journal reputation is related to the impact factor.

Quote

"*Some scientists find this revolution in the name of increased transparency and openness distasteful – they do not see a problem with the current system, and fear that this movement will undermine the public's trust in science. I would argue that these scientists have lost touch with what the public expects of science. For many non-scientists, learning that transparency is not the norm in science comes as a surprise. To anyone outside of the power hubs of science, it must seem obvious that scientists should be held to a higher standard than used car salespeople.*" (p.3)

Abstract

When consumers of science (readers and reviewers) lack relevant details about the study design, data, and analyses, they cannot adequately evaluate the strength of a scientific study. Lack of transparency is common in science, and is encouraged by journals that

place more emphasis on the aesthetic appeal of a manuscript than the robustness of its scientific claims. In doing this, journals are implicitly encouraging authors to do whatever it takes to obtain eye-catching results. To achieve this, researchers can use common research practices that beautify results at the expense of the robustness of those results (e.g., p-hacking). The problem is not engaging in these practices, but failing to disclose them. A car whose carburetor is duct-taped to the rest of the car might work perfectly fine, but the buyer has a right to know about the duct-taping. Without high levels of transparency in scientific publications, consumers of scientific manuscripts are in a similar position as buyers of used cars – they cannot reliably tell the difference between lemons and high quality findings. This phenomenon – quality uncertainty – has been shown to erode trust in economic markets, such as the used car market. The same problem threatens to erode trust in science. The solution is to increase transparency and give consumers of scientific research the information they need to accurately evaluate research. Transparency would also encourage researchers to be more careful in how they conduct their studies and write up their results. To make this happen, we must tie journals' reputations to their practices regarding transparency. Reviewers hold a great deal of power to make this happen, by demanding the transparency needed to rigorously evaluate scientific manuscripts. The public expects transparency from science, and appropriately so – we should be held to a higher standard than used car salespeople.

APA Style Reference

Vazire, S. (2017). Quality Uncertainty Erodes Trust in Science. *Collabra: Psychology, 3*(1), 1. DOI: http://doi.org/10.1525/collabra.74

You may also be interested in

- ➔ Trust Your Science? Open Your Data and Code (Stodden, 2011)
- ➔ Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results (Wicherts et al., 2011)
- ➔ Is science really facing a reproducibility crisis, and do we need it to? (Fanelli, 2018)
- ➔ Psychologists Are Open to Change, yet Wary of Rules (Fuchs et al., 2012)
- ➔ The Nine Circles of Scientific Hell (Neuroskeptic, 2012)
- ➔ Don't let transparency damage science (Lewandowsky & Bishop, 2016)
- ➔ How scientists can stop fooling themselves (Bishop, 2020b)
- ➔ CJEP Will Offer Open Science Badges (Pexman, 2017)
- ➔ Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency (Kidwell et al., 2016)

- ➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)
- ➔ False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant (Simmons et al., 2011)
- ➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)
- ➔ Seven Easy Steps to Open Science: An Annotated Reading List (Crüwell et al., 2019)
- ➔ Promoting an open research culture (Nosek et al., 2015)
- ➔ Promoting Transparency in Social Science Research (Miguel et al., 2014)
- ➔ Fallibility in Science: Responding to Errors in the Work of Oneself and Others (Bishop, 2018)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)

Rein in the four horsemen of irreproducibility (Bishop, 2019)

***Main Takeaways*.**

- Publication bias harms patients. The tendency not to publish negative results misleads readers.
- Low statistical power misleads readers as well, as a small sample size and manipulation is small, there is low probability one will detect an effect even if it exists.
- Resources and time is wasted for these underpowered studies.
- P-hacking distorts findings-choose a finding that looks exciting and write a paper about it. P values when removed are pointless. These problems are exacerbated in older than junior staff.
- Social media is allowing us to criticise the papers. Most journals are adopting registered reports and funders are encouraging strict guidelines, with data and scripts being made open and methods being fully described.

Abstract

Dorothy Bishop describes how threats to reproducibility, recognized but unaddressed for decades, might finally be brought under control.

APA Style Reference

Bishop, D. (2019). Rein in the four horsemen of irreproducibility. *Nature, 568*(7753), 435-436. http://doi.org/10.1038/d41586-019-01307-2

You may also be interested in

➔ Trust Your Science? Open Your Data and Code (Stodden, 2011)

➔ Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results (Wicherts et al., 2011)

➔ Is science really facing a reproducibility crisis, and do we need it to? (Fanelli, 2018)

➔ Psychologists Are Open to Change, yet Wary of Rules (Fuchs et al., 2012)

➔ The Nine Circles of Scientific Hell (Neuroskeptic, 2012)

➔ Don't let transparency damage science (Lewandowsky & Bishop, 2016)

➔ How scientists can stop fooling themselves (Bishop, 2020b)

➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)

- ➔ False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant (Simmons et al., 2011)
- ➔ Registered Reports: Realigning incentives in scientific publishing (Chambers et al., 2015)
- ➔ Registered Reports: A new publishing initiative at Cortex (Chambers, 2013)
- ➔ Registered Reports: A step change in scientific publishing (Chambers, 2014)
- ➔ Registered reports: a method to increase the credibility of published results (Nosek & Lakens, 2014)
- ➔ Registered reports (Jamieson et al., 2019)
- ➔ Quality Uncertainty Erodes Trust in Science (Vazire, 2017)
- ➔ Seven Easy Steps to Open Science: An Annotated Reading List (Crüwell et al., 2019)
- ➔ Promoting an open research culture (Nosek et al., 2015)
- ➔ Promoting Transparency in Social Science Research (Miguel et al., 2014)
- ➔ Fallibility in Science: Responding to Errors in the Work of Oneself and Others (Bishop, 2018)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)
- ➔ On the persistence of low power in psychological science (Vankov et al., 2014)
- ➔ The life of p: "Just significant" results are on the rise (Leggett et al., 2013)

Seven Easy Steps to Open Science: An Annotated Reading List (Crüwell et al., 2019)

Main Takeaways.

- Students and academics with little knowledge of open science may not easily find and make use of resources.
- Transparency and robustness may not guarantee increased rigour.
- Readers should plan data collection and analysis, be aware of assumptions of statistical models and understanding of statistical tools.
- Credibility of scientific claims depend on replicability.
- Open access removes barriers to access and distributes research.
- The gold route refers to publicly available articles, while green route relates to self-archiving or the works are made publicly available by people who created the (e.g. preprints).
- Open access articles are cited between 36%-600% more than non-open access work. It is given more coverage and discussed more in non-scientific settings.
- Researchers need to consider how they share their data. Is it findable, accessible, inter-operable and reusable (FAIR)?
- All steps of data analysis should be recorded in open source programs (e.g. R or Python) or placed in a reproducible syntax file.
- Pre-registration is an open science practice protecting people from biases, encourages transparency about analytic decision-making, supporting rigorous scientific research, enabling more replicable and reproducible work.
- Open science increases confidence and replicability of scientific results. Direct replication duplicates necessary elements to produce original findings, whereas conceptual replication changes one component of the original procedure- sample or measure.

Quote

"We hope that this paper will provide researchers interested in open science an accessible entry point to the practices most applicable to their needs. For all of the steps presented in this annotated reading list, any time taken by researchers to understand the issues and develop better practices will be rewarded in orders of magnitude. On an individual level, time and effort are ultimately saved, errors are reduced, and one's own research is improved through a greater adherence to openness and transparency. On a field-wide level, the more researchers invest in adopting these practices, the closer the field will come toward adhering to scientific norms and the values it claims to espouse." (p.245)

Abstract

The open science movement is rapidly changing the scientific landscape. Because exact definitions are often lacking and reforms are constantly evolving, accessible guides to open science are needed. This paper provides an introduction to open science and related reforms in the form of an annotated reading list of seven peer-reviewed articles, following

the format of Etz, Gronau, Dablander, Edelsbrunner, and Baribault (2018). Written for researchers and students – particularly in psychological science – it highlights and introduces seven topics: understanding open science; open access; open data, materials, and code; reproducible analyses; preregistration and registered reports; replication research; and teaching open science. For each topic, we provide a detailed summary of one particularly informative and actionable article and suggest several further resources. Supporting a broader understanding of open science issues, this overview should enable researchers to engage with, improve, and implement current open, transparent, reproducible, replicable, and cumulative scientific practices

APA Style Reference

Crüwell, S., van Doorn, J., Etz, A., Makel, M. C., Moshontz, H., Niebaum, J. C., ... & Schulte-Mecklenbeck, M. (2019). Seven Easy Steps to Open Science. *Zeitschrift für Psychologie*. https://doi.org/10.1027/2151-2604/a000387

You may also be interested in

- ➔ Trust Your Science? Open Your Data and Code (Stodden, 2011)
- ➔ Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results (Wicherts et al., 2011)
- ➔ Psychologists Are Open to Change, yet Wary of Rules (Fuchs et al., 2012)
- ➔ How scientists can stop fooling themselves (Bishop, 2020b)
- ➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)
- ➔ Registered Reports: Realigning incentives in scientific publishing (Chambers et al., 2015)
- ➔ Registered Reports: A new publishing initiative at Cortex (Chambers, 2013)
- ➔ Registered Reports: A step change in scientific publishing (Chambers, 2014)
- ➔ Registered reports: a method to increase the credibility of published results (Nosek & Lakens, 2014)
- ➔ Registered reports (Jamieson et al., 2019)
- ➔ Quality Uncertainty Erodes Trust in Science (Vazire, 2017)
- ➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)
- ➔ Promoting an open research culture (Nosek et al., 2015)
- ➔ Promoting Transparency in Social Science Research (Miguel et al., 2014)

- ➔ Fallibility in Science: Responding to Errors in the Work of Oneself and Others (Bishop, 2018)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)

Many hands make tight work (Silberzahn & Uhlmann, 2015)

***Main Takeaways*.**

- It is argued that re-running the analysis produces the same outcome. An analysis run by a single team-researchers takes on several roles: inventor creates ideas and hypotheses; analysts scrutinise data to support hypotheses and devil's advocate use different approaches to show weaknesses in the findings.
- A team of skilled researchers validate findings, inform policymakers and balance discussions.
- Several teams work with the same dataset with the hypotheses and results are held close.
- All researchers discuss results via email exchanges and researchers add notes to individual reports in others' work. A broad range of effect size is disturbing but any single analysis is too seriously mistaken.
- Crowdsourcing is not optimal, demands huge resources for one question. Decisions have to be made concerning hypotheses, data collection and which variables can or cannot be collected.
- Strong storylines are favoured over a messy reality. Scientists are hungry for reliable methods of discovery and to improve their networking.

Abstract

Crowdsourcing research can balance discussions, validate findings and better inform policy, say Raphael Silberzahn and Eric L. Uhlmann.

