Industrial and Organizational Psychology (2020), **13**, 28–31
doi:10.1017/iop.2020.6

CAMBRIDGE UNIVERSITY PRESS

COMMENTARY

Using results-blind reviewing to support the peer review competency framework

Liana Kreamer* and Steven G. Rogelberg

University of North Carolina at Charlotte
*Corresponding author. Email: lkreamer@uncc.edu

The focal article by Köhler et al. (2020) develops a highly compelling competency framework for excellent reviewing. It is certainly a key piece of a robust and open science ecosystem (Banks et al., 2019). The framework can and should be used as a guide for novice and seasoned reviewers "in defining what they should pay attention to, upon what their review should touch, how they should word their review, what kind of advice they need to provide, and how they should overall go about peer reviewing" (Köhler et al., 2020, p. 17). Creating specific reviewer training around the competencies takes this framework to the next level by actively teaching reviewers how to implement the proposed guidelines. Developing these skills is a critical component for bettering the publication process by enabling reviewers to provide feedback to colleagues that is more organized, reliable, and helpful.

The focal article, by design, has an individual focus. Thus, this commentary will serve as an *extension* to the focal article, pushing the call for better reviewing further by making the case for results-blind reviews (RBR) as an environmental mechanism that better supports the framework and training of competencies discussed in the focal article.

A conducive environment

Literature on transfer-of-training explores under what conditions knowledge and skills learned in training are applied in action (Baldwin & Ford, 1988). This area of research emphasizes the importance of situational cues and environmental factors, noting that these largely determine whether or not learned competencies are actually applied in the workplace (Blume, Ford, Baldwin, & Huang, 2010; Colquitt, Lepine, & Noe, 2000). It is our contention that the training of competencies alone may not be enough for successful transfer of reviewing skills as situation/context matters to what behaviors are exhibited by individuals (Dalal, Alaybek, Sheng, Holland, & Tomassetti, in press). Furthermore, the training focuses on explicit behaviors of reviewers, but does not target the implicit or unconscious biases that may exist in the system itself.

Bias in the dominant peer-review system

The current dominant peer review system in the fields of industrial and organizational psychology (I-O), organizational behavior (OB), and human resources (HR) requires authors to submit full manuscripts, including results, to a journal. There is good reason to believe that manuscript results can bias reviewers' judgments toward the rest of the article. This, in turn, can undermine the efficacy of the focal article's framework.

Research suggests that favoritism toward significant and novel findings in the organizational sciences is widespread (Antonakis, 2017; Murphy & Aguinis, 2019; O'Boyle, Banks, Carter, Walter, & Yuan, 2019; O'Boyle, Banks, & Gonzalez-Mulé, 2017), and null findings are typically

© Society for Industrial and Organizational Psychology, Inc. 2020.

https://doi.org/10.1017/iop.2020.6 Published online by Cambridge University Press

undervalued (Coursol & Wagner, 1986; Emerson, Warme, Wolf, Heckman, Brand, & Leopold, 2010). Emerson et al. (2010) illustrated this bias, having conducted a randomized experiment where two versions of a fabricated manuscript were submitted to peer review at two different journals. The manuscripts were identical, with the only difference being that one version reported positive, statistically significant results while the other version reported no significant differences (aka, the null-outcome version). The authors also purposely placed five identical errors throughout the two versions. Reviewers at both journals favored the positive-outcome manuscript version, more frequently recommending it for publication than the null-outcome version (97% vs. 80%). Moreover, reviewers detected more of the errors in the null-outcome version than the positive-outcome manuscript. Furthermore, the null-outcome version received lower scores for methodological validity than the positive-outcome manuscript. These results suggest that reviewers often interpret statistically significant findings as a consequence of methodological rigor and are more likely to recommend positive-outcome manuscripts for publication (Emerson et al., 2010). This ultimately contributes to a publication bias in the literature, where knowledge is not entirely representative of research.

As the focal article suggests, training reviewers to look past the significance of findings and focus solely on the rigor of the research design could potentially reduce this bias. However, is a cautionary warning truly enough to eliminate implicit bias? It is our contention that blinding the research results of a manuscript under review will invariably reduce the bias caused by significant, positive, and novel results—as the results are not shown prior to the decision. This is akin, for example, to what was found in the sex-based hiring study conducted by Goldin and Rouse (2000). The researchers investigated systematic bias in the selection of symphony orchestras in the United States by implementing blind auditions, concealing candidates' identities from the orchestra's selection committee. The authors demonstrated that the blind audition reduced sex-based implicit biases in hiring decisions, increasing the probability of a women being selected to the orchestra. Blind auditioning promoted impartiality in the committees' hiring decisions, increasing the proportion of women in the symphony orchestra by roughly 50%. Other examples of studies like this exist across the selection, performance appraisal, and decision-making literature that strengthen the case for changing the system to yield better and more accurate outcomes.

