Solving the Replication Problem in Psychology Requires Much More Than a Website

GORDON B. SCHMIDT
Indiana University Purdue University Fort Wayne

RICHARD N. LANDERS
Old Dominion University

As Kepes and McDaniel (2013) describe, there is significant concern regarding the trustworthiness of the industrial and organizational (I–O) psychology research base. A primary cause of this concern is the lack of replications, which currently account for only about 1% of published studies in psychology (Makel, Plucker, & Hegarty, 2012). This is a significant issue, as replication is needed for science to progress by examining the rigor and reproducibility of research results. As part of their vision, Kepes and McDaniel call for more exact replications, suggesting that such replications might be included as back matter in

Correspondence   concerning this article should be addressed to Gordon B. Schmidt.
E-mail: schmidtg@ipfw.edu
  Address: IPFW, OLS Neff 288D, 2101 Coliseum Blvd., Fort Wayne, IN 46805

journals with a similar placement as book reviews.

Lack of Replications in Psychology Is a Symptom of a Larger Problem

Although we agree with Kepes and McDaniel that increased exact replication is needed, we would argue that poor attitudes toward the value of replication and an associated lack of incentives are the underlying causes of this particular problem. Psychology does not incentivize researchers to conduct replications. Studies that find null results, regardless of research method quality, are less likely to be accepted for publication in lead journals than those finding statistical significance (Greenwald, 1975; Orlitzky, 2012). Journals seek research papers that offer major theoretical contributions rather than examine the reproducibility of previous work (Academy of Management Journal, 2012; Kozlowski, 2009). Because null results and replications are generally less valued, the probability of publication for replications is low, which ultimately means that there is reduced incentive for scholars to pursue them, particularly those facing tenure. We believe this problem is systemic to many fields in modern academia. Although replications are valuable to advance the science of I–O psychology as a whole, they are not rationally desirable for many academic I–O psychologists, and the solution presented by Kepes and McDaniel to increase replications—adding a special section at the back of journals—is probably not sufficient by itself to encourage them.

Creating a Home for Replications Alone Does Not Solve This Problem

The lack of respect for replications and resulting lack of incentive to conduct replications in psychology are likely the reasons for the general lack of success of the websites cited by Kepes and McDaniel as previously attempted solutions to the replication problem. Psychfiledrawer.org offers a place for researchers to post replication study results and discuss their results, as did an earlier attempt at a similar site (now defunct) by the second author (R.N.L.) of this commentary called filedrawer.org. Open Science Framework offers a place for researchers to document their studies, share data, and share results (Open Science Collaboration, 2012). Although these sites offer a venue for presenting replication results, researcher use of such sites has remained quite low. In its 5 years of existence (2006–2010), only five studies were uploaded to filedrawer.org. Later efforts have been marginally more successful. Psychfiledrawer.org and Open Science Framework, both created in 2011, report 28 and 29 studies, respectively.

The challenge faced by these sites is not design or a lack of technology to offer a robust system for reporting replications; instead, their relative unpopularity is due to lack of incentive. Posting data on one of these open source sites does not count as a "publication" to universities and thus does not provide the traditional employment-related rewards for researchers (e.g., tenure, merit raises). When filedrawer.org was operational, a typical concern reported to the second author (R.N.L.) of this paper was that once a researcher's replication or null result study was uploaded, it would be considered "published" and thus unable to be later published in traditional journals, should the opportunity ever arise. To try to combat this, this author contacted editorial boards of several APA journals, who reported that this concern was valid and a compelling reason not to use such online websites. Psychfiledrawer.org addresses this issue by licensing all its content under a creative commons license (psychfiledrawer.org, 2013), effectively preventing any researcher who posts there from ever releasing the rights to that work to a publisher. Open Science Framework has tried to combat this issue through their current Reproducibility Project, where teams of researchers are replicating each study from 2008 in three major psychology journals

(*Journal of Personality and Social Psychology*, *Psychological Science*, and *Journal of Experimental Psychology: Learning, Memory, and Cognition*). They plan to submit a paper summarizing all of the replications with the authorship of that summary paper including all authors of the replications studies. Although this is a noble effort, the lag from doing the replication study to even ultimate submission of the summary paper will be long, with the site planning to wait until all research groups are done with their replications. This puts the timeline for ultimate career-related value gained from the replication study relatively far in the future. It is also limited to a handful of participating research labs with interests in this particular domain of psychology. We see this as a worthy cause but unlikely to be a general solution to the lack of replication problem.

These open source sites also lack critical mass to be broadly valuable. In their current states, these sites only sporadically cover the research base. More studies are needed to provide value, but people will only be motivated to provide such coverage when their submissions look like they will be meaningful. Thus, purely technology-based solutions face a catch-22: For such sites to grow, they must be large enough for users to find them valuable and credible, but they also need a substantial influx of users to provide that value. This influx is unlikely to occur given the current lack of respect paid to replication efforts.

A Solution That Increases Replication and Provides Career Incentives

From this, we conclude that current open source options do not provide adequate incentives to fulfill the vision of increased production of replications envisioned by Kepes and McDaniel, and that their current lack of popularity reflects this. However, we recognize the value that such Internet-based organizational efforts could bring to increasing replication efforts. As such, we propose an alternate Internet-based structure integrated into the traditional publication process that could achieve the goal of increased replications in the psychology research literature. We propose that each issue of industrial and organizational (I–O) psychology journals contain a call for replication of empirical studies within that issue, with journal websites or a centralized database integrated with those journal websites used as a means to organize these efforts. Once a replication has been submitted, it would be published as a short (a page or less, perhaps) ''replication report'' in the back pages of a later issue.

