Effective Peer Review: Who, Where, or What?

Peer review is widely viewed as one of the most critical elements in assuring the integrity of scientific literature (Baldwin, 2018; Smith, 2006). Despite the widespread acceptance and utilization of peer review, many difficulties with the process have been identified (Hames, 2014; Horrobin, 2001; Smith, 2006). One of the primary goals of the peer review process is to identify flaws in the work and, by so doing, help editors choose which manuscripts to publish. It is surprising that one of the persistent problems in peer review is assessing the quality of the reviews. Both authors and journal editors expect peer review to detect errors in experimental design and methodology and to ensure that the interpretation of the findings is presented in an objective and thoughtful manner. In traditional peer review, two or more reviewers are asked to evaluate a manuscript on the basis of the expectation that if the two reviewers agree on the quality of the submission, the likelihood of a high-quality review is increased. Unfortunately, studies have not consistently confirmed a high degree of agreement among reviewers. Rothwell and Martynn (2000) evaluated the reproducibility of peer review in neuroscience journals and meeting abstracts and found that agreement was approximately what would be expected by chance. Similarly, Scharschmidt et al. (1994) found similar results in the evaluation of 1,000 manuscripts submitted to the *Journal of Clinical Investigation,* where clustering of grades in the middle resulted in an agreement being "…only marginally…" better than chance. These observations suggest that we cannot rely on the agreement of reviewers to be an indication of the quality of the reviews. Another potential way to evaluate the quality of reviews would be to assess the ability of reviewers to detect errors in submissions. It is generally accepted that detection of intentional fraud is beyond the scope of typical peer review, but we do expect reviewers to detect major and minor errors as a primary function of the traditional peer review system (Hwang, 2006; Weissman, 2006). Schroter et al. (2008) evaluated the ability of reviewers to detect major and minor errors by introducing errors into three previously published papers describing randomized controlled clinical trials. Reviewers detected approximately three of the nine errors introduced in each manuscript. Unfortunately, reviewers who had undergone training in how to conduct a high-quality peer review were not significantly better than untrained reviewers. Similar results have been reported by Godlee et al. (1998) and Baxt et al. (1998). Baxt et al. (1998) did report that reviewers who rejected or suggested revision of a manuscript identified more errors than those who accepted the manuscript (decision: 17.3% of major errors detected [accept], 29.6% of major errors detected [revise], and 39.1% of major errors detected [reject]). It is almost certainly true that the extent of the failure to recognize errors in submitted manuscripts may differ among scientific disciplines and journals. It also however seems likely that these observations do have some applicability to journals such as *JID Innovations*. It is critical that both authors and editors are cognizant of these limitations of peer review in their assessment of reviews. These findings compel journals to continue to work to develop new strategies to train and evaluate reviewers. The findings also suggest that factors beyond the failure to detect objective mistakes in a manuscript may be playing a role in the discrepancy in reviewers' evaluations. One area of ongoing concern in the peer review process is the role of reviewer bias in assessing the scientific work of colleagues (Kuehn, 2017; Lee et al, 2013; Tvina et al, 2019).

Bias in the peer review process can take many forms, including collaborator/competitor bias, affiliation bias based on an investigator's institution or department, geographical bias based on the region or country of origin, racial bias, and gender or sex bias (Kuehn, 2017; Lee et al, 2013; Tvina et al, 2019). All of these forms of bias present the risk that a decision of the reviewer will not be based solely on the quality or merit of the work but rather be influenced by a bias of the reviewer. We and other journals routinely seek to avoid selecting individuals to review work from their own institutions and ask all reviewers to declare any potential personal conflicts of interest. All these methods require either the editor or the reviewer to identify a bias and fail to address the issue of implicit or unconscious reviewer bias. The dominant method currently utilized for peer review is the so-called single-blind review,

JID Innovations (2022);**2:**100162
doi:10.1016/j.xjidi.2022.100162

in which the identity and affiliations of the authors are known to the reviewers, whereas the identity of reviewers remains unknown to the authors. This has led to concern that knowledge of the identity of the authors and their institutions may be the source of significant reviewer bias, especially implicit bias, in the evaluation of manuscripts. Double anonymized peer review (DAPR), also known as double-blind peer review, has been suggested as a way to address this issue (Bazi, 2020; Lee et al, 2013). Studies have compared single-blind with double-blind reviewing and reported that there is no significant difference in the quality of the reviews (Alam et al, 2011; Godlee et al, 1998; Justice et al, 1998; van Rooyen et al, 1998). Although these studies looked at measures such as the number of errors detected, acceptance rate, and distribution of initial reviewer scores, they were not designed to address specific sources of bias such as authors' gender, institution, or geographic location. Other studies have been undertaken to directly address the issue of bias in the peer review process. Ross et al (2006) compared the acceptance of abstracts submitted to the American Heart Association's annual scientific meeting during a period when the reviewers knew the identity and origin of the authors (i.e., single-blind review) with when this information was not known by the reviewers (i.e., double-anonymized peer review). They found a significant increase in acceptance of non–United States abstracts and abstracts from non-English speaking countries when the reviewers were unaware of the country of origin of the abstracts (Ross et al, 2006). They also found a significant decrease in the acceptance of abstracts from prestigious institutions when the reviewers were unaware of the institutions where the work was done. In a similar study, Tomkins et al. (2017) found that papers submitted to a prestigious computer science meeting were more likely to be accepted if they were from famous authors, top universities, and top companies. Okike et al. (2016) documented similar results for manuscripts submitted to the orthopedic literature. They submitted a fabricated manuscript that was presented as being written by two prominent orthopedic surgeons (past Presidents of the American