Making prepublication independent replication mainstream

Warren Tierney

University of Limerick

Martin Schweinsberg

ESMT Berlin

Eric Luis Uhlmann

INSEAD

(Commentary on Rolf A. Zwaan, Alexander Etz, Richard E. Lucas,

and Brent Donnellan, “Making Replication Mainstream”)

ABSTRACT: 56 words

MAIN TEXT: 984 words

REFERENCES: 317 words

ENTIRE TEXT: 1455 words

CONTACT:

Warren Tierney
Kemmy Business School
University of Limerick
Castletroy, Limerick
Ireland
Phone: +35387329150
E-mail: warrentierney@hotmail.com

Martin Schweinsberg
ESMT Berlin
Schlossplatz 1
10178 Berlin, Germany
Phone: +49 30 2123 115 49
E-mail: martin.schweinsberg@esmt.org

Eric Luis Uhlmann
INSEAD, Organisational Behaviour Area
1 Ayer Rajah Avenue
138676 Singapore
Phone: 65 8468 5671
E-mail: eric.luis.uhlmann@gmail.com

Abstract

The widespread replication of research findings in independent laboratories prior to publication is suggested as a complement to traditional replication approaches. The pre-publication independent replication approach further addresses three key concerns from replication skeptics by systematically taking context into account, reducing reputational costs for original authors and replicators, and increasing the theoretical value of failed replications.

The reproducibility of scientific findings, whereby a study is replicated by independent investigators in order to assess the robustness of the research and of its findings, is fundamental to the scientific process (Dunlap, 1926; Popper, 1959). Overall, we strongly agree with the authors of the target article that replication should be made mainstream. Although replication is typically discussed in terms of reproducing previously published work, we further advocate for making mainstream the independent replication of findings prior to publication (see also Schooler, 2014). Pre-Publication Independent Replication (PPIR) is a collaborative, crowdsourced approach to science where original study authors nominate their own findings to be replicated in independent laboratories around the world. This is a complementary approach to existing replication initiatives that focus on published findings, one with different strengths and weaknesses. Importantly, PPIR further addresses three of the key concerns from replication skeptics counterargued so effectively in the target article.

In our first Pre-Publication Independent Replication initiative (Schweinsberg et al., 2016; Tierney et al., 2016), 10 unpublished moral judgement effects from the last author's research pipeline were replicated by 25 independent research groups who collected data from over 11,000 participants. The findings were mixed— while some studies replicated successfully, others did not replicate according to the a priori established criteria. Overall, six findings successfully replicated, one study replicated but with a much smaller effect size than the original (a decline effect; Schooler, 2011), two findings were not supported, and one study was culturally moderated (replicating consistently in the original country but not in five other countries). The culturally moderated effect provides evidence that contextual factors can play an important and unexpected role in replications. In total, 40% of the original findings failed at least one major criterion for reproducibility.

We have expanded the scope of our crowdsourcing approach in a second PPIR initiative, the Pipeline Project 2. This initiative opens pre-publication independent replication to the world, providing original authors the opportunity to nominate their unpublished work for replication in partner laboratories as well as graduate methods classes. We currently have 14 original findings being replicated at over 50 sites around the world (Schweinsberg, et. al., in progress). Original authors opt into the PPIR process and help select replicators they regard as suitable and as having access to relevant subject populations, leading to collaborative rather than adversarial interactions. Notably, original authors are asked to specify beforehand in what cultures and research sites they do and do not expect their effect to emerge. We are further conducting a prediction market (Dreber et al., 2015) to see if members of the scientific community at large can anticipate contextual variability in effects. These aspects of the PPIR process further addresses a key challenge raised by replication skeptics, by systematically taking into account context.

Concerns have also been raised about reputational damage to those involved in replications, both to original authors whose published findings are not reproduced by other research groups, and replicators whose results question established findings (Bohannon, 2014; Kahneman, 2014; Schnall 2014a/b/c). By replicating findings in independent laboratories before (rather than after) the findings are published, PPIRs minimizes reputational costs to both original authors and replicators since 1) no one's reputation depends on the outcome, and 2) original authors voluntarily opt into the PPIR process and help select their replicators.

