Meta-Psychology, 2025, vol 9, MP.2024.4216
https://doi.org/10.15626/MP.2024.4216
Article type: Commentary

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

Edited by: Felix D. Schönbrodt
Reviewed by: Rene Bekkers, Gilad Feldman
Analysis reproduced by: Not Applicable
All supplementary files can be accessed at OSF:
https://doi.org/10.17605/OSF.IO/ED65R

Use and Misuse of a Fast Approximation: Not a Criticism, but a Caution

Carlisle Rainey[1]
[1]Florida State University

P. Isager et al. (2025) propose a fast approximation of replication value, RV_{Cn}, that relies on citation count and sample size. This approximation is simple, transparent, and easy to implement across many studies. It can potentially help metascientists evaluate large collections of studies. However, RV_{Cn} is not a precise statement of fact; it should not substitute for detailed substantive and methodological arguments. I make two counterclaims: (1) studies with few citations might be worth replicating and (2) studies with large samples might be worth replicating. While RV_{Cn} can helpfully supplement researchers' judgments, it should not *substitute for* researchers' judgments.

Keywords: citation count, replication, replication value, RV_{Cn}, sample size, study selection

P. Isager et al. (2025) propose that researchers use the sample size and citation count to quickly estimate the value of replicating a study. In this commentary, I discuss their fast approximation of replication value, which they denote RV_{Cn}. While RV_{Cn} helps quickly evaluate a large collection of studies, it can easily become a *harmful crutch*; it might stifle detailed substantive and methodological arguments.

The Fast Approximation as a Helpful Tool

P. Isager et al. (2025) explain the benefits of a fast approximation clearly (see also P. M. Isager et al. (2023, p. 440)). While I do not reiterate their explanations here, I endorse their fast approximation as a potentially helpful tool. In brief, P. Isager et al. (2025) offer a useful rule of thumb to quickly approximate the replication value for many studies. They develop their approximation carefully and demonstrate its potential (see their Table 1).

The Fast Approximation as a Harmful Crutch

But the fast approximation is ripe for misuse. Although RV_{Cn} is a precise equation, P. Isager et al. (2025) clearly explain that RV_{Cn} is not a precise statement of fact. And while the authors recommend in-depth evaluations when these are available or practical, it is tempting to push the hard work of evaluating potential (or even completed) replications entirely onto their simple equation. After all, why struggle carefully through the messy, complex details of concepts, theory, and evidence when a precise equation produces a clear answer so quickly? Their fast approximation sings a siren song of "effortless objectivity"—an easy, exact equation that researchers and referees might use uncritically.

As with most useful guidelines, researchers can make a serious error if they treat RV_{Cn} as literal or precise. And P. Isager et al. (2025) agree, writing that "[i]n-depth evaluation allows for quality control and nuance during study selection. . . a certain amount of quality control will likely always be beneficial."

We could imagine that researchers carefully craft a detailed argument that a replication would bolster the evidence for an important claim, only for an overworked referee to dismiss the proposal because the original study has a large sample or relatively few citations.[1] What a disheartening outcome! Perhaps worse, the researchers themselves might defer entirely to these cues rather than inject their deep expertise into the evaluation.[2]

We must remember that P. Isager et al. (2025) write from the perspective of metascientists evaluating dozens or hundreds of potential replications across many subjects. While this perspective is useful and important, many researchers study specific substantive

[1]And we need not imagine. Feldman (2025, p. 3) notes: "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."

[2]Feldman (2025, p. 2) makes a strong case that "we should be doing and publishing a lot more replications" and that RV_{Cn} "may unintentionally limit the publication of replications compared to novel research." Indeed, Feldman (2025, p. 3) writes that RV_{Cn} should *not* be used "as a criteria for evaluation, comparison, and/or rejection of replication work that has already been conducted."

problems slowly and carefully. It would be a grave mistake for these substantive experts to treat the fast approximation RV_{Cn} as a precise statement of fact. For these substantive experts, careful and detailed assessments are available or within reach.

