Meta-Psychology, 2025, vol 9, MP.2024.4326
https://doi.org/10.15626/MP.2024.4326
Article type: Commentary
Published under the CC-BY4.0 license

Open data: Yes
Open materials: Yes
Open and reproducible analysis: Not Applicable
Open reviews and editorial process: Yes
Preregistration: No

Edited by: Felix D. Schönbrodt
Reviewed by: Bert N. Bakker, Edith Beerdsen
Analysis reproduced by: Not Applicable
All supplementary files can be accessed at OSF:
https://doi.org/10.17605/OSF.IO/5BZEH

The value of replications goes beyond replicability and is associated with the value of the research it replicates: Commentary on Isager et al. (2025)

Gilad Feldman[1]
[1]Department of Psychology, University of Hong Kong

Replications are essential for rigorous credible science yet are still grossly undervalued and very rare. The value of replications is directly tied to the value of the research they aim to replicate, and replications offer many benefits that go far beyond the mere testing of replicability, such as including verifications and error detection, promoting long-term reproducibility of all research outputs, clarifying theory, refining measurement, and testing generalizability. We need far more independent pre-registered well-powered direct replications to strengthen the credibility of scientific findings. Isager et al. (2025)'s aim to define a formula for the value of replications based on over-simplified metrics of citation count and sample size is misaligned, already misunderstood, and may backfire by hindering the pursuit and publication of replications.

Keywords: replications, value, commentary, replicability

Replications are very rare: We just do not do replications

Replications are still often misunderstood and undervalued. Despite ongoing discussions regarding the importance of replications, there has been little to no progress in making replications mainstream. One of the strongest indicators is the replication-nonreplication publication ratio, with recent estimates putting the ratio of replications of publications at around 0.2% in psychology (Clarke et al., 2023) and 0.54% in education (Cook et al., 2024), unfortunately closely resembling the rates reported a decade ago with 0.1% in psychology (Makel et al., 2012) and 0.13% in education (Makel and Plucker, 2014). Accumulating meta-scientific evidence indicates replications are still an anecdote, at best, with most research not subject to independent direct replications.

To ensure scientific credibility we should be doing and publishing a lot more replications (Nosek et al., 2022; Vazire, 2018; Zwaan et al., 2018). There are many systemic challenges hindering replications: there is a strong bias for novelty in publishing, hiring, promotion, and funding, there are sensitivities around conducting replications where replicators are perceived as having some kind of an agenda, and strong prestige, hindsight, and outcome biases where - for example - replicators are criticized as incompetent when replications fail, or replication work is regarded as having no value and unsurprising when replications succeed (Chandrashekar and Feldman, 2024), with fierce debates even regarding the very definition of replication success and failure. Very few scholars are choosing to conduct replications, fewer are submitting replications that they did do, and of replications submitted even fewer are eventually published. An example from the field of social psychology, Journal of Experimental Social Psychology reported that in 2019 1.7% (12 of 670) of *submissions* to the journal were replications, and 1.5% (10 of 700) in 2020 (Giner-Sorolla, 2021a, 2021b). Quite a few of those 2019-2020 submissions were from our replication team (CORE, 2024), and of which only about a third made it to publication in that journal. Even lower rates were reported for other journals like Journal of Personality and Social Psychology (Corker, 2020).

At the moment, given how scarce replications are, almost any replication of published articles that were not yet replicated is of value. We need to conduct, submit, and publish more replications. It would be best to prioritize which replications to run to maximize utility and impact, yet this should not be a publication selection tool. It would take years if not decades and a mindset shift to come anywhere close to a novel-replication publication ratio that makes sense for credible science.

Replication value is tied to research value

Isager et al. (2025) and Isager et al. (2023) tackle the value of replications asking the following - with limited resources, how do we choose which *replications* to prioritize? An important question, yet it is only part of

a much broader question - with limited resources, what criteria should we apply to choose what *research* to prioritize?

We must shift our priorities from an emphasis on novelty to a balance between novel and replication work, and so a value function for replication alone seems misaligned.

Some might ask: Why not start from first establishing a criteria for replications as a case study and then move to a broader criteria for other research?