Results blind reviews

There are two stages in RBR. In Stage 1 of RBR, researchers send in a manuscript including an introduction, theoretical/conceptual rationale, relevant literature review, hypotheses, detailed methodology, and plan for analyses (Grand et al., 2018). The actual results do not matter since compelling questions answered in appropriate ways is what should drive what is published, not the actual findings. On the basis of these criteria, reviewers may reject, invite for further revision, or recommend an in-principle acceptance to the researchers. An in-principle acceptance is a formal agreement stating that the research will be published as long as the procedures are followed and the proposed/agreed upon analyses are conducted as described. Overall, RBR is intended to alter the lens through which researchers and reviewers approach and evaluate research (Grand et al., 2018). Assuming the researchers fulfilled the promises indicated in their manuscript at Stage 1, their article would be accepted—regardless of supported hypotheses or significant findings.

In Stage 2, only upon receiving an in-principle acceptance, the researchers submit the full manuscript, including the results and discussion. The manuscript is assigned to the same reviewers from Stage 1. Assuming there are no unjustifiable deviations from the original proposal and all analyses are handled competently, the manuscript is accepted and published. Reviewers must focus on the quality and rationale of the study, evaluating on the basis of the appropriateness of research questions and hypotheses, the practical or theoretical importance, and the rigor of the methodology without knowledge of (or being biased by) results. It is important to note that RBR

https://doi.org/10.1017/iop.2020.6 Published online by Cambridge University Press

processes are highly akin to how we train our graduate students. Namely, students create thesis and dissertation proposals. The theoretical rational, research questions/hypotheses, and proposed methodology are highly scrutinized. But, once accepted by the committee (similar to an in-principle acceptance), the students' passing of these milestones is not actually affected by what is found in the research.

There is another potential benefit of the RBR process. It could lead to authors being more transparent and honest in their reporting of results. The preference for significant results may incentivize researchers to engage in questionable research practices (QRPs), such as presenting post hoc findings as a priori, HARKing, selectively reporting fit indices, and strategically deleting outliers for the purpose of reaching statistical significance (Banks, Rogelberg, Woznyj, Landis, & Rupp, 2016). The QRPs mentioned above lead to a publication bias (i.e., file drawer problem) in academic journals, potentially hindering the advancement of knowledge by only displaying supportive findings or meaningful outcomes. The RBR process generally eliminates implicit or explicit incentives for researchers to engage in QRPs, instead encouraging scholars to submit manuscripts regardless of whether they found significant results (Button et al., 2016).

Woznyj, Grenier, Ross, Banks, and Rogelberg (2018) explored the field's reactions (including practitioners) to the RBR initiative and found that a majority of stakeholders in the organizational sciences (roughly 75%) believes the field would benefit from blind reviewing. Seventy-two percent of stakeholders believe the initiative would help advance cumulative knowledge, while 65% believe RBR would combat against questionable research practices. In regard to RBR benefiting the peer-review process, about one-fourth (22%) of stakeholders indicated blind results would shift how reviewers evaluate research. One respondent stated that the initiative would "focus reviewer attention on theory and the testing of theory rather than on statistical support," while another respondent noted "reviewers and editors would probably focus a lot more on the quality of the arguments and hypothesis development than they currently do" (Woznyj et al., 2018, p. 567).

With this all said, it is important to acknowledge some potential unknowns with the RBR process that warrant further examination. First, will reviewers, both academics and practitioners, like to review articles in this fashion? It is our contention that they will for a couple of reasons: (1) The Stage 1 manuscript is actually shorter than a full manuscript, thus it should take less time to read and review; and (2) RBR processes are basically identical to what is done for a thesis and dissertation process in graduate school. For example, if students effectively carry out their proposal (akin to a Stage 1 manuscript), they are given an in-principal acceptance—they just need to deliver on what was approved during the initial proposal meeting. Thus, the RBR process in many regards should feel comfortable and familiar. Second, does the RBR process require a longer reviewer commitment? This is certainly worthy of examination. We do not believe it will. Namely, the RBR process just changes the rhythms of reviewing. If a Stage 1 manuscript is given an in-principal acceptance, reviewer time with the Stage 2 manuscript should be absolutely minimal, as reviewers are just confirming the article was carried out as described in the Stage 1 manuscript and that the discussion section is appropriate and well-done. Typically, the Stage 2 review processes should be completed in one or, at most, two rounds.