APA Style Reference

Silberzahn, R., & Uhlmann, E. L. (2015). Crowdsourced research: Many hands make tight work. *Nature News, 526*(7572), 189. https://doi.org/10.1038/526189a

You may also be interested in

➔ Publishing Research With Undergraduate Students via Replication Work: The Collaborative Replications and Education Project (CREP; Wagge et al., 2019)

➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)

A user's guide to inflated and manipulated impact factor (Ioannidis & Thombs, 2019)

Main Takeaways.

- A widely misused metric is the impact factor, reflecting the importance of the publications in a specific journal.
- Promotion and funding depends on impact factor (cf. Goodhart's Law). Over 14000 say let us remove it as an individual article quality based on the Declaration on Research Assessment.
- Tricks are used to inflate the impact factor of journals. Higher impact factors lead to more articles submitted and published.
- Inappropriate use of impact factor is unlikely to stop, especially with a large number of papers being cited without being counted.
- Self citations or requesting submitting articles to include citations to other recent articles without justification will lead to this inappropriate use of the impact factor.
- Review articles get more citations than research articles. Papers with questionable scientific value will get cited as standard reference.
- Authors should submit papers to target journals based on the journal's relevance, scientific rigour and quality, not impact factors.
- Authors who submit to journals with high impact inflation become members of bubbles and likely to publish in journals that may be discredited.
- We need to replace impact factor with median citations per item indicators separately for articles, reviews and other article types.
- We need metrics that can exclude self-citation, that are difficult to game and are appropriate.

Abstract

This is a view on impact factor by Professor John P.A. Ioannidis and Dr Brett D. Thombs. It contains a discussion of the impact factor being misused, how it is misused by journals and reviewers but provides solutions to overcome the use of this metric. In addition, we should base journals not on the impact factor but the relevance, scientific rigour and quality of the journal.

APA Style Reference

Ioannidis, J. P., & Thombs, B. D. (2019). A user's guide to inflated and manipulated impact factors. *European journal of clinical investigation, 49*(9), e13151. https://doi.org/10.1111/eci.13151

You may also be interested in

➔ Six principles for assessing scientists for hiring, promotion, and tenure (Naudet et al, 2018)

Promoting an open research culture (Nosek et al., 2015)

Main Takeaways.

- The reward system focuses on innovation, as opposed to replication and being open, transparent and reproducible.
- There is no means to align individual and communal incentives via universal scientific policies and procedures.
- We should reward researchers for time and effort spent in open practices.
- Citation stands should extend to data, code and research materials. Regular and rigorous citation of these materials and credit them as original intellectual contribution.
- Reproducibility increases confidence in results and allows scholars to learn more about data interpretation.
- The transparency guidelines are used to improve transparency about the research process, while reducing vague or incomplete reporting of methodology.
- Pre-registration of studies facilitates discovery of research, allowing the study to be recorded in a public registry.
- Several levels are used to encourage open science policy: Level 1-no barrier or incentive to open science. This reduces the effort on journal efficiency and workflow.
- Level 2 has stronger expectations but avoids cost to editors and open data is placed in a trusted repository.
- Level 3 is the strongest standard but provides some barriers.
- Quality of publication increases by reducing time spent on communication with authors and reviewers, improving standard of reporting.

Quote

"The journal article is central to the research communication process. Guidelines for authors define what aspects of the research process should be made available to the community to evaluate, critique, reuse, and extend. Scientists recognize the value of transparency, openness, and reproducibility. Improvement of journal policies can help those values become more evident in daily practice and ultimately improve the public trust in science, and science itself." (p.1425).

Abstract

Author guidelines for journals could help to promote transparency, openness, and reproducibility.

APA Style Reference

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., ... & Contestabile, M. (2015). Promoting an open research culture. *Science, 348*(6242), 1422-1425. http://doi.org/10.1126/science.aab2374

You may also be interested in

- ➔ Quality Uncertainty Erodes Trust in Science (Vazire, 2017)
- ➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)
- ➔ Seven Easy Steps to Open Science: An Annotated Reading List (Crüwell et al., 2019)
- ➔ Promoting Transparency in Social Science Research (Miguel et al., 2014)
- ➔ Fallibility in Science: Responding to Errors in the Work of Oneself and Others (Bishop, 2018)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)

Promoting Transparency in Social Science Research (Miguel et al., 2014)

Main Takeaways.

- The incentives, institutions and norms that social science makes it difficult to improve research design.
- Social science journals do not instruct adherence to reporting standards, data sharing or study registrations.
- Researchers have incentives to analyse and present data to make them more publishable at the expense of being accurate.
- This article surveys recent progress towards research transparency in the social sciences and provides standards and rules to realign scientific and scholarly incentives with scientific and scholarly values.
- Social scientists should use pre-registration to provide detailed documents, specify statistical models, dependent variables, covariates, interaction terms and multiple testing corrections to reduce biases.
- Open data and materials allows researchers to test alternative approaches on the data, reproduce results, identify misreported or fraudulent results; reuse or adapt materials for replication to improve interventions measures and so on.
- We need to move towards greater research transparency to pre-registration.

Quote

"Scientific inquiry requires imaginative exploration. Many important findings originated as unexpected discoveries. But findings from such inductive analysis are necessarily more tentative because of the greater flexibility of methods and tests and, hence, the greater opportunity for the outcome to obtain by chance. The purpose of prespecification is not to disparage exploratory analysis but to free it from the tradition of being portrayed as formal hypothesis testing. New practices need to be implemented in a way that does not stifle creativity or create excess burden. Yet we believe that such concerns are outweighed by the benefits that a shift in transparency norms will have for overall scientific progress, the credibility of the social science research enterprise, and the quality of evidence that we as a community provide to policy-makers" (p.31).

Abstract

Social scientists should adopt higher transparency standards to improve the quality and credibility of research.

APA Style Reference

Miguel, E., Camerer, C., Casey, K., Cohen, J., Esterling, K. M., Gerber, A., ... & Laitin, D. (2014). Promoting transparency in social science research. *Science, 343*(6166), 30-31. http://doi.org/10.1126/science.1245317

You may also be interested in

- ➔ Quality Uncertainty Erodes Trust in Science (Vazire, 2017)
- ➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)
- ➔ Seven Easy Steps to Open Science: An Annotated Reading List (Crüwell et al., 2019)
- ➔ Promoting an open research culture (Nosek et al., 2015)
- ➔ Promoting Transparency in Social Science Research (Miguel et al., 2014)
- ➔ Fallibility in Science: Responding to Errors in the Work of Oneself and Others (Bishop, 2018)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)

Rebuilding Ivory Tower: A bottom-up experiment in aligning research with societal needs (Hart & Silka, 2020) ◈

Main Takeaways.

- Scientists are trained to conduct good science, develop interesting research questions, be impartial to data, sceptical about conclusions and open to criticisms from our peers.
- We are taught good science is a reward in itself for improving our world.
- We need strong collaborations with diverse stakeholders in the public and private sectors, non-governmental organisations and civil society in order to identify and solve problems.
- There is greater attention to specific human dimensions: interactions between society and nature, along with university and diverse stakeholders.
- Many universities have a shared focus on research, teaching and service, but it is not enough to unite people for sustained collaborations.
- Academics share a commitment to excellence but not offer guidance for why and where to establish such excellence.
- We try to create an atmosphere of learning from successes and failures. There is no sure-fire formula to match research with societal needs.
- Older faculty are retiring but are being replaced by younger students who are able to move the initiative forward as a result of their skills to be interdisciplinary researchers.

Quote

"Two fundamental commitments [have emerged]: 1) In addition to the traditional focus on the biophysical components underpinning a problem, a much greater emphasis is needed on the human dimensions, including the complex interactions between society and nature; and 2) productive collaborations must be built between the university and diverse stakeholders to develop a sufficient understanding of sustainability problems and viable strategies for solving them."

Abstract

Academic scientists can transcend publish-or-perish incentives to help produce real-world solutions. Here's how one group did it.

APA Style Reference

Hart, D. D., & Silka, L. (2020). Rebuilding the ivory tower: bottom-up experiment in aligning research with societal needs. *Issues Sci Technol, 36*(3), 64-70. https://issues.org/aligning-research-with-societal-needs/ [accessed 14/08/2020]

You may also be interested in

- ➔ Promoting an open research culture (Nosek et al., 2015)
- ➔ Promoting Transparency in Social Science Research (Miguel et al., 2014)

Fallibility in Science: Responding to Errors in the Work of Oneself and Others (Bishop, 2018)

Main Takeaways.

- Senior colleagues and institutions try to hide errors.
- Retraction produces fear in a scientist, as it is associated with shame.
- Errors can be reduced with open science practices.
- Raw data can never be made completely open due to confidentiality but we can modify it to remove identifiable information, so other researchers reproduce what was done.
- Stigma needs to be removed concerning error detection.
- Making an analysis program open does not mean they are error-free. A reproducible result means when the same data is conducted, the same result is produced but this is incorrect.
- Researchers whose error is noticed may respond with denial, anger or silence. It damages reputation for integrity, thus should be done via journal but in practice rarely done smoothly.
- Findings may be due to methodological concern, as opposed to errors in calculation or scripts.
- A study without a control group is underpowered, uses unreliable measures or has a major confound.
- Methodological errors due to ignorance not bad faith. These could be honest errors in data, analysis or method compromising conclusions inferred.
- Replication is important, as confidence in robustness of a finding cannot depend on a single study. When there is a failure to replicate, we should uncover why this happened (e.g. contextual factors or research expertise).
- We should not say original researchers are incompetent, frauds etc., but we should not say that critics had malevolent motives and lack expertise. We need to be impartial.
- We should avoid bias and identify publications that are ignored, as positive findings produce more citations than null findings.
- Investigating misconduct is important but challenging. It is a difficult endeavour and requires evidence that takes time to accumulate.
- Academic institutions take an accusation of misconduct against a staff member seriously but takes a long time. We should consider whether people could have vested interests against this academic.
- We should not mock or abuse other scientists who make honest errors, as this would encourage poor research practices and people may be less likely to be open about these errors.

Quote

"Criticism is the bedrock of the scientific method. It should not be personal: If one has to point to problems with someone's data, methods, or conclusions, this should be done

without implying that the person is stupid or dishonest. This is important, because the alternative is that many people will avoid engaging in robust debate because of fears of interpersonal conflict—a recipe for scientific stasis. If wrong ideas or results are not challenged, we let down future generations who will try to build on a research base that is not a solid foundation. Worse still, when the research findings have practical applications in clinical or policy areas, we may allow wrongheaded interventions or policies to damage the well-being of individuals or society. As open science becomes increasingly the norm, we will find that everyone is fallible. The reputations of scientists will depend not on whether there are flaws in their research, but on how they respond when those flaws are noted." (p.6)

Abstract

This is a view on the fallibility of science, response to self-errors and errors made by others by Professor Dorothy Bishop. It contains discussion on how open science is the norm but being open and honest about oneself is not. It informs us that we should not mock or be hurtful to others concerning honest mistakes and that misconduct is a serious issue but we need to be supportive of both the researcher who is being accused and the individual who is accusing them.