First, the credibility and value of published research depends on its replications. As Isager et al. (2025) noted the value of replications is directly tied to the value of the research being replicated, yet the opposite also holds true, the value of published research depends on its reproducibility and replicability, and so it requires replications. One of the very few articles that suggested a benchmark was Makel et al. (2012) who noted that "if a publication is cited 100 times, we think it strange if no attempt at replication has been conducted and published" (p.541). Suggesting a threshold is bold and appreciated, yet the reference to 100 seems arbitrary. If research reporting evidence supporting a scientific claim has been conducted, peer reviewed, and endorsed or accepted for publication in an outlet the scientific community considers credible, then it would be strange if there was no attempt at an independent direct replication of this work. Especially so for older studies with outdated standards of rigor and transparency. Currently, for an article to be accepted for publication in a reputable journal, the editor as a representative of the journal and its scientific community saw sufficient value in it for it to be publicly disseminated with their endorsement. The value of research is positively associated with the value of its replication. If novel research was considered of sufficient value to be published, then its replications should have sufficient value to be published ("Pottery Barn Rule"; summarized in Edlund et al., 2022). The scientific process warrants that one should follow the other, for any finding to be considered solid and credible.

Second, establishing a criteria that only applies to replications may unintentionally limit the publication of replications compared to novel research where no clear criteria exists or is proposed. Therefore, it is important to make clear that "Deciding what to replicate" is about replication *resource optimization* to support replicators in suggesting what replications to prioritize given scarce resources, and not as a criteria for evaluation, comparison, and/or rejection of replication work that has already been conducted. This is not a hypothetical, as our team already had reviewers referring to this replication value criteria in their assessment of the value of our replications submitted for publication.

Replications go beyond replicability

What is it that direct replications do that is of "central value of empirical science"? Isager et al., 2025 focused on value of and uncertainty regarding the original findings, which they linked to citation count and sample size, respectively. This is a good starting point, yet does not fully capture much of the value that replications offer in addressing the many urgent challenges identified by the science reform movement and recent meta-scientific evidence.

Direct replications assess replicability, testing whether others are able to observe similar findings with a different sample at a different time using similar processes, methods, and materials. They contribute additional evidence that helps update our confidence regarding the original's results, and our expectations regarding how solid, robust, and generalizable the effects are. Replications help identify uncertainties and gaps, and then - in the aggregate, help make progress towards a better estimation of the effect size, and identify possible moderators.

Replications include verifications. In closely replicating a study, replicators do very detailed analyses of the original study, like in reconstructing the methods and analyses and conducting a power analysis. They help identify inadequate methods, errors, signs of questionable research practices, and even fraud. This captures a different kind of uncertainty that can only be found with an in-depth analysis of an article. This can be done with external red-teams and assessments, yet is already inherent in replications. For example, some of our replications identified errors in published articles that have led to publisher and original authors issuing an expression of concern regarding the article (e.g., Imada et al., 2022).

Replications assess the reproducibility of the target article, aid in patching reproducibility gaps, and promote long-term reproducibility of process, materials, methods, data, and code, often adhering to much stricter open-science standards. This allows others to more easily build on this research and reduce potential for errors and misunderstandings. For example, several teams from around the world have built on our replication materials to conduct their own replications in their own country and language (e.g., Borborema et al., 2023 using our Chen et al., 2020).

Replications contribute to generalizability. Our team conducts replications of many seminal findings going as far back as the 1970s with thousands of citations that do not have any published independent direct replications. Even if researchers indeed found solid evidence

for a phenomenon using a specific sample from a certain country decades ago, that says little regarding whether these findings hold true years later, with other samples, in other parts of the world, and in varying contexts. Repeating the same process, methods, and stimuli in a different context and continuously over time is of value. Uncertainty is related to many factors beyond the sample size of a study from a specific time and context.

Replications also contribute to theory. They help assess the alignment between theory, hypotheses, and the provided evidence. Replications often require making the original's implicit theory explicit, clearly articulating testable falsifiable hypotheses tied to the empirical tests, with a close evaluation and possible rethinking of the theory and hypotheses. For example, one of our common replication adjustments was when original theories were implicitly arguing for a null hypothesis and used no signal using null hypothesis significance testing with small samples as evidence in support of the null (e.g., Imada et al., 2022; Ip and Feldman, 2025).