Conclusion

The focal article develops a much-needed competency framework for training reviewers, offering concrete advice for how we can improve our reviewing skills. Training reviewers via the framework is a great step toward solving some of the issues with the peer-review process in the organizational sciences. This commentary extends the focal argument by suggesting a complementary solution, results-blind review, that addresses the scientific ecosystem at large. Using RBR, reviewers can more easily and successfully enact the competencies proposed in the focal article, ultimately advancing our science and improving the published research that represents our field.

https://doi.org/10.1017/iop.2020.6 Published online by Cambridge University Press

References

Antonakis, J. (2017). On doing better science: From thrill of discovery to policy implications. *The Leadership Quarterly*, **28**(1), 5–21.

Baldwin, T., & Ford, J. (1988). Transfer of training: A preview and directions for future research. *Personnel Psychology*, **41**(1), 43–63.

Banks, G. C., Field, J. G., Oswald, F. L., O'Boyle, E. H., Landis, R. S., Rupp, D. E., & Rogelberg, S. G. (2019). Answers to 18 questions about open science practices. *Journal of Business and Psychology*, **34**, 257–270.

Banks, G. C., Rogelberg, S. G., Woznyj, H. M., Landis, R. S., & Rupp, D. E. (2016). Evidence on questionable research practices: The good, the bad, and the ugly. *Journal of Business & Psychology*, **31**(3), 323–338.

Blume, B., Ford, J., Baldwin, T., & Huang, J. (2010). Transfer of training: A meta-analytic review. *Journal of Management*, **36**(4), 1065–1105.

Button, K. S., Bal, L., Clark, A., & Shipley, T. (2016). Preventing the ends from justifying the means: Withholding results to address publication bias in peer-review. *BMC Psychology*, **4**(1), 59.

Colquitt, J., Lepine, J., & Noe, R. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *The Journal of Applied Psychology*, **85**(5), 678–707.

Coursol, A., & Wagner, E. (1986). Effect of positive findings on submission and acceptance rates: A note on meta-analysis bias. *Professional Psychology: Research and Practice*, **17**(2), 136–137.

Dalal, R., Alaybek, B., Sheng, Z., Holland, S., & Tomassetti, A. (In press). Extending situational strength theory to account for situation-outcome mismatch. *Journal of Business and Psychology*. doi:10.1007/s10869-019-09632-z

Emerson, G. B., Warme, W. J., Wolf, F. M., Heckman, J. D., Brand, R. A., & Leopold, S. S. (2010). Testing for the presence of positive-outcome bias in peer review: A randomized controlled trial. *Archives of Internal Medicine*, **170**(21), 1934–1939.

Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of 'blind' auditions on female musicians. *American Economic Review*, **90**(4), 715–741.

Grand, J., Rogelberg, S., Allen, T., Landis, R., Reynolds, D., Scott, J., ... **Truxillo, D.** (2018). A systems-based approach to fostering robust science in industrial-organizational psychology. *Industrial and Organizational Psychology*, **11**(1), 4–42.

Köhler, T., González-Morales, M. G., Banks, G. C., O'Boyle, E. H., Allen, J. A., Sinha, R., ... Gulick, L. M. V. (2020). Supporting robust, rigorous, and reliable reviewing as the cornerstone of our profession: Introducing a competency framework for peer review. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, **13**(1), 1–27.

Murphy, K. R., & Aguinis, H. (2019). HARKing: How badly can cherry-picking and question trolling produce bias in published results? *Journal of Business and Psychology*, **34**, 1–17.

O'Boyle, E., Banks, G. C., Carter, K., Walter, S., & Yuan, Z. (2019). A 20-year review of outcome reporting bias in moderated multiple regression. *Journal of Business and Psychology*, **34**, 19–37.

O'Boyle, E., Banks, G., & Gonzalez-Mulé, E. (2017). The Chrysalis Effect: How ugly initial results metamorphosize into beautiful articles. *Journal of Management*, **43**(2), 376–399.

Woznyj, H., Grenier, K., Ross, R., Banks, G., & Rogelberg, S. (2018). Results-blind review: A masked crusader for science. *European Journal of Work and Organizational Psychology*, **27**(5), 561–576.

Cite this article: Kreamer, L. and Rogelberg, S.G. (2020). Using results-blind reviewing to support the peer review competency framework. *Industrial and Organizational Psychology* **13**, 28–31. https://doi.org/10.1017/iop.2020.6

https://doi.org/10.1017/iop.2020.6 Published online by Cambridge University Press