APA Style Reference

Bishop, D. V. M. (2018). Fallibility in science: responding to errors in the work of oneself and others. *Advances in Methods and Practices in Psychological Science, 1*(3), 432-438. https://doi.org/10.1177/2515245918776632

You may also be interested in

➔ Don't let transparency damage science (Lewandowsky & Bishop, 2016)

➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)

➔ Quality Uncertainty Erodes Trust in Science (Vazire, 2017)

➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)

➔ Seven Easy Steps to Open Science: An Annotated Reading List (Crüwell et al., 2019)

➔ Promoting an open research culture (Nosek et al., 2015)

➔ Promoting Transparency in Social Science Research (Miguel et al., 2014)

➔ Signalling the trustworthiness of science should not be a substitute for direct action against research misconduct (Kornfeld & Titus, 2020)

➔ Reply to Kornfeld and Titus: No distraction from misconduct (Jamieson et al., 2020)

- ➔ Stop ignoring misconduct (Kornfeld & Titus, 2016)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)

Imagine a Research Future Defined by Open Values: Introducing the Open Science MOOC (Tennant, 2019)◈

Main Takeaways.

- Research does not work as well as it could. We have to be better scientists-we need to focus on reproducibility crisis,questionable research practices, lack of open access, wasteful research and flawed incentives and reward systems.
- Expectations are changing how to perform and communicate research.
- Modern research demands transparency and collaboration.
- At Open Science MOOC, a peer-to-peer and value-based community works towards better science for society.
- There is a community based on learning, sharing and collaboration that empowers researchers with knowledge and skills to save time and effort, solve research issues and advance global research.

Abstract

This is a blog by Jon Tennant, who argues about a research future defined by open values and introducing open science MOOC. It contains information about open science, the benefit of MOOC and that to solve global research we need to be a community based on learning, sharing and collaboration to empower researchers to learn new skills and consolidate further knowledge in order to save time and effort.

APA Style Reference

Tennant, J. (2019, Feb). Imagine a Research Future Defined by Open Values: Introducing the Open Science MOOC. *Generation R*. https://genr.eu/wp/imagine-a-research-future-defined-by-open-values-introducing-the-open-science-mooc/ https://doi.org/10.25815/6hyr-g583

You may also be interested in

➔ See

Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)

Main Takeaways.

- The paper communicates about the replication crisis and argues that the publication bias exaggerates effect size.
- Small sample sizes are used to conduct multiple studies. Researchers need to decide a priori to collect more data, exclude subjects, conditions, measures or observations and transform a measure or add a covariate to an analysis.
- Meta-analysis estimates an effect effect size that combines results across many experiments to correct for an effect-size exaggeration.
- We need to have improved statistical sophistication for researchers to test hypotheses about populations based on the samples- reward quality and accuracy, as opposed to quantity and flashiness.
- We need to be honest and advocate research-if the idea was inspired by data, we should state it. Report effect size with 95% confidence intervals around them.
- We need to create a detailed research plan providing a priori hypotheses, sample size planning, data exclusion rules, analyses, transformations, covariates etc. We need to be transparent and register a research plan (cf. Pre-registration).
- Pre-registration might be vague, leaving many implicit decisions to be made.. A vague plan helps researchers think through the project in advance and protect them from believing they had a priori hypothesis based on the data.
- Registered report should be considered. Stage 1 is based on the importance of the question to address and rigour of the methodology, if deemed worthy, it is in principle Stage 1 accepted. If work is completed as planned, Stage 2 is accepted, irrespective of the findings.
- However registered reports require a lot of resources to complete and results will possess value irrespective of findings.
- It is important to develop a lab manual that encourages replicability and transparency, it is important to make data, materials and analysis scripts transparent. They should be Findable, Accessible, Interoperable and Reusable (FAIR).
- Researchers place time and effort to make data, scripts and materials to be shared but the difficulty is in making them accessible.
- We also need to address constraints on generality of findings-published findings fail to replicate, failure to replicate could be due to differences in procedures, albeit original work did not indicate such differences modulate effect.

Quote

"The aim of the methodological reform movement is not to restrict psychological research to procedures that meet some fixed criterion of replicability. Replicability is not in itself the goal of science. Rather, the central aim of methodological reform is to make research reports more transparent, so that readers can gain an accurate understanding of how the

data were obtained and analyzed and can therefore better gauge how much confidence to place in the findings. A second aim is to discourage practices that contribute to effect-size exaggeration and false discoveries of non-existent phenomena. As per Vazire's analogy, the call is not for car dealerships to sell nothing but new Ferraris, but rather for dealers to be forthcoming about the weaknesses of what they have on the lot. The grand aim of science is to develop better, more accurate, and more useful understandings of reality. Methodological reform cannot in and of itself deliver on that goal, but it can help." (p.19).

Abstract

Psychological scientists strive to advance understanding of how and why we animals do and think and feel as we do. This is difficult, in part because flukes of chance and measurement error obscure researchers' perceptions. Many psychologists use inferential statistical tests to peer through the murk of chance and discern relationships between variables. Those tests are powerful tools, but they must be wielded with skill. Moreover, research reports must convey to readers a detailed and accurate understanding of how the data were obtained and analyzed. Research psychologists often fall short in those regards. This paper attempts to motivate and explain ways to enhance the transparency and replicability of psychological science. Specifically, I speak to how publication bias and p hacking contribute to effect-size exaggeration in the published literature, and how effect-size exaggeration contributes, in turn, to replication failures. Then I present seven steps toward addressing these problems: Telling the truth; upgrading statistical knowledge; standardizing aspects of research practices; documenting lab procedures in a lab manual; making materials, data, and analysis scripts transparent; addressing constraints on generality; and collaborating.

APA Style Reference

Lindsay, D. S. (2020). Seven steps toward transparency and replicability in psychological science. Canadian Psychology/Psychologie canadienne. Advance online publication. https://doi.org/10.1037/cap0000222 [ungated]

You may also be interested in

- ➔ The Statistical Crisis in Science (Gelman & Loken, 2014)
- ➔ Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)
- ➔ False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant (Simmons et al., 2011)
- ➔ A 21 Word Solution (Simmons et al., 2012)◈
- ➔ Quality Uncertainty Erodes Trust in Science (Vazire, 2017)
- ➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)

- ➔ Seven Easy Steps to Open Science: An Annotated Reading List (Crüwell et al., 2019)
- ➔ Many hands make tight work (Silberzahn & Uhlmann, 2015)
- ➔ Promoting an open research culture (Nosek et al., 2015)
- ➔ Promoting Transparency in Social Science Research (Miguel et al., 2014)
- ➔ Psychologists Are Open to Change, yet Wary of Rules (Fuchs et al., 2012)
- ➔ Registered Reports: Realigning incentives in scientific publishing (Chambers et al., 2015)
- ➔ Registered Reports: A new publishing initiative at Cortex (Chambers, 2013)
- ➔ Registered Reports: A step change in scientific publishing (Chambers, 2014)
- ➔ Registered reports: a method to increase the credibility of published results (Nosek & Lakens, 2014)
- ➔ Registered reports (Jamieson et al., 2019)
- ➔ Attitudes Toward Open Science and Public Data Sharing: A Survey Among Members of the German Psychological Society (Abele-Brehm et al., 2019)
- ➔ Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results (Wicherts et al., 2011)
- ➔ Constraints on Generality (COG): A Proposed Addition to All Empirical Papers (Simons et al., 2017)
- ➔ How scientists can stop fooling themselves (Bishop, 2020b)
- ➔ CJEP Will Offer Open Science Badges (Pexman, 2017)
- ➔ Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency (Kidwell et al., 2016)
- ➔ Trust Your Science? Open Your Data and Code (Stodden, 2011)◈
- ➔ Publishing Research With Undergraduate Students via Replication Work: The Collaborative Replications and Education Project (CREP; Wagge et al., 2019)
- ➔ Is science really facing a reproducibility crisis, and do we need it to? (Fanelli, 2018)
- ➔ Fallibility in Science: Responding to Errors in the Work of Oneself and Others (Bishop, 2018)
- ➔ A consensus-based transparency checklist (Aczel et al., 2020)
- ➔ Tell it like it is (Anon, 2020)

- ➔ Is pre-registration worthwhile? (Szollosi et al., 2020)

Measurement Schmeasurement: Questionable Measurement Practices and How to Avoid Them (Flake & Fried, 2019)◈

Main Takeaways.

- Questionable measurement practices are ignored in the literature and provide researchers the degrees of freedom to obtain the desired results. This poses a serious threat to cumulative psychological science.
- The research process is very flexible and exists. This is done implicitly (cf. Garden of forking paths).
- The definition of a measurement is broad and we need decisions to theorise the nature of a phenomenon to operationalise and analyse.
- Qualitative measurements do not need to be discussed but face questionable measurement practice. The present study measured transparent reporting practices.
- We need to provide information about measurements but it is lacking. This undermines internal and external validity, the statistical conclusion and construct made.
- Transparency does not make science more rigorous but facilitates it to evaluate it more thoroughly and accurately.
- We need to ask what the construct is measuring? Reporting what it is and how it is defined.
- We need to consider theoretical definitions, how it aligns with measures and the existing validity evidence for these measures.
- Once the construct is defined and measures are selected, we must report its origins, exact number of items, the wording, response format, short or long version, language, how it was presented, and hardware and software specification.
- We need to consider how it was transformed, which items form which scores and how was it calculated.
- We need to report decision rules to facilitate others to reproduce and evaluate work such as how it was calculated and why was it calculated.
- Before data collection, potential degrees of freedom are changing response type, changing response style or options, changing item wording or content.
- We need to declare and justify modifications that threaten the validity of inferences.
- When using existing scales, we need to make many decisions. When we create and use a scale-there are threats to its validity. We need to disclose why we created it, why we used it over an existing measure.
- If there is no or little validity, we discuss it as a limitation of the study. Without the transparency, we have to wonder what is significant or replicable. Transparency promotes new lines of inquiry, strengthening the validity of measures and improving the quality of work.

Quote

"The increased awareness and emphasis on QRPs, such as p-hacking, have been an important contribution to improving psychological science. We echo those concerns, but also see a grave need for broadening our scrutiny of current practices to include QMPs (Fried & Flake, 2018). Recalling our example of depression at the outset, even if we increase the sample size of our depression trials, adequately power our studies, pre-register our analytic strategies, and stop p-hacking — we can still be left wondering if we were ever measuring depression at all." (p.22)

Abstract

Academic scientists can transcend publish-or-perish incentives to help produce real-world solutions. Here's how one group did it.

APA Style Reference

Flake, J. K., & Fried, E. I. (2019). Measurement schmeasurement: Questionable measurement practices and how to avoid them. https://psyarxiv.com/hs7wm/

You may also be interested in

➔ See

A consensus-based transparency checklist (Aczel et al., 2020)

Main Takeaways.