Similarly, replications contribute to measurement. Replications allow reassessing, validating, and adjusting the original's measurement, examining appropriateness and alignment with theory, and retesting using better validated measures. For example, in our replications of studies with scale predictors we often run the original's alongside different types of measurement or more comprehensive scales in random order and then test them all with comparisons against each other (e.g., Koppel et al., 2023; Zhu and Feldman, 2023).

Finally, replications may help reignite interest in important research that was forgotten or pushed aside because of gatekeeping. Traditional imbalanced power structures in academia at times led to emphasizing a specific type or view of research and target topic, impacting academic discourse through biased publications and citations. Citation counts reflect these old imbalanced power structures. Replications can help bring attention to highly valuable neglected research and address these structural weaknesses (e.g., Aiyer et al., 2024). We need a broad clear assessment of value for all types of research that goes beyond citation counts (e.g., Dougherty and Horne, 2022), which will also help inform the value of replications.

Conclusion

Replications offer far more value than just testing replicability. Replications check for mistakes and issues, validate and ensure long term reproducibility, update methods to best-practices, and clarify, test, and improve theory, measurement, and generalizability. The value of published research depends on replications, and the value of replications is tied to the value of the research they revisit. Given that replications are so rare and most research is never replicated, we need to catch up, so it would be helpful to have a tool that helps replicators decide which articles to *prioritize*, and the debate is appreciated. Yet, value assessment should not start with replications and based on biased metrics like citations, but rather with an assessment of the value of research overall, the research that replications help make more credible.

Author Contact

Gilad Feldman; Department of Psychology, University of Hong Kong, Hong Kong SAR; gfeldman@hku.hk / giladfel@gmail.com; ORCID: 0000-0003-2812-6599

Conflict of Interest and Funding

I coordinate the Collaborative Open-science and meta REsearch team (CORE, 2024) in which we completed over 120 replications of seminal studies in social psychology and judgment and decision making. Despite much advice and many warnings against pursuing this path, I have been very invested in conducting replications and in promoting the practice of replications. I acknowledge that this likely greatly influenced my views on the importance of replications.

I received no financial support for the research and/or authorship of this article.

Open Science Practices

This article is conceptual and as such is not eligible for Open science badges. The entire editorial process, including the open reviews, is published in the online supplement.

References

Aiyer, S., Chan, W. Y., & Feldman, G. (2024). Outcome Bias in Evaluations of Ethical Decisions: Replication and Extensions of Gino, Moore, and Bazerman (2009). *Collabra: Psychology*, *10*(1). https://doi.org/10.1525/collabra.126266

Borborema, R. S., de Moraes Ferreira, M., da Silva, A. L. M., Bastos, R. V. S., Fatori, D., Feldman, G., Seda, L., & Batistuzzo, M. C. (2023). Inaction Inertia Effect: Foregoing Opportunities as a Consequence of an Initial Failure to Act - a Replication-Extension Study in the Brazilian Population. *OSF*. https://doi.org/10.31234/osf.io/4u98x

Chandrashekar, S., & Feldman, G. (2024). On the process and value of direct close replications: Reply to shafir and cheek (2023) commentary on chandrashekar et al. https://doi.org/10.17605/OSF.IO/RVPZF

Chen, J., Hui, L. S., Yu, T., Feldman, G., Zeng, S., Ching, T. L., Ng, C. H., Wu, K. W., Yuen, C. M., Lau, T. K., Cheng, B. L., & Ng, K. W. (2020). Foregone Opportunities and Choosing Not to Act: Replications of Inaction Inertia Effect. *Social Psychological and Personality Science*, *12*(3), 333–345. https://doi.org/10.1177/1948550619900570

Clarke, B., Lee, P., Schiavone, S., Rhemtulla, M., & Vazire, S. (2023). The prevalence of direct replication articles in top-ranking psychology journals. In *Advances in methods and practices in psychological science*. https://doi.org/10.1037/amp0001385

Cook, B., Therrien, W., Waterfield, D., McClain, S., Fleming, J., Robinson, H., & Boyle, J. (2024). Same as it ever was. *An Updated Review of Replication Studies in Special Education Journals. Remedial and Special Education*, *07419325241248766*. https://doi.org/10.1177/07419325241248766