The Value of a Simple Equation

RV_{Cn} is a simple measure. This simplicity allows metascientists to implement the measure across dozens or hundreds of studies. Citations are readily available, and sample size is usually reported.

However, the simplicity has another subtle and important benefit. RV_{Cn} (and its two components) clearly correspond to the concept of replication value (and its two components).

$$RV_{Cn} = \text{value} \times \text{uncertainty} = \frac{w(C_S)}{Y+1} \times \frac{1}{\sqrt{n}}$$

The close, obvious correspondence between the concept and the measure (and their components) allows researchers to clearly see how RV_{Cn} works. By implication, it is also clear how RV_{Cn} *can fail*.[3] For any particular study, topic, or field, researchers can immediately ask: under what conditions is the number of citations a poor measure of value? And under what conditions is sample size a poor measure of uncertainty?

We could instead use a black box procedure. For example, a web application could take the full text of an article, implement a complex algorithm, and return an estimate of the replication value. Perhaps this algorithm would be more accurate. But it would not be clear to users how the input relates to the output. It would not be clear *how the procedure can fail*.

As Box and Draper (2007, p. 414) note: "the approximate nature of the model must always be borne in mind." With RV_{Cn}, it is (relatively) *easy* to keep the approximate nature of the model at the top of mind.

Two Counterclaims

The fast approximation RV_{Cn} can helpfully *supplement* detailed substantive and methodological arguments, but RV_{Cn} should not *substitute for* such detailed arguments. To make my point of caution clearly, I make two counterclaims.

Counterclaim 1. *Studies that have not been heavily cited might be worth replicating.*

Citations might suggest that *others* have found a result important to build upon, but this is a crude metric. Indeed, factors unrelated to the importance of the claim can affect citations, such as the author's gender (Dion et al., 2018).

If we view social science as a cumulative exercise, then we view new theories and studies as *building* upon existing studies. This leaves the question: what are researchers building upon? As Chambers (2019, p. 49) writes: "To rely solely on extrapolation at the expense of direct replication is to build a house on sand." If we want to build on a firm foundation of existing research, it follows that researchers ought to prioritize replicating studies that they hope to build upon.

Pathologically, high citations might indicate that many closely related, subsequent studies have accumulated evidence closely related to the original experiment, even if no direct replication exists. On the other hand, low citations might indicate a relatively original experiment that did not immediately and obviously connect to existing research streams. In this case, low citations might indicate sparse evidence and a *greater* need for replication. Again, careful and detailed assessment is valuable.

Further, claims can be important for social, clinical, scientific, or other reasons, and a claim's importance might not have been previously appreciated. Although it requires hard work, the detailed arguments and in-depth evaluations are essential.

Counterclaim 2. *Studies with large samples might be worth replicating.*

Large sample sizes suggest compelling evidence, but this is a crude metric. Many factors affect uncertainty besides sample size. For example, list experiments require very large sample sizes (Blair et al., 2020). Pre-post designs require smaller sample sizes (Clifford et al., 2021). Interaction effects require a much larger sample size to estimate precisely (Rainey, 2025, pp. 25–27).

And moving beyond direct or close replication, there are many reasons why we might view existing evidence for a claim as insufficient, including the design, measurement, sample, and statistical methods. Again, detailed assessment by substantive experts is critical.

Summary

P. Isager et al. (2025) propose RV_{Cn} to estimate replication value at scale. This is a simple, transparent approximation. The simplicity is a great virtue; it is clear how RV_{Cn} works and thus how it can help and how it can fail. We should keep this potential for failure top-of-mind. We should not allow an easy, objective equation

[3] Putting it bluntly, Bekkers (2025) writes: "citation counts and sample size are bad indicators of replication value." Bakker et al. (2025) discuss specific ways that RV_{Cn} can fail in greater detail.

to tempt us away from detailed, substantive arguments about whether (1) the claim, if correct, would be a valuable foundation to build upon, (2) the body of evidence in support of the claim is inadequate, and (3) a replication study would bolster the evidence for the claim (P. M. Isager et al., 2023; Pittelkow et al., 2023).