- There is an erosion of trust that affects credibility of specific articles and the discipline as a whole.
- There is a lack of transparency, which is required to evaluate and reproduce findings, but also for research synthesis and meta analysis from the raw data.
- The lack of transparency is not meant to be deceptive or intentional. Human reasoning is prone to biases (e.g. confirmation bias and motivated reasoning).
- Few journals ask about statistical and methodological practices and transparency.
- Journals can support open practices by offering badges, using the transparency and openness promotion guidelines, promote availability of all research items, including data, materials and codes.
- The consensus-based transparency checklist can be submitted with the manuscript to provide critical information about the process to evaluate the robustness of a finding.
- The checklist can be modified by deleting, adding and rewording items with high level of acceptability and consensus with no strong counter argument for single items.
- Researchers can explain the choices at the end of each 36 section. There is a shortened 12-item version to reduce demands on the researchers' time and facilitate broader adoption that fosters transparency and asks authors to complete a 36-item list.

Abstract

We present a consensus-based checklist to improve and document the transparency of research reports in social and behavioural research. An accompanying online application allows users to complete the form and generate a report that they can submit with their manuscript or post to a public repository.

APA Style Reference

Aczel, B., Szaszi, B., Sarafoglou, A., Kekecs, Z., Kucharský, Š., Benjamin, D., ... & Ioannidis, J. P. (2020). A consensus-based transparency checklist. *Nature human behaviour, 4*(1), 4-6. https://doi.org/10.1038/s41562-019-0772-6

You may also be interested in

➔ How scientists can stop fooling themselves (Bishop, 2020b)

➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)

- Tell it like it is (Anon, 2020)
- Is pre-registration worthwhile? (Szollosi et al., 2020)

Tell it like it is (Anon, 2020)

Main Takeaways.

- A manuscript answers a question(s) based on findings and how they support or contradict hypotheses.
- Current research culture is defined by pressure to present research projects as conclusive narratives leave no room for ambiguity.
- Clean narratives represent a threat to validity and counter reality of what science looks like.
- Report only outcomes to confirm original predictions or excluding research findings that provide messy results.
- These questionable research practices create a distorted picture of research that prevents cumulative knowledge.
- Pre-registration has little value if not heeded or transparently reported.
- It is evident during peer review that a pre-registered analysis is inappropriate or suboptimal. Authors have to provide deviations and explain why they did these deviations.
- A pre-registered analysis plan is flawed, authors report results of pre-registered alongside new analyses.
- Authors report multi-study research papers and authors report all work they executed, irrespective of outcomes.

Quote

"*No research project is perfect; there are always limitations that also need to be transparently reported. In 2019, we made it a requirement that all our research papers include a limitations section, in which authors explain methodological and other shortcomings and explicitly acknowledge alternative interpretations of their findings… Science is messy, and the results of research rarely conform fully to plan or expectation. 'Clean' narratives are an artefact of inappropriate pressures and the culture they have generated. We strongly support authors in their efforts to be transparent about what they did and what they found, and we commit to publishing work that is robust, transparent and appropriately presented, even if it does not yield 'clean' narratives*" p.1

Abstract

Every research paper tells a story, but the pressure to provide 'clean' narratives is harmful for the scientific endeavour.

APA Style Reference

Anon (2020). Tell it like it is. *Nat Hum Behav 4,* 1. https://doi.org/10.1038/s41562-020-0818-9

You may also be interested in

- ➔ How scientists can stop fooling themselves (Bishop, 2020b)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)
- ➔ A consensus-based transparency checklist (Aczel et al., 2020)
- ➔ Tell it like it is (Anon, 2020)
- ➔ Is pre-registration worthwhile? (Szollosi et al., 2020)

Is pre-registration worthwhile? (Szollosi et al., 2020)

Main Takeaways.

- Pre-registration solves statistical problems and forces people to think more deeply about theories, methods, and analyses.
- Diagnosticity of statistical tests depend on how well statistical models map onto theories and improve statistical techniques does little to improve theories when mapping is weak.
- Models are useful depending on how accurately the theory is matched to the model. Many statistical models (e.g. general linear model) in psychology are poor estimates of the theory.
- Bad theories are pre-registered, predictions that are barely better than randomly picking an outcome but which bear out in experiments.
- There is no problem with post-hoc scientific inference when the theories are strong.
- Pre-registration does not improve theories but should allow them to think more deeply on how to improve theories through better planning, more precise operationalisation of constructs, and clear motivation for statistical planning.
- We should improve theories when encountering difficulties with pre-registration or when pre-registered predictions are wrong.
- Any improvement depends on a good understanding of how to improve a theory, and pre-registration provides no understanding. Pre-registration encourages thinking, but it is unclear whether the thinking is better or worse.
- Pre-registration could harm the progress in our field.
- Transparency is important but other solutions to solve problems and asking researchers to disclose studies and methods when publishing.
- Pre-registration should be an option to improve research, needing, rewarding or promoting it is not worthwhile.
- Scientific inference is the process to develop better theories.
- Statistical models are simplified mathematical abstractions of scientific problems, simplifications to aid scientific inference but to allow abstraction.
- Poor operationalisation, imprecise measurement, weak connection between theory and statistical method take precedence over problems of statistical inference.

Abstract

Proponents of preregistration argue that, among other benefits, it improves the diagnosticity of statistical tests. In the strong version of this argument, preregistration does this by solving statistical problems, such as family-wise error rates. In the weak version, it nudges people to think more deeply about their theories, methods, and analyses. We argue against both: the diagnosticity of statistical tests depend entirely on how well statistical models map onto underlying theories, and so improving statistical techniques does little to improve theories when the mapping is weak. There is also little

reason to expect that preregistration will spontaneously help researchers to develop better theories (and, hence, better methods and analyses).

APA Style Reference

Szollosi, A., Kellen, D., Navarro, D. J., Shiffrin, R., van Rooij, I., Van Zandt, T., & Donkin, C. (2020). Is Preregistration Worthwhile?. *Trends in cognitive sciences, 24*(2), 94.https://doi.org/10.1016/j.tics.2019.11.009

You may also be interested in

- ➔ How scientists can stop fooling themselves (Bishop, 2020b)
- ➔ Seven Steps Toward Transparency and Replicability in Psychological Science (Lindsay, 2020)
- ➔ A consensus-based transparency checklist (Aczel et al., 2020)
- ➔ Tell it like it is (Anon, 2020)
- ➔ Arrested theory development: The misguided distinction between exploratory and confirmatory research (Szollosi & Donkin, 2019)
- ➔ From pre-registration to publication: a non-technical primer for conducting meta-analysis to synthesize correlation data (Quintana, 2015)
- ➔ Pre-registration is Hard, And Worthwhile (Nosek et al., 2019)
- ➔ Easy preregistration will benefit any research (Mellor & Nosek, 2018)
- ➔ Preregistration of Modeling Exercises May Not Be Useful (MacEachern & Van Zandt, 2019)

Arrested theory development: The misguided distinction between exploratory and confirmatory research (Szollosi & Donkin, 2019)◈

Main Takeaways.

- It is difficult to provide arguments for why direct replication or pre-registration are important, we do not know whether they aim to achieve coincides with the goals of science.
- We need to think about theories. What makes a good theory? They are a good explanation about phenomena that occur in the world and need exploratory power. They allow us to correct flaws in our theories.
- Good explanation is hard to change consistent with other good theories and cannot be adapted to explain anything.
- A theory limited by existing knowledge without benefit of flexibility can tailor explanation to any possible observation.
- A good theory resists change. A theory can be criticised by argument and rejection of explanations.
- A good theory should observe what has been shown in the past and what we should observe in the future.
- Both pre-registrations and direct replications need researchers to fix what their theories predict in current forms but practice is futile when theory can be changed to explain experimental outcomes.
- If when testing good theories, we expect unchanged predictions and show, the theory remains good, if we do not observe it, it becomes problematic. Flexible theories are inconsequential.
- It is not possible to reconcile expected and observed data, no change should be required to explain data or for which any explanation is possible.
- Theory is rendered problematic and new theories are required.

Quotes

"Rather than quick-fix methodological suggestions, we advocate for an overhaul of how theories are developed and evaluated. In order to move forward, we need to confront the real problems of the field, wherein theories are not held accountable for their flexibility. This is not an easy feat to achieve, but we are optimistic that a deeper understanding of how science progresses can help behavioral scientists develop better theories" (p.10).

Abstract

Starting from the view that progress in science consists of the improvement of our theories, in the current paper we ask two questions: what makes a theory good,and how much do the current method-oriented solutions to the replication crisis contribute to the development of good theories? Based on contemporary philosophy of science, we argue that good theories are hard-to-vary: they (1) explain what they are supposed to explain,

(2) are consistent with other good theories, and (3) cannot easily be adapted to explain anything.Theories can be improved by identifying problems in them either by argument or by experimental test,and then correcting these problems by changing the theory. Importantly, such changes and the resultant theory should only be assessed based on whether they are hard-to-vary.An assessment ofthe current state of the behavioral sciences reveals that theory development is arrested by the lack of consideration for how easy it is to change theories to account for unexpected observations. Further, most of the current method-oriented solutions are unlikely to contribute much to the development of good theories, because they do not work towards eliminating this problem.Instead, they reward only temporary inflexibility in theories,and promote the assessment of theory change based on whether the theory was changed before (confirmatory) or after (exploratory) an experimental test, but not whether that change yields a hard-to-vary theory.Finally, we argue that these methodological solutions would become irrelevant if we turned our focus to the explicit aim of developing theories that are hard-to-vary.

APA Style Reference

Szollosi, A., & Donkin, C. (2019). Arrested theory development: The misguided distinction between exploratory and confirmatory research. https://doi.org/10.31234/osf.io/suzej

You may also be interested in

- ➔ Is pre-registration worthwhile? (Szollosi et al., 2020)

From pre-registration to publication: a non-technical primer for conducting meta-analysis to synthesize correlation data (Quintana, 2015)

Main Takeaways.

- Meta analysis is an integration of evidence from numerous studies. Effect size and measures of variance, numerous outcomes can be formed to compute a summary effect size. They have not been combined in a single resource targeting psychologists.
- This review will discuss how to conduct meta-analysis via PRISMA guidelines. Pre-registration allows a study rationale to be created that forms good systematic review and helps avoid bias. Meta analyses are prone to hypothesising after the results-HARKING.
- Pre-registration is important for submission, however, few journals need to consider meta-analysis registration. It is more important to pre-register meta analysis than clinical trial pre-registration as meta-analyses drive treatment for practice and health policy.
- Pre-registration avoids unintended meta-analysis duplication. Boolean operators and search limits help literature research. Databases are available with researchers to choose most suitable sources. Numerous scientists use duplicate search terms within two or more databases to cover numerous sources.
- Gray literature is difficult with numerous universities posting dissertations. Fields and research questions exclude blanket recommendation. Meta-analyses detail search strategy for study protocols and methods.
- Model selection centre around assumptions of study homogeneity, how much variation of studies can be due to variations in true effect size. Variation is from random error and true study heterogeneity.
- Forest plots visualise effect sizes and confidence intervals from included studies with summary effect size.
- Publication bias produces stronger effects size and are more likely to be published and included in meta-analysis. A funnel plot is a visual tool to investigate potential publication bias in meta analyses.
- Funnel plots offer a useful visualisation for potential publication bias, it is important to consider asymmetry may represent other types of bias like study quality, location bias and study size.
- Funnel plots suffer from subjective measures of potential publication bias. Two tests used to calculate objective measures of potential bias: trim and fill method and moderating variables.