Another common argument is that failed replications are uninterpretable and low in theoretical value (Schnall 2014a/b/c). Although in our view replications are always

informative and valuable (Dreber et al., 2015), it is at the same time true that there are other plausible explanations for null findings other than the original effect being false (Open Science Collaboration, 2015). We suggest that the theoretical value of PPIR in terms of identifying false positives is even higher than for traditional replications, since most alternative explanations for null effects are ruled out. In particular, defenders of the original finding have little basis to attribute an unsuccessful replication to a lack of replicator expertise or use of irrelevant subject populations, since the original authors helped select what they regarded as qualified replicators and specified a priori which participant populations they expected to exhibit the effect. However, informational value is correspondingly lower for successful PPIRs, relative to traditional replications, since the original authors participate in selecting their own replicators who may be biased in favor of the hypothesis. Indeed, research demonstrates that the theories investigators endorse strongly predict the effect sizes they obtain (Berman & Reich, 2010).

The biggest challenge to making pre-publication independent replication mainstream is the lack of professional incentives, especially for replicators. One potential solution is to build PPIRs into the education of graduate students (Everett & Earp, 2015) as part of crowdsourced projects on which they and the instructors of their methods courses are co-authors. These student PPIRs can examine findings that the original authors identify as straightforward for a junior researcher to conduct. To facilitate the integration of pre-publication independent replication into graduate methods courses, as part of the Pipeline Project 2 we have developed an open source curriculum on Crowdsourcing Science including instructions for student PPIR projects (https://osf.io/hj9zr/). Researchers of any level of experience who wish to initiate projects can use the Study Swap website (https://osf.io/view/StudySwap/), a new forum where interested parties can engage with the PPIR process, both as original authors

looking for labs to replicate their findings or as independent investigators looking to replicate findings. Networks of partner laboratories such as the Psychological Science Accelerator (Chartier, 2017) might also be leveraged to conduct replications of unpublished, rather than published findings.

In sum, conducting independent replications earlier in the research process – before findings are even submitted for publication – can further address what the target article identifies as three of the key concerns raised by skeptics of replication. The pre-publication independent replication approach minimizes reputational costs to original authors and replicators, systematically takes into account context, and maximizes the informational value of failed replications.

References

Berman, J. S., & Reich, C. M. (2010). Investigator allegiance and the evaluation of psychotherapy outcome research. *European Journal of Psychotherapy and Counselling, 12*, 11-21.

Bohannon, J. (2014). Replication effort provokes praise—and 'bullying' charges. *Science, 344*, 788-789.

Chartier, C.R. (2017). The Psychological Science Accelerator: A distributed laboratory network. Retrieved at: https://christopherchartier.com/2017/09/21/the-psychological-science-accelerator-a-distributed-laboratory-network

Dreber, A., Pfeiffer, T., Almenberg, J., Isaksson, S., Wilson, B., Chen, Y. Nosek B.A., & Johannesson, M. (2015). Using prediction markets to estimate the reproducibility of scientific research. *Proceedings of the National Academy of Sciences, 112*, 15343-15347.

Dunlap, K. (1926). The experimental methods of psychology. In: C. Murchison (Ed.) *Psychologies of 1925* (pp. 331-353). Worchester: Clark University Press.

Everett, J. A., & Earp, B. D. (2015). A tragedy of the (academic) commons: interpreting the replication crisis in psychology as a social dilemma for early-career researchers. *Frontiers in Psychology, 6*, 1–4.

Kahneman, D. (2014). A new etiquette for replication. *Social Psychology, 45(4)*, 310-311.

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science, 349(6251)*. DOI: 10.1126/science.aac4716

Popper, K. R. (1959). *Logic of scientific discovery*. New York: Basic Books.

Schnall, S. (2014a). An experience with a registered replication project. Available at: http://www.psychol.cam.ac.uk/cece/blog#anchor-1

Schnall, S. (2014b). Further thoughts on replications, ceiling effects and bullying. Available

at: http://www.psychol.cam.ac.uk/cece/blog

Schnall, S. (2014c). Social media and the crowd-sourcing of social psychology. Available at: http://www.psychol.cam.ac.uk/cece/blog

Schooler, J. (2011). Unpublished results hide the decline effect. *Nature, 470*, 437.

Schooler, J. (2014). Metascience could rescue the 'replication crisis'. *Nature, 515*, 9.

Schweinsberg, M., et al. (2016). The pipeline project: Pre-publication independent replications of a single laboratory's research pipeline. *Journal of Experimental Social Psychology. 66*, 55–67.

Schweinsberg, M., et al. (in progress). The pipeline project 2: Opening up pre-publication independent replication to the world.

Tierney, W. et al. (2016). Data from a pre-publication independent replication initiative examining ten moral judgement effects. *Nature Scientific Data 3*:160082 doi: 10.1038/sdata.2016.82.