

Restoring confidence in psychological science findings: A call for direct replication studies

Denis Cousineau ,^a

^a School of Psychology, Université d'Ottawa

Abstract ■ The *Quantitative Methods for Psychology* journal begins the publication of replication studies. The replication of a target experiment must be direct (as truthful as possible to the target experiment), performed by independent researchers not involved in the target experiment and have at least the same statistical power. The manuscript will be evaluated on these criteria only, not on the result(s) of the replication. A typical manuscript is very short, as the theoretical context just put the target experiment in context, the methodology only highlight the modifications or methodological aspects that needed clarifications, the results list the findings, mostly in terms of effect sizes (raw and standardized), and the discussion judges how well the original findings replicated. No extra narrative is needed.

Keywords ■ Replication, direct replications

 denis.cousineau@uottawa.ca

Introduction

The Bem incident (Bem, 2011) attracted much attention (e.g., Yong, 2014) not so much because an editor presumably published a type-I error (this is bound to happen roughly once in twenty experiments!) but because subsequent replications invalidating the target article were rejected at triage, not being sent to reviewers. This practice opens up the possibility that psychology as a scientific body of knowledge may contain flawed entries that are not purged away with time. This led Ionannidis to title a recent article "Why science is not necessarily self-correcting" (2012). This problem in turn can be magnified by meta-analyses, the famous file-drawer problem (see a recent discussion in Francis, 2012). Some psychology scientists went even further and announced "A crisis of confidence?" (Pashler and Wagenmakers, 2014).

The solution is however very simple. Publish replications of an experiment. The gold standard of science after all is reproducibility (as reminded recently by Jasny, Chin, Chong & Vignieri, 2011). Makel, Plucker and Hegarty (2012) analyzed a large sample of articles published in psychological sciences and found that approximately 1% of the published papers are replications and of this number, only one over seven are direct replications (i. e., as close as possible to the target experiment). This is very few, not mentioning the fact that only half of these are from independent

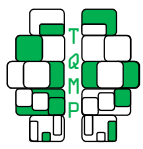
authors (not related to the authors of the target experiment).¹ Considering that about 33% of the independent replications are unsuccessful, there are reasons for worries...

To avoid this problem, there should be systematic replications and journals should have obligations to publish replications (and there should be an obligation to cite these replications studies, at least for some years, Koole and Lakens, 2014). Some argue that the efforts for our discipline would be minimal (Grahe, Reifman, Hermann, Walker, Oleson, Nario-Redmond and Weibe, 2012).

"I certainly agree that it's desirable that replications are published," the editor of Bem's article told Aldous (2011). "The question is where. There are hundreds of journals in psychology". This quote suggests that the problem of replication studies is not as much a question of "Should we publish them?" but rather "Who should put the efforts to publish them?"

The *Quantitative Methods for Psychology* will publish replication studies. The editor's decision to publish will not be based on rejections of null hypotheses, or on how well the target study replicates.

¹ Approximately 170,000 articles were published in 2010. These figures then suggests that there was about 120 independent direct replications.



Instead, the following criteria will be used: (1) The replication study must be a direct replication (as opposed to a conceptual replication; see Make, Plucker, & Hegarty, 2014, or Schmidt, 2009); (2) it must also have a sample size that at least matches the sample size of the original study; (3) the exact replication study must not be conducted by people related to the authors of the target experiment and the author(s) of the replication must disclose this possibility explicitly in the cover letter accompanying the manuscript.

After identifying the target experiment, the introduction of such an article must replace the target experiment in context, why it was done, the impact it had right after publication, the impact it still has today if some years have elapsed.

The method section should contain only precisions when methodological details were unclear and list changes brought to the target experiment (e. g., if an old apparatus is replaced by a newer one). Of course, changes should be as minimal as possible, and dictated by necessity (for example, who would be able to make a tachistoscope works nowadays?).