Abstract

Starting from the view that progress in science consists of the improvement of our theories, in the current paper we ask two questions: what makes a theory good,and how much do the current method-oriented solutions to the replication crisis contribute to the

development of good theories? Based on contemporary philosophy of science, we argue that good theories are hard-to-vary: they (1) explain what they are supposed to explain, (2) are consistent with other good theories, and (3) cannot easily be adapted to explain anything.Theories can be improved by identifying problems in them either by argument or by experimental test,and then correcting these problems by changing the theory. Importantly, such changes and the resultant theory should only be assessed based on whether they are hard-to-vary.An assessment ofthe current state of the behavioral sciences reveals that theory development is arrested by the lack of consideration for how easy it is to change theories to account for unexpected observations. Further, most of the current method-oriented solutions are unlikely to contribute much to the development of good theories, because they do not work towards eliminating this problem.Instead, they reward only temporary inflexibility in theories,and promote the assessment of theory change based on whether the theory was changed before (confirmatory) or after (exploratory) an experimental test, but not whether that change yields a hard-to-vary theory.Finally, we argue that these methodological solutions would become irrelevant if we turned our focus to the explicit aim of developing theories that are hard-to-vary.

APA Style Reference

Quintana, D. S. (2015). From pre-registration to publication: a non-technical primer for conducting a meta-analysis to synthesize correlational data. Frontiers in psychology, 6, 1549. https://doi.org/10.3389/fpsyg.2015.01549

You may also be interested in

- ➔ Is pre-registration worthwhile? (Szollosi et al., 2020)
- ➔ Pre-registration is Hard, And Worthwhile (Nosek et al., 2019)
- ➔ Easy preregistration will benefit any research (Mellor & Nosek, 2018)
- ➔ Preregistration of Modeling Exercises May Not Be Useful (MacEachern & Van Zandt, 2019)

Pre-registration is Hard, And Worthwhile (Nosek et al., 2019)

***Main Takeaways*.**

- Pre-registration allows us to make exploratory and confirmatory analyses.
- Pre-registration allows us to make the transparent uncertainty more certain, how many statistical tests were conducted and familywise error rate to be corrected.
- Pre-registration reduces influence of publication bias and pre-registration is a skill that needs experience to be improved.
- Pre-registration promotes intellectual humility and better calibration of scientific claims.
- It allows us to provide information on how methodology is implemented, how hypotheses are tested, the exclusion rules, how variables are combined and what to use concerning the statistical model, covariates and characteristics.
- Pre-registration converts general sense into precise and explicit plans that predict what has not yet occurred and decide what will be done.
- It allows us to stop data collection. What are the steps required to assess questions of interest? What are the outcomes?
- Having a plan is better than no plan, sharing plans to advance is better than not sharing them.
- Planning will improve and benefits will increase for oneself and consumers of research.
- Deviations make it harder to interpret with confidence what occurred to what was planned.
- Transparency is important and all deviations should be reported, this is difficult due to narrative coherence, reviewer expectations and word limits.
- We need to maximise credibility of reporting findings when possible, update pre-registration, deviations before observing data, mention all planned analyses to explain why a planned analysis was not reported.
- Use supplements to share in full not hide inconvenient information and during analysis.

Abstract

Preregistration clarifies the distinction between planned and unplanned research by reducing unnoticed flexibility. This improves credibility of findings and calibration of uncertainty. However, making decisions before conducting analyses requires practice. During report writing, respecting both what was planned and what actually happened requires good judgment and humility in making claims.

APA Style Reference

Nosek, B. A., Beck, E. D., Campbell, L., Flake, J. K., Hardwicke, T. E., Mellor, D. T., ... & Vazire, S. (2019). Preregistration is hard, and worthwhile. *Trends in cognitive sciences, 23*(10), 815-818.https://doi.org/10.1016/j.tics.2019.07.009 [ungated]

You may also be interested in

- ➔ Is pre-registration worthwhile? (Szollosi et al., 2020)
- ➔ From pre-registration to publication: a non-technical primer for conducting meta-analysis to synthesize correlation data (Quintana, 2015)
- ➔ Easy preregistration will benefit any research (Mellor & Nosek, 2018)
- ➔ Preregistration of Modeling Exercises May Not Be Useful (MacEachern & Van Zandt, 2019)

Easy preregistration will benefit any research (Mellor & Nosek, 2018)

Main Takeaways.

- It is important to commit to study participants, the public and basic science, we need to be transparent, rigorous and reproducible.
- Pre-registration discriminates between confirmatory and exploratory research.
- We need to address publication bias. Widespread pre-registration addresses key contributions and increases interpretability of most empirical research.
- New practices need to improve and accelerate, not interfere with it, knowledge accumulation.
- Open science framework has more than 10000 registrations and supports multiple registration formats and iterative registrations.

Abstract

This view was written by David Mellor and Professor Brian Nosek. They discuss the benefits of pre-registration and how it addresses publication bias. Open-science framework provides support to multiple registration formats..

APA Style Reference

Mellor, D. T., & Nosek, B. A. (2018). Easy preregistration will benefit any research. *Nature Human Behaviour, 2*(2), 98-98. https://doi.org/10.1038/s41562-018-0294-7

You may also be interested in

- ➔ Is pre-registration worthwhile? (Szollosi et al., 2020)
- ➔ From pre-registration to publication: a non-technical primer for conducting meta-analysis to synthesize correlation data (Quintana, 2015)
- ➔ Pre-registration is Hard, And Worthwhile (Nosek et al., 2019)
- ➔ Preregistration of Modeling Exercises May Not Be Useful (MacEachern & Van Zandt, 2019)

Preregistration of Modeling Exercises May Not Be Useful (MacEachern & Van Zandt, 2019)

***Main Takeaways*.**

- The present study focuses on modelling and data analysis and how each round improves analysis that builds richer understanding of data and processes that give rise to data.
- Powerful software and improved graphical capabilities allows us to explore many more features of data.
- The ease with which data is transformed and cleaned, with which a model can be fit may lead to overfitting.
- Model development is intrinsically exploratory and creative.
- The present article disagrees with pre- and post-registration of models. In highly exploratory settings, there is greater difficulty to pre-register a model and analysis.
- Modelling depends on the modeller's perspective and data collected. Each author performs exploratory analysis and settles on the same transformation for the response variable.
- Two authors would not realistically pre-register models or exploratory plans except in general terms.
- When the model is combined with the Bayesian model averaging, the overall model provides a better description of the entire dataset than any single model on its own.
- Reality is too complicated and covariates are sparse enough that it would be a challenge to identify the right model. Models are tools. Different models are used differently to different ends.
- Model construction and development depend on analysing and re-analysing a dataset to determine which of its properties are crucial to understand a phenomenon or make predictions.
- Confirmatory model implies truth to be discovered among models in competition but there is favouritism, as the models invested develops, how we view the world, ease of implementation and so on.
- One model is not true in that some data is captured, and other data is not captured.
- Bayesian methods need to be used, as datasets grow. If pre-registration is required, Bayesian analysts pay attention to influence prior distributions such as Bayes factor.
- Underfitting of the data is as problematic as overfitting. Pre-registration of model development will lead to less and less creative and exploratory analyses that are achieved by needing publication of raw data and code.
- Psychology departments should devote more resources to training in quantitative areas and training include explicit content on under- and over-modelling. Also, we should partner with the statistics department to improve our modelling skills.

Abstract

This is a commentary on Lee et al.'s (2019) article encouraging preregistration of model development, fitting, and evaluation. While we are in general agreement with Lee et al.'s characterization of the modeling process, we disagree on whether preregistration of this process will move the scientific enterprise forward. We emphasize the subjective and exploratory nature of model development, and point out that "under-modeling" of data (relying on black-box approaches applied to data without data exploration) is as big a problem as "over-modeling" (fitting noise, resulting in models that generalize poorly). We also note the potential long-run negative impact of preregistration on future generations of cognitive scientists. It is our opinion that preregistration of model development will lead to less, and to less creative, exploratory analysis (i.e., to more under-modeling), and that Lee at al.'s primary goals can be achieved by requiring publication of raw data and code. We conclude our commentary with suggestions on how to move forward.

APA Style Reference

MacEachern, S. N., & Van Zandt, T. (2019). Preregistration of modeling exercises may not be useful. *Computational Brain & Behavior, 2*(3-4), 179-182. https://doi.org/10.1007/s42113-019-00038-x

You may also be interested in

- Is pre-registration worthwhile? (Szollosi et al., 2020)
- From pre-registration to publication: a non-technical primer for conducting meta-analysis to synthesize correlation data (Quintana, 2015)
- Easy preregistration will benefit any research (Mellor & Nosek, 2018)
- Pre-registration is Hard, And Worthwhile (Nosek et al., 2019)

Sample size and the fallacies of classical inference (Friston, 2013)

Main Takeaways.

- Authors and reviewers need to consider sample size and effect size. One should get as much data as possible.
- Trivial effect sizes can be resolved by reporting confidence intervals or interval null hypotheses. Simple point hypotheses are tested with p value.
- Best studies use large numbers of subjects and report them in terms of confidence intervals.
- Large sample sizes increase the power of model comparison.
- Increasing sample size will increase the number of true positives.
- If trivial effect sizes prevail, the positive predictive values measure the proportion of significant results that are not important.
- A trivial effect size does not mean small effect.
- Cross-validation in neuroimaging needs further discussion to validate a model to predict using new observations.

Abstract

I would like to thank Michael Ingre, Martin Lindquist and their co-authors for their thoughtful responses to my ironic Comments and Controversies piece. I was of two minds about whether to accept the invitation to reply — largely because I was convinced by most of their observations. I concluded that I should say this explicitly, taking the opportunity to consolidate points of consensus and highlight outstanding issues.

APA Style Reference

Friston, K. (2013). Sample size and the fallacies of classical inference. *Neuroimage, 81*, 503-504.https://doi.org/10.1016/j.neuroimage.2013.02.057

You may also be interested in

- ➔ Why small low-powered studies are worse than large high-powered studies and how to protect against “trivial” findings in research: Comment on Friston (2012) (Ingre, 2013)
- ➔ Ironing out the statistical wrinkles in “ten ironic rules” (Lindquist et al., 2013)
- ➔ Ten ironic rules for non-statistical reviewers (Friston, 2012)

Why small low-powered studies are worse than large high-powered studies and how to protect against "trivial" findings in research: Comment on Friston (2012) (Ingre, 2013)

Main Takeaways.