Zwaan et al. (2018, p. 9) write: "It goes without saying that *scientific judgment* should be used to assess the validity and importance of a study before deciding whether it is worth replicating." While the fast approximation can helpfully *supplement* researchers' judgments, the fast approximation should not *substitute for* researchers' judgments. With or without RV_{Cn}, researchers must not shy away from detailed substantive and methodological arguments about theories, evidence, and implications.

Author Contact

Carlisle Rainey; Department of Political Science, Florida State University, Tallahassee, FL, USA; crainey@fsu.edu; ORCID: 0000-0002-8728-4696

Conflict of Interest and Funding

There are no conflicts of interest to report. There was no specific funding for this research.

Author Contributions

This is a solo-authored paper; the author completed all aspects of the work.

Open Science Practices

This article is purely 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

Bakker, B. N., Bomm, L., & Peterson, D. J. (2025). Commentary on isager et al. (2025): Reflections on the replication value (rv) and a proposal for revision. *Meta-Psychology*, *9*. https://doi.org/10.15626/MP.2024.4324

Bekkers, R. (2025). Replication value increases with transparency, test severity, and societal impact. *Meta-Psychology*, *9*. https://doi.org/10.15626/MP.2024.4194

Blair, G., Coppock, A., & Moor, M. (2020). When to worry about sensitivity bias: A social reference theory and evidence from 30 years of list experiments. *American Political Science Review*, *114*(4), 1297–1315. https://doi.org/10.1017/S0003055420000374

Box, G. E. P., & Draper, N. R. (2007). *Response surfaces, mixtures, and ridge analyses* (2nd). Wiley. https://doi.org/10.1002/0470072768

Chambers, C. (2019). *The seven deadly sins of psychology* (2nd). Princeton University Press.

Clifford, S., Sheagley, G., & Piston, S. (2021). Increasing precision without altering treatment effects: Repeated measures designs in survey experiments. *American Political Science Review*, *115*(3), 1048–1065. https://doi.org/10.1017/S0003055421000241

Dion, M. L., Sumner, J. L., & Mitchell, S. M. (2018). Gendered citation patterns across political science and social science methodology fields. *Political Analysis*, *26*(3), 312–327. https://doi.org/10.1017/pan.2018.12

Feldman, G. (2025). The value of replications goes beyond replicability and is associated with the value of the research it replicates: Commentary on Isager et al. (2025). *Meta-Psychology*, *9*. https://doi.org/10.15626/MP.2024.4326

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. M., van Aert, R. C. M., Bahník, Š., Brandt, M. J., DeSoto, K. A., Giner-Sorolla, R., Krueger, J. I., et al. (2023). Deciding what to replicate: A decision model for replication study selection under resource and knowledge constraints. *Psychological Methods*, *28*(2), 438–451. https://doi.org/10.1037/met0000438

Pittelkow, M.-M., Field, S. M., Isager, P. M., Van't Veer, A. E., Anderson, T., Cole, S. N., Dominik, T., Giner-Sorolla, R., Gok, S., Heyman, T., Jekel, M., Luke, T. J., Mitchell, D. B., Peels, R., Pendrous, R., Sarrazin, S., Schauer, J. M., Specker, E., Tran, U. S., ... van Ravenzwaaij, D. (2023). The process of replication target selection in psychology: What to consider? *Royal Society Open Science*, *10*(2). https://doi.org/10.1098/rsos.210586

Rainey, C. (2025). Power rules: Practical advice for computing power (and automating with pilot data) [Preprint]. https://doi.org/10.31219/osf.io/5am9q

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