CORE. (2024). Collaborative open-science and meta research [Retrieved from]. http://osf.io/5z4a8

Corker, K. (2020). Post of twitter/x [Retrieved from]. https://x.com/katiecorker/status/1222592990411214851

Dougherty, M., & Horne, Z. (2022). Citation counts and journal impact factors do not capture some indicators of research quality in the behavioural and brain sciences. *Royal Society Open Science*, *9*(8), 220334. https://doi.org/10.1098/rsos.220334

Edlund, J., Cuccolo, K., Irgens, M., Wagge, J., & Zlokovich, M. (2022). Saving science through replication studies. *Perspectives on Psychological Science*, *17*(1), 216–225. https://doi.org/10.1177/1745691620984385

Giner-Sorolla, R. (2021a). Post on twitter/x [Retrieved from]. https://x.com/RogertheGS/status/1222847696140218369

Giner-Sorolla, R. (2021b). Time for change, part i reflections on the makeup and historical methodological practices in social [Retrieved from]. *Society for Personality and Social Psychology talk (timestamp*, *20*(45). https://www.youtube.com/watch?v=0KIVLn4C5U8&t=1245s

Imada, H., Chan, W. F., Ng, Y. K., Man, L. H., Wong, M. S., Cheng, B. L., & Feldman, G. (2022). Rewarding More Is Better for Soliciting Help, Yet More So for Cash Than for Goods: Revisiting and Reframing the Tale of Two Markets With Replications and Extensions of Heyman and Ariely (2004). *Collabra: Psychology*, *8*(1). https://doi.org/10.1525/collabra.32572

Ip, H. C., & Feldman, G. (2025). The complex misestimation of others' emotions: Underestimation of emotional prevalence versus overestimation of emotional intensity and their associations with well-being. *Personality and Social Psychology Bulletin*. https://doi.org/10.17605/OSF.IO/BWMTR

Isager, P., van 't Veer, A., & Lakens, D. (2025). Replication value as a function of citation impact and sample size. *Meta-Psychology*, *9*. https://doi.org/10.15626/MP.2022.3300

Isager, P., Aert, R., Bahník, Š., Brandt, M., DeSoto, K., Giner-Sorolla, R., & Lakens, D. (2023). Deciding what to replicate: A decision model for replication study selection under resource and knowledge constraints. *Psychological Methods*, *28*(2), 438. https://doi.org/10.1037/met0000438

Koppel, L., Andersson, D., Tinghög, G., Västfjäll, D., & Feldman, G. (2023). We are all less risky and more skillful than our fellow drivers: Successful replication and extension of Svenson (1981). *Meta Psychology*, *7*. https://doi.org/10.15626/MP.2021.2932

Makel, M., & Plucker, J. (2014). Facts are more important than novelty: Replication in the education sciences. *Educational Researcher*, *43*(6), 304–316. https://doi.org/10.3102/0013189X14545513

Makel, M., Plucker, J., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives on Psychological Science*, *7*(6), 537–542. https://doi.org/10.1177/1745691612460688

Nosek, B., Hardwicke, T., Moshontz, H., Allard, A., Corker, K., Dreber, A., & Vazire, S. (2022). Replicability, robustness, and reproducibility in psychological science. *Annual Review of Psychology*, *73*, 719–748. https://doi.org/10.1146/annurev-psych-020821-114157

Vazire, S. (2018). Implications of the credibility revolution for productivity, creativity, and progress. *Perspectives on Psychological Science*, *13*(4), 411–417. https://doi.org/10.1177/1745691617751884

Zhu, M., & Feldman, G. (2023). Revisiting the Links Between Numeracy and Decision Making: Repli-

cation Registered Report of Peters et al. (2006) With an Extension Examining Confidence. *Collabra: Psychology*, *9*(1). https://doi.org/10.1525/collabra.77608

Zwaan, R., Etz, A., Lucas, R., & Donnellan, M. (2018). Making replication mainstream. *Behavioral and Brain Sciences*, *41*, 120. https://doi.org/10.1017/S0140525X17001972