- Small underpowered studies provide better evidence due to a lack of small and trivial effect sizes, leading to poor research practices.
- Underpowered studies are less likely to find a true effect in data and failure does not come without consequences.
- Failure to detect true effects means when significant findings are reported-it may be due to type I error.
- Findings from small low-powered studies are weaker than high-powered studies due to the fact that poor statistical power increases false positive.
- Small studies have poor estimates and produce false positives with large effect sizes.
- Researchers should make use of at least one additional statistic value (e.g. t value, effect size or confidence intervals).
- We need to be cautious about effect size seen as trivial, meaningful or unimportant.

Quote

"From a strictly scientific point of view, you can never have too much precision, and consequently, never too many subjects or too much statistical power (unless a researcher is doing something wrong when reporting and interpreting data). The limiting factors are cost (time, resources and money) and potential harm for the subjects involved in the study. The real question you need to ask is how much cost and harm you can afford to get as good answer as possible." (p.498)

Abstract

It is sometimes argued that small studies provide better evidence for reported effects because they are less likely to report findings with small and trivial effect sizes (Friston, 2012). But larger studies are actually better at protecting against inferences from trivial effect sizes, if researchers just make use of effect sizes and confidence intervals. Poor statistical power also comes at a cost of inflated proportion of false positive findings, less power to "confirm" true effects and bias in reported (inflated) effect sizes. Small studies ($n = 16$) lack the precision to reliably distinguish small and medium to large effect sizes ($r < .50$) from random noise ($\alpha = .05$) that larger studies ($n = 100$) does with high level of confidence ($r = .50$, $p = .00000012$). The present paper presents the arguments needed for researchers to refute the claim that small low-powered studies have a higher degree of scientific evidence than large high-powered studies.

APA Style Reference

Ingre, M. (2013). Why small low-powered studies are worse than large high-powered studies and how to protect against “trivial” findings in research: Comment on Friston (2012). *Neuroimage, 81*, 496-498.https://doi.org/10.1016/j.neuroimage.2013.03.030

You may also be interested in

- ➔ Sample size and the fallacies of classical inference (Friston, 2013)
- ➔ Ironing out the statistical wrinkles in “ten ironic rules” (Lindquist et al., 2013)
- ➔ Ten ironic rules for non-statistical reviewers (Friston, 2012)

Ironing out the statistical wrinkles in "ten ironic rules" (Lindquist et al., 2013)

Main Takeaways.

- A Collaborative environment encourages good and bad ideas about statistics. Similar to large sample sizes, small sample sizes can detect large effects. However, small effects cannot be detected in small sample sizes, this requires larger sample sizes.
- Large sample sizes are prone to biases masking small effects.
- It is difficult to interpret significant results in small samples and makes it difficult to check certain assumptions or perform sensitivity analyses.
- Increasing sample size, leads to small effects being significant, producing more positive effects.
- More data allows us to detect subtle effects.
- Hypothesis test should not be exploratory, but confirmatory.
- Small sample sizes are better when considering important non-statistical issues such as the lives of animals or side effects.
- Hypothesis testing cannot discriminate between important and trivial effects.
- We need estimates, confidence intervals, exploratory plots and other summaries of data with careful scientific thinking.
- We need to consider Lindley's paradox-Bayesian and frequentist approaches lead to different results for certain types of prior distributions.

Abstract

The article "Ten ironic rules for non-statistical reviewers" (Friston, 2012) shares some commonly heard frustrations about the peer-review process that all researchers can identify with. Though we found the article amusing, we have some concerns about its description of a number of statistical issues. In this commentary we address these issues, as well as the premise of the article.

APA Style Reference

Lindquist, M. A., Caffo, B., & Crainiceanu, C. (2013). Ironing out the statistical wrinkles in "ten ironic rules". *Neuroimage, 81*, 499-502. https://doi.org/10.1016/j.neuroimage.2013.02.056

You may also be interested in

- ➔ Sample size and the fallacies of classical inference (Friston, 2013)
- ➔ Why small low-powered studies are worse than large high-powered studies and how to protect against "trivial" findings in research: Comment on Friston (2012) (Ingre, 2013)
- ➔ Ten ironic rules for non-statistical reviewers (Friston, 2012)

Ten ironic rules for non-statistical reviewers (Friston, 2012)

Main Takeaways.

- Reviewers may not have adequate statistical expertise to provide a critique during peer review to reject a manuscript.
- Handling editors are happy to decline a paper and are placed under pressure to maintain a high rejection rate.
- All journals maximise rejection rates, increase quality of submission and impact factor to explain long-term viability. There are ten rules to follow.
- 1. You must dismiss self-doubt. You may feel underqualified. An expert reviewer's opinion cannot be challenged with reference to the paper. Write with authority in the firm and friendly fashion.
- 2. Avoid dispassionate statements. Avoid phrases that are dispassionate but use phrases such as I feel and I do not trust. No one questions feelings.
- 3. Submit comments late. It brings delays to the editorial process and creates an air of frustration.
    - Create an impression of being busy and that you gave the paper due consideration due to thinking about it over a month.
- 4. Discuss under-sampled groups. base inference based on less than 16 subjects and authors did not resort to usual anecdotes that charm editors.
- 5. Over-sampled subjects. If the number of subjects exceeds 32, one can use a rarer criticism.
- 6. Untenable assumptions or use of non-parametric analysis. Neyman-Pearson lemma is more conservative than the original likelihood ratio, making significant results disappear.
- 7. Question validity or cross validation: question the fundamentals of the statistical analysis and move the author out of their comfort zone.
- 8. Exploit superstitious thinking. Authors should report effect sizes to supplement inferential statistics and enable a Catch-22 for authors who have not reported effect sizes.
- 9. Missing procedures: Neuroimaging is blessed with specialist procedures and you can ask for additional procedures.
- 10. Last Resort: Make no personal agenda and appreciate the author's efforts. Make the final value judgment should secure the desired editorial decision but reject in the end, saying it does fit the ethos, aims or mission of the journal.
- We should take results from small samples more seriously than the results in large sample studies.
- Studies with 50 participants may make trivial effects significant.
- Overpowered studies lose integrity and should be interpreted with caution.
- Sample size must be big enough that the effect of scientific significance will be important. The study must not be too big, an effect of little scientific importance is statistically detectable.

Abstract

As an expert reviewer, it is sometimes necessary to ensure a paper is rejected. This can sometimes be achieved by highlighting improper statistical practice. This technical note provides guidance on how to critique the statistical analysis of neuroimaging studies to maximise the chance that the paper will be declined. We will review a series of critiques that can be applied universally to any neuroimaging paper and consider responses to potential rebuttals that reviewers might encounter from authors or editors.

APA Style Reference

Friston, K. (2012). Ten ironic rules for non-statistical reviewers. *Neuroimage, 61*(4), 1300-1310. https://doi.org/10.1016/j.neuroimage.2012.04.018

You may also be interested in

- ➔ Sample size and the fallacies of classical inference (Friston, 2013)
- ➔ Why small low-powered studies are worse than large high-powered studies and how to protect against “trivial” findings in research: Comment on Friston (2012) (Ingre, 2013)
- ➔ Ironing out the statistical wrinkles in “ten ironic rules” (Lindquist et al., 2013)

Using OSF to Share Data: A Step-by-Step Guide (Soderberg, 2018)

Main Takeaways.

- Materials should findable, accessible, interoperable and reusable forms. Researchers should look for repositories to decide where and how to share their data.
- A repository should contain unique and persistent identifiers, so data can be cited.
- The data is publicly searchable with licenses clarifying how data is reused.
- Rich meta-data descriptions are provided to allow data to be understandable and reusable.
- Open science framework is a free and open-source Web tool to help researchers collaboratively manage, store and share the research process and the files related to their research.
- Step 1: create an account on https://osf.io
- Step 2: Sign-in your account. Enter name and password or login through your institution.
- Step 3: Create a project. Press the green button to create a new project.
- Step 4: Add Collaborators to the project.
- Click on Contributors and press +Add green button. Search for contributors by name and click on the green + button.
- If a collaborator does not come up in search, add them to the project by clicking add as an unregistered contributor link.
- Step 5: upload files- there is no more than 5GB.
- Step 6: Add a description of the project so you and other users know what files relate to.
- Step 7: Add a License: reuse is one of the main purposes of the data sharing. Other researchers know how they are allowed to reuse your work.
- Step 8: Add component: place data, analysis script and study materials should be placed in the project. Click the Add contributors checkbox before clicking on the green Create button.
- Step 9: Share your project with reviewers. Project is set up that you may want or need to give reviewers access to the contents of your project before you make it public.
- Step 10: Make a project public. To make a project public, press the "make public" button in the top right corner of the project page. Anyone will be able to view and download all files.
- Step 11: Reference open science files in your work. Include the links in the manuscript, lab website or the published article. Ensure readers find the specific links in manuscript, lab website or published article to make the data accessible and useful.

Abstract

Sharing data, materials, and analysis scripts with reviewers and readers is valued in psychological science. To facilitate this sharing, files should be stored in a stable location, referenced with unique identifiers, and cited in published work associated with them. This Tutorial provides a step-by-step guide to using OSF to meet the needs for sharing psychological data.

APA Style Reference

Soderberg, C. K. (2018). Using OSF to share data: A step-by-step guide. *Advances in methods and practices in psychological science, 1*(1), 115-120. https://doi.org/10.1177/2515245918757689

You may also be interested in

- ➔ Trust Your Science? Open Your Data and Code (Stodden, 2011)
- ➔ Attitudes Toward Open Science and Public Data Sharing: A Survey Among Members of the German Psychological Society (Abele-Brehm et al., 2019)
- ➔ Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results (Wicherts et al., 2011)

On supporting early-career black scholars (Roberson, 2020)

Main Takeaways.

- Non-black researchers need to take immediate support for early-career Black scholars.
- Black doctoral students hear an offensive joke and try to push against this but there is a racist disciplinary norm.
- You silently agreed and followed this up to support them. These issues push Black scholars out of academia, since these silent signals indicate the space was not created for them.
- We must challenge white supremacy in terms of science, it will be costly to our careers but it is worse for early-career black scholars, who face an onslaught of racist micro- and macro-aggressions on a daily basis.
- We should be proactive in our outreach. We should invite early-career Black scholars, if they have expertise to improve our research project. Our careers and science will benefit from this help.
- Do not encourage us to apply, provide us material support, share funded grants, work with us on developing aims page and write a persuasive letter of support. Invite Black scholars to series, colloquia or conference programs.
- Inviting them will increase their credibility as experts and expand the audience's familiarity with their scholarship. Manels are now being prohibited but we need to eliminate all-white speaker panels.
- Educate yourself on rising Black scholars in your field, learn from early-career Black researchers, investigate journals that publish their scholarships, be familiar with the Black community's professional societies, affinity groups and diversify your following list on Twitter.
- Incorporate their work into your syllabi. This is necessary to eliminate structural racism. However, it requires individuals with the most amount of power. These steps will promote Black people to thrive among trainees and early-career scholars.
- This will remove barriers to promote a more inclusive environment!

Abstract

Professor Mya Roberson provides a detailed commentary about the struggles that Black people encounter in academia and starting steps to eliminate structural racism.

APA Style Reference

Roberson, M. L. (2020). On supporting early-career Black scholars. Nature Human Behaviour, 1-1. https://doi.org/10.1038/s41562-020-0926-6

You may also be interested in

➔ See

On the persistence of low power in psychological science (Vankov et al., 2014)

Main Takeaways.

- We adhere to 5% for false positives but we pay little attention to false negatives.
- Scientists are humans and respond to incentives where personal success is related to quality and quantity of publications produced.
- A single transformative study might produce great prestige if accepted in a highly regarded journal.
- This prestige depends on whether a high-risk high-reward strategy is used. If the desired results are not produced, the journal may not accept it for publication.
- Authors may conduct a larger number of smaller studies, likely to produce publishable findings instead of using limited resources in a smaller number of larger studies.
- If experiments are repeated with the same sample size, on average, 50% of the time, these studies will be replicated.
- There is a need for structural change that is necessary to enforce rigorous requirements and editors of guidelines.
- Journals are introducing registered reports or registered replication reports.

Abstract

A comment by Dr Ivan Vankov, Professors Jeffrey Bowers and Marcus Munafo on the persistence of low power in psychological sciences. They discuss issues concerning false negatives, the importance of highly-regarded journals and that power is an issue to be discussed. They state that we need structural changes in journals in order to avoid the replicability crisis.

APA Style Reference

Vankov, I., Bowers, J., & Munafò, M. R. (2014). Article commentary: On the persistence of low power in psychological science. *Quarterly journal of experimental psychology, 67*(5), 1037-1040. https://doi.org/10.1080/17470218.2014.885986

You may also be interested in

➔ Is science really facing a reproducibility crisis, and do we need it to? (Fanelli, 2018)

➔ Registered Reports: Realigning incentives in scientific publishing (Chambers et al., 2015)

➔ Registered Reports: A new publishing initiative at Cortex (Chambers, 2013)

➔ Registered Reports: A step change in scientific publishing (Chambers, 2014)

➔ Registered reports: a method to increase the credibility of published results (Nosek & Lakens, 2014)
➔ Registered reports (Jamieson et al., 2019)
➔ Rein in the four horsemen of irreproducibility (Bishop, 2019)

Publication Decisions and their possible effects on inferences drawn from tests of significance or vice versa (Sterling, 1959)

Main Takeaways.

- There is a risk to reject the null hypothesis.
- Test of significance evaluates observed difference and the probability of results.
- Depending on the confidence of methodology and data collection, readers can reject or accept the null hypothesis.
- Acceptance and rejection of null hypothesis is taken at $p < .05$.
- When a fixed level is used as a criterion, it may result in embarrassing and surprising results.
- What are the inferences drawn from the statistical tests, if the reader is not aware of the dependent variable? Can the reader justify the same level of significance? The author will reject the null hypothesis.
- What risks for type I or II error happen by rejecting null hypothesis? Risk from the author cannot be accepted at face value once printed.

Abstract

There is some evidence that in fields where statistical tests of significance are commonly used, research which yields non-significant results is not published. Such research being unknown to other investigators may be repeated independently until eventually by chance a significant result occurs - an "error of the first kind" - and is published. Significant results published in these fields are seldom verified by independent replication. The possibility thus arises that the literature of such a field consists in substantial part of false conclusions resulting from errors of the first kind in statistical tests of significance.

APA Style Reference

Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance—or vice versa. *Journal of the American statistical association*, *54*(285), 30-34. https://doi.org/10.1080/01621459.1959.10501497

You may also be interested in

- ➔ See

Replicability as a Publication Criterion (Lubin, 1957)

***Main Takeaways*.**

- How can we reduce publication lag?
- Perform replications to show the results were repeated.
- Replicability and generalisability are not new criteria and these need to be performed to judge the rigour of an article.
- Articles with replication designs are not adequate to the editor and made the lowest publication priority.
- If results are replicated, it is not important to discuss other small variables.
- If these results are not replicated, it could result from several factors (e.g. time of day, temperature, kind of food eaten etc.)

Abstract

A comment by Dr Ardie Lubin on replicability being a criterion of publication. Replications are seen as fundamental but may not be seen as adequate to the editor, thus are seen as the lowest publication priority. However, replications are important to remove any trivial variables that may explain the findings.

APA Style Reference

Lubin, A. (1957). Replicability as a publication criterion. *American Psychologist, 12*(8), 519-520. https://doi.org/10.1037/h0039746

You may also be interested in

➔ See

The life of p: “Just significant” results are on the rise (Leggett et al., 2013)

Main Takeaways.

- We use computers to execute analyses to produce precise p values. Modern software enables simple and instantaneous calculations, allowing researchers to monitor data while collecting it.
- This ease to analyse data produces issues of optional stopping and selective exclusion of outliers.
- All practices manipulate p-values and drive them towards significance.
- The current study measured whether there is a spike in p values at .05 and whether there is an over-representation of p values below .05 for 2005 than 1965.
- *Method*: P values were collected for all articles published between 1965 and 2005.
- *Method*: P values that were incorrectly reported, categorised as p < .05 and .01 or values reported only 2 decimal places were recalculated to an accuracy of 6 decimal places.
- *Method*: Insufficient information available to determine the exact p values, thus data was excluded from analyses. P values around significance cut-off point of .05 values and between .01 and .10 were the main focus.
- *Results*: Frequency of p <.05 was greater than expected compared to p frequencies in other ranges.
- *Results*: Over-representation found for values published in both 1965 and 2005, much greater for the latter.
- *Results*: P values close but over .05 were more likely to be rounded down or incorrectly reported in 2005 than in 1965.
- Magnitude of spike at .05 is larger in recent articles than in 1965. Majority was inaccurately rounding p values.
- Changes in how statistical analyses are executed due to shifting research climates may explain this spike.
- Values outside this cut-off point should not be seen as significant. Trends should not be seen as trends but non-significant.
- Advances in analytical procedure make it easier to engage in suboptimal research practices.
- We need to use confidence intervals and effect sizes, mandatory methods disclosure and registered reporting.
- Reduce focus on p values and encourage use of optimal research practices. P values along is prone to human fallibility.

Abstract

Null hypothesis significance testing uses the seemingly arbitrary probability of .05 as a means of objectively determining whether a tested effect is reliable. Within recent psychological articles, research has found an overrepresentation of p values around this

cut-off. The present study examined whether this overrepresentation is a product of recent pressure to publish or whether it has existed throughout psychological research. Articles published in 1965 and 2005 from two prominent psychology journals were examined. Like previous research, the frequency of p values at and just below .05 was greater than expected compared to p frequencies in other ranges. While this overrepresentation was found for values published in both 1965 and 2005, it was much greater in 2005. Additionally, p values close to but over .05 were more likely to be rounded down to, or incorrectly reported as, significant in 2005 than in 1965. Modern statistical software and an increased pressure to publish may explain this pattern. The problem may be alleviated by reduced reliance on p values and increased reporting of confidence intervals and effect sizes.

APA Style Reference

Leggett, N. C., Thomas, N. A., Loetscher, T., & Nicholls, M. E. (2013). The life of p:" just significant" results are on the rise. *Quarterly journal of experimental psychology (2006), 66*(12), 2303. https://doi.org/10.1080/17470218.2013.863371

You may also be interested in

- How scientists can stop fooling themselves (Bishop, 2020b)
- The Statistical Crisis in Science (Gelman & Loken, 2014)
- Only Human: Scientists, Systems, and Suspect Statistics A review of: Improving Scientific Practice: Dealing With The Human Factors, University of Amsterdam, Amsterdam, September 11, 2014 (Hardwicke et al., 2014)
- A 21 Word Solution (Simmons et al., 2012)
- Rein in the four horsemen of irreproducibility (Bishop, 2019)
- False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant (Simmons et al., 2011)

Experimental power comes from powerful theories – the real problem in null hypothesis testing (Ashton, 2013)

Main Takeaways.

- Although null hypothesis testing is a powerful tool for decision making.
- Power analysis of an effect size must be carried out before an experiment to test the null hypothesis against an alternative-based hypothesis based on effect size.
- The alternative hypothesis is left open with the effect size being estimated from data.
- A vague open hypothesis in neuroscience is open to being amenable, we need more replications.
- Alternative hypothesis tests depend on theory and the problems being investigated.
- Observations are retained if below significance threshold moderated by spurious statistical significance if open-ended search for more nuanced levels of effect lead to false positives.

Abstract

A commentary by John C. Ashton who discusses the paper written by Professor Kate Button on small sample sizes. Ashton argues that power analyses and effects sizes should be used to estimate the alternative hypothesis.

APA Style Reference

Ashton, J. C. (2013). Experimental power comes from powerful theories—the real problem in null hypothesis testing. *Nature Reviews Neuroscience, 14*(8), 585-585. https://doi.org/10.1038/nrn3475-c2

You may also be interested in

- ➔ Negative results are disappearing from most disciplines and countries (Fanelli, 2011)
- ➔ Negativity towards negative results: a discussion of the disconnect between scientific worth and scientific culture (Matosin et al., 2014)

Negative results are disappearing from most disciplines and countries (Fanelli, 2011)

***Main Takeaways*.**

- There is a worry that scientific knowledge might have to endure the loss of negative results.
- *Method*: 4600 papers in all disciplines between 1990 and 2007 were used, including variables such as frequency of papers to test a hypothesis and a report to support it.
- *Method*: Country of location was included, information on year of publication and country was coded. Whether the evidence was positive or negative.
- *Results*: Frequency of positive findings increased between 1990 and 2007 by 22%. This increase was larger in social and some biomedical disciplines.
- *Results*: There were fewer positive results published by in American than Asian countries. More positive results in American than European countries.
- Negative results decreased in frequency across disciplines due to publication bias.
- The authors seem to suggest that science is now closer to truth today than 20 years ago.
- There is an editorial bias that favours the United States that enables them to publish as many or more negative results than any other country, not fewer. The United States has a stronger bias against negative findings than Europe.
- Self-correcting principles do not work efficiently where theoretical predictions are less accurate, methodologies are less codified and true replications are rare.