The result section should contain a list of the effects found in the target article and compares them with the effects present or absent in the replication study, expressed in terms of raw effect sizes, standardized effect sizes, and p values. The comparison should be accompanied by a table containing two columns, one for the target experiment, one for the replication experiment. There is no need to discuss the results at length; bare facts are sufficient. Hence, the result section could be composed of the above table only.

The discussion should indicate how well the results replicate qualitatively as well as quantitatively. Little focus is given to p values as they are probably the least replicable statistics (Cumming, 2014). If the replication fails, the author(s) may conjecture why this is so if they wish.

Overall, a replication study should result in a short paper, a manuscript 5 to 8 pages long in APA format.

Every replication study article will have the subtitle "A replication of " followed by a reference to the name and year of the target article (and the experiment number if this is a multi-experiment article). The journal's web site will add next to the subtitle an icon, ☺, ☹, or ☹, to highlight whether the replication is judge successful, mix or unsuccessful, as evaluated by the authors of the replication study.

The journal might on occasion publish more than one replication of a given target article. At some point,

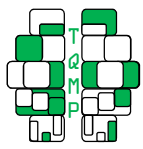
the editor may stop publishing replications if the previously published replications converge onto an unambiguous conclusion. At this time, we do not have a firm and objective methodology which would determine when enough replications have been published. It might be based on Wald's (1947) ideas of sequential sampling. Each replication could be considered a new "observation" of the effect under scrutiny. Each new observation would be entered into a random walk model with two criteria: the positive criterion indicates that the effect is reproducible; the negative criterion indicates that the effect is not reproducible. The *Quantitative Methods for Psychology*, in addition to be calling for reproduction studies, is also asking for contributions on this specific question: *how many replications are enough replications?*

Author's note

I would like to thank David Asselin and Camille Blais-Rochette for discussions and comments on a previous version of this editorial, and Camille Blais-Rochette, Geneviève Boudreault, Samantha Burns, Natasha Plourde, Julie Raymond, Sabrina Schmiedel, Caroline Seguin and Marie-Pier Vandette, for constructive discussions regarding replications and statistics.

References

- Aldhous, P. (2011). *Journal rejects studies contradicting precognition*, Internet resource found at http://Journal_rejects_studies_contradicting_precognition.html, last consulted 3/mai/2014.
- Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, *100*, 407-425.
- Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*, 7-29.
- Francis, G. (2012). Publication bias and the failure of replication in experimental psychology. *Psychonomic Bulletin & Review*, *19*, 975-991.
- Grahe, J. E., Reifman, A., Hermann, A. D., Walker, M., Oleson, K., Nario-Redmond, M., & Wiebe, R. P. (2012). Harnessing the undiscovered resource of student research projects. *Perspectives on Psychological Science*, *7*, 605-607.
- Ioannidis, J. P. A. (2012). Why Science Is Not Necessarily Self-Correcting. *Perspectives on Psychological Science*, *7*, 645-654.
- Jasny, B., R., Chin, G., Chong, L., & Vignieri, S. (2011). Again, and again, and again. *Science*, *334*, 1225-



1225.

- Koole, S. L., & Lakens, D. (2012). Rewarding replications: A sure and simple way to improve psychological science. *Perspectives on Psychological Science*, 7, 608-614.
- Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur?. *Perspectives on Psychological Science*, 7, 537-542.
- Pashler, H., & Wagenmakers, E.-J. (2012). Editor's

introduction to the special section on replicability in psychological science: A crisis of confidence?. *Perspectives on Psychological Science*, 7, 528-530.

- Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences, *Review of General Psychology*, 13, 90-100.
- Wald, A. (1947). *Sequential analysis*. New York: John Wiley and sons.
- Yong, E. (2012). Bad copy. *Nature*, 485, 298-300.

Citation

Cousineau, D (2014). Restoring confidence in psychological science findings: A call for direct replication studies. *The Quantitative Methods for Psychology*, 10 (2), 77-79.

Copyright © 2014 Cousineau. This is an open-access article distributed under the terms of the *Creative Commons Attribution License (CC BY)*. The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Received: 10/06/14