This retains many of the strengths of Kepes and McDaniel's vision but recognizes the realities faced by academics in pursuit of published work. Journal websites or a connected hub website organized by one of our academic societies would be available for researchers to sign up to conduct replications. For each focal article, a limited number of slots for researchers to sign up to perform replications might be available. Such replications could be mixed in purpose if desired: In addition to the gold standard of exact replications, conceptual replications, replications in a field setting, and replications with a different but reasonable target population might be requested. Journal editors and authors could, if desired, have great latitude in defining where and how additional replications would be conducted. These replications would thus provide a more directed path for submission, with clear guidelines provided for how replications should be presented and discussed, briefly but informatively. The quality of replications is kept high by designing this process up front; in a sense, these replications are peer reviewed before they are started. Once established, such websites might also serve as a foundation on which to build the broader research registries also described by Kepes and McDaniel.

If replication studies become sufficiently popular in a particular journal that they could not be easily contained in a few pages in the end of an issue, that journal might create a mirror journal exclusively

for replications of studies (e.g., *Journal of Applied Psychology Replications*). Such a mirror journal could be in print or presented online only. Critical to the attractiveness of this approach is that the replication journal must be indexed by PsycINFO; to maintain the incentive structure, such mirror journals must be treated equivalently or near equivalently to traditional journals by publishers.

Remaining Challenges

Certainly this system does not resolve all potential problems. Although such replications would be formally published, this does not necessarily mean that they would or should be regarded as of equal scholarly value to nonreplications. Would a publication in *Journal of Applied Psychology* be equivalent to one in *Journal of Applied Psychology Replications*? How many replications represent the same scholarly impact as a single theoretical or rigorous empirical publication? For the system we describe to function as intended, replication studies must be viewed as providing a sufficient contribution for it to be worthwhile for researchers to spend time and resources conducting them, and this may vary greatly even from department to department.

This system also goes against the grain of current I–O journal emphasis on theoretical contributions being essential for research work (Academy of Management Journal, 2012; Kozlowski, 2009). Although we see the value in both theory-driven work and replications in I–O psychology, the system here would represent an ideological shift from current trends. We suspect that resistance to the idea of ''just'' replicating studies will initially be rather high, even with the improved incentive structure we describe. However, there is no path to increase replications that will not face such a challenge. With support from editors of top-tier journals, and formal efforts to support replications by those journals, we suspect this resistance would be much reduced.

Finally, this system will require a buy in and additional effort from editors and publishers. Much of this effort would happen on the front end, as websites and related tools would need to be created and deployed. A basic process of replications being organized through the site, tracked, and ultimately submitted and published would need to be established. Once the system was running however, the work would mostly involve adding new issues to the replication website and examining replications for compliance with guidelines as they were submitted. If a journal received sufficient replications to warrant a mirror journal, the workload would be much increased. However, this increased effort would be in response to increased replications, so this would be a desirable challenge to face. Overall, we feel that such efforts are required to have any hope to make replications a part of mainstream research in I–O psychology.

Conclusions

Overall, we agree with Kepes and McDaniel that the lack of replications in I–O psychology is a critical problem, but we believe the allocation of a few pages at the back of journals is not sufficient by itself to increase such replications. We propose instead that Kepes and McDaniel's vision can be better achieved if journals establish formal processes to publish replications, with organization of such effort conducted online, explicitly recognizing the value of these replications to scientific progress. Such a system would allow replications to proceed in a systematic fashion and give researchers proper incentive to do so. Although this system will increase the effort required from editors and publishers, the benefits to our field and the rigor of its science could be sizable.

In many ways, we face a problem potentially solvable by an application of goal setting theory. In academia, a driving goal of researchers is to produce work publishable in traditional journals. Creating an entirely new outlet—and one that will be seen as inferior by many traditionalists—creates an entirely new goal for researchers. We propose here that such a goal is unlikely to be

pursued and that the current lack of popularity of open source websites is evidence of this shortcoming. This problem is larger than psychology alone; it would be a very large task indeed to change the negative attitudes towards work posted on websites among all academics. Even with the support of colleagues in psychology, convincing cross-disciplinary promotion and tenure committees to recognize replications published only on open source websites as ''real'' contributions to science may be unachievable for many years. But if we can instead recognize such work as valuable by publishing it alongside other high-quality papers in traditional journals, emphasizing the value provided by these replications, embedding them within a larger formally recognized system, we can incentivize researchers to produce replications without changing their publication goals or asking them to potentially sacrifice their future career success.

References

Academy of Management Journal. (2012). *Information for contributors*. Retrieved from http://aom.org/ Publications/AMJ/Information-for-Contributors. aspx

Greenwald, A. G. (1975). Consequences of prejudice against the null hypothesis. *Psychological Bulletin*, *82*, 1–20. doi: 10.1037/h0076157

Kepes, S., & McDaniel, M. A. (2013). How trustworthy is the scientific literature in I-O psychology? *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *6*(3), 252–268.

Kozlowski, S. W. J. (2009). Editorial. *Journal of Applied Psychology*, *94*, 1–4. doi: 10.1037/a0014990

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. doi: 10.1177/ 1745691612460688

Open Science Collaboration (2012). An open, large-scale, collaborative effort to estimate the reproducibility of psychological science. *Perspectives on Psychological Science*, *7*, 657–660. doi: 10.1177/1745691612462588

Orlitzky, M. (2012). How can significance tests be deinstitutionalized? *Organizational Research Methods*, *15*, 199–228. doi: 10.1177/10944281 11428356

Psychfiledrawer.org. (2013). *FAQ*. Retrieved from http://www.psychfiledrawer.org/faq.php