Abstract

Concerns that the growing competition for funding and citations might distort science are frequently discussed, but have not been verified directly. Of the hypothesized problems, perhaps the most worrying is a worsening of positive-outcome bias. A system that disfavours negative results not only distorts the scientific literature directly, but might also discourage high-risk projects and pressure scientists to fabricate and falsify their data. This study analysed over 4,600 papers published in all disciplines between 1990 and 2007, measuring the frequency of papers that, having declared to have "tested" a hypothesis, reported a positive support for it. The overall frequency of positive supports has grown by over 22% between 1990 and 2007, with significant differences between disciplines and countries. The increase was stronger in the social and some biomedical disciplines. The United States had published, over the years, significantly fewer positive results than Asian countries (and particularly Japan) but more than European countries (and in particular the United Kingdom). Methodological artefacts cannot explain away these patterns, which support the hypotheses that research is becoming less pioneering and/or that the objectivity with which results are produced and published is decreasing.

APA Style Reference

Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics, 90*(3), 891-904. https://doi.org/10.1007/s11192-011-0494-7

You may also be interested in

- ➔ Experimental power comes from powerful theories – the real problem in null hypothesis testing (Ashton, 2013)
- ➔ Negativity towards negative results: a discussion of the disconnect between scientific worth and scientific culture (Matosin et al., 2014)

Negativity towards negative results: a discussion of the disconnect between scientific worth and scientific culture (Matosin et al., 2014)

Main Takeaways.

- There is pressure on scientists to choose investigative avenues in high-impact knowledge.
- Scientists pursue research that is high in impact, not that is hypothesis-driven.
- Negative results are not given the same value as positive results.
- Scientific principles are under reconsideration and occasions- new evidence refutes old hypotheses.
- Negative findings are seen as an inconvenient truth, ignoring equivocal findings is only human.
- Science is a collaborative discipline and we should report negative findings, so we do not waste time and resources repeating our findings.
- When time is money, research output is judged based on impact and citations, why waste time?
- We face resistance at scientific conferences, when disseminating evidence as we were criticised. Why is a negative finding viewed as a bad thing? A negative is seen as philosophical than practical.
- If negative questions are rephrased as positive questions, does that mean a negative finding is a positive finding?
- Negative findings are seen as taboo and worthy of publication and clinical relevance translated to other related research fields.
- Negative results are not worthy of attention, thus placed in a file drawer and seen as less important.
- We should determine the importance not by the process of "publish or perish" but hypothesis-driven science to fill holes in our knowledge.

Quote

"At the core, it is our duty as scientists to both: (1) publish all data, no matter what the outcome, because a negative finding is still an important finding; and (2) have a hypothesis to explain the finding. If the experiment has been performed to plan, the data has not been manipulated or pulled out of context and there is compiled evidence of a negative result, then it is our duty to provide an explanation as to why we are seeing what we are seeing. Only by truly rethinking the current scientific culture, which clearly favours positive findings, will negative results be esteemed for their entire value. Only then can we work towards an improved scientific paradigm." (p.173)

Abstract

"*What gets us into trouble is not what we don't know, it's what we know for sure that just ain't so.*" – Mark Twain. Science is often romanticised as a flawless system of knowledge

building, where scientists work together to systematically find answers. In reality, this is not always the case. Dissemination of results are straightforward when the findings are positive, but what happens when you obtain results that support the null hypothesis, or do not fit with the current scientific thinking? In this Editorial, we discuss the issues surrounding publication bias and the difficulty in communicating negative results. Negative findings are a valuable component of the scientific literature because they force us to critically evaluate and validate our current thinking, and fundamentally move us towards unabridged science.

APA Style Reference

Matosin, N., Frank, E., Engel, M., Lum, J. S., & Newell, K. A. (2014). Negativity towards negative results: a discussion of the disconnect between scientific worth and scientific culture. *Disease Models & Mechanisms, 7*(2), 171. https://doi.org/10.1242/dmm.015123

You may also be interested in

- ➔ Experimental power comes from powerful theories – the real problem in null hypothesis testing (Ashton, 2013)
- ➔ Negative results are disappearing from most disciplines and countries (Fanelli, 2011)

A farewell to Bonferroni: the problems of low statistical power and publication bias (Nakagawa, 2004)

Main Takeaways.

- The statistical power of a null hypothesis is the probability to reject the null hypothesis when the alternative is correct.
- Using several effect sizes: Cohen's d and Pearson's r. One can assess the mean difference using Cohen's d or Pearson's r to assess the strength of the relationship.
- There is a greater opportunity to make a false negative.
- Bonferroni correction tries to reduce false positives when multiple tests or comparisons are performed.
- Sequential Bonferroni procedure leads to a reduction in power.
- Standard Bonferroni correction- a statistical power of each t-test drops to as low as 33%
- Sequential Bonferroni is not as severe as Standard Bonferroni correction.
- Reviewers may demand Bonferroni to remove irrelevant variables and reduce the number of false positives but it can still lead to publication bias.
- We should discourage Bonferroni or practice of reviewers demanding Bonferroni. We are focusing on p values, not biological or statistical values. You need effect sizes and p values for the importance of these effects.
- Also, report confidence intervals for effect sizes.

Abstract

Professor Shinichi Nakagawa provides a commentary on low statistical power and the need to discourage Bonferroni corrections. In addition, we should rely on effect sizes and their confidence intervals to determine the value of science findings.

APA Style Reference

Nakagawa, S. (2004). A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behavioral ecology, 15*(6), 1044-1045.
https://doi.org/10.1093/beheco/arh107

You may also be interested in

➔ See

The File-drawer problem revisited: A general weighted method for calculating fail-safe numbers in meta analysis (Rosenberg, 2005)

Main Takeaways.

- There is a file-drawer problem in which studies that observe no significant effects are not published.
- One measure to assess the number of non-significant findings is a fail-safe number. These unpublished studies need to be added to a meta-analysis to reduce significant results to non-significance.
- This approach is the best avenue to approach publication bias but identify whether more complex processes are necessary. We should consider unweighted effect sizes.
- Studies with larger sample size or small variance are given higher weight than small sample sizes or large variance.
- Sample size is equivalent to studies with null effect and mean weight necessary to reduce significance level.
- N is the size of a single study of no effect. The relative size means a single study would need to be weighted.
- Degrees of freedom need to be considered.If N is interpreted as multiple studies of mean weight.
- N must be solved for iteratively and variations in t are small with moderate degree of freedom, only a few iterations are needed for convergence.
- Fail-safe calculation is not a method to identify publication bias or explain the publication that exists; it is a procedure to estimate publication bias if it exists and may be safely ignored.

Abstract

Quantitative literature reviews such as meta-analysis are becoming common in evolutionary biology but may be strongly affected by publication biases. Using fail-safe numbers is a quick way to estimate whether publication bias is likely to be a problem for a specific study. However, previously suggested fail-safe calculations are unweighted and are not based on the framework in which most meta-analyses are performed. A general, weighted fail-safe calculation, grounded in the meta-analysis framework, applicable to both fixed- and random-effects models, is proposed. Recent meta-analyses published in Evolution are used for illustration.

APA Style Reference

Rosenberg, M. S. (2005). The file-drawer problem revisited: a general weighted method for calculating fail-safe numbers in meta-analysis. *Evolution, 59*(2), 464-468.https://doi.org/10.1111/j.0014-3820.2005.tb01004.x

You may also be interested in

➔ The "File Drawer Problem" and Tolerance for Null Results (Rosenthal, 1979)

The "File Drawer Problem" and Tolerance for Null Results (Rosenthal, 1979)

Main Takeaways.

- The file drawer problem is that journals have 5% of articles with false positives, while file drawers have 95% non-significant results.
- We need to calculate the number of studies with null findings before the overall false positives are made.
- Another conservative alternative is that when the exact p levels are not present is to set Z = .00 for any non-significant and Z = 1.645 for $p < .05$.
- Small number of studies that are not significant, even when combined can be distorting and misleading with a few studies making a significant effect, non-significant.
- Currently, there are no firm guidelines that can constituent an unlikely number of unretrieved or unpublished studies.It could be 100 or 500, while for others even 10 or 20 seem unlikely.

Quote

" more and more reviewers of research literature are estimating average effect sizes and combined ps of the studies they summarize. It would be very helpful to readers if for each combined p they presented, reviewers also gave the tolerance for future null results associated with their overall significance level." (p.640)

Abstract

For any given research area, one cannot tell how many studies have been conducted but never reported. The extreme view of the "file drawer problem" is that journals are filled with the 5% of the studies that show Type I errors, while the file drawers are filled with the 95% of the studies that show nonsignificant results. Quantitative procedures for computing the tolerance for filed and future null results are reported and illustrated, and the implications are discussed.

APA Style Reference

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological bulletin, 86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638.

You may also be interested in

➔ The File-drawer problem revisited: A general weighted method for calculating fail-safe numbers in meta analysis (Rosenberg, 2005)

At random: Sense and Nonsense (McNemar, 1960)

Main Takeaways.

- Psychology is seen as a quintessential behavioural science.
- 1% of the American Psychological Association are statisticians and 3% of abstracts deal with statistics.
- Survival of psychology can either be due to little discoveries that are scientifically and statistically significant and to the number of little men who are significant to insignificant others.
- There are too many test builders that think to repel statisticians to neglect statistical and psychometric theory of test construction. They produce difficult and voluminous writings and may not be suitable to promote the use of the statistical methods.
- Chi square test was involved in psychology until the 1930s but it was misused, leading to mistaken significance level.
- ANOVAs became more popular in the late 1930s. During this time, more complex designs were included and seen as a better approach to psychology.
- Significance tests are necessary but not sufficient conditions for the development of a science.Co-variance method is a real blessing to social, child, clinical and educational psychology.
- We have to use an interval scale for means, standard deviation and correlation coefficients such as Pearson's r with log-transformed scores.
- Spearman's rho is between two sets of rank.
- A simple design with a simple statistical treatment all lead to a simple conclusion. What could be simpler? Why not a simple approach?
- Mathematical statisticians can no longer keep up with each other, so what is the hope for psychology?
- Statistics teaching will incomprehensible to most psychology students, they need a sound understanding to have an intelligent and critical reading of the literature.
- Journal editors may reject manuscripts with negative findings.

Quote

" more and more reviewers of research literature are estimating average effect sizes and combined ps of the studies they summarize. It would be very helpful to readers if for each combined p they presented, reviewers also gave the tolerance for future null results associated with their overall significance level." (p.640)

Abstract

For any given research area, one cannot tell how many studies have been conducted but never reported. The extreme view of the "file drawer problem" is that journals are filled with the 5% of the studies that show Type I errors, while the file drawers are filled with

the 95% of the studies that show nonsignificant results. Quantitative procedures for computing the tolerance for filed and future null results are reported and illustrated, and the implications are discussed.

APA Style Reference

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological bulletin, 86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638.

You may also be interested in

- ➔ The File-drawer problem revisited: A general weighted method for calculating fail-safe numbers in meta analysis (Rosenberg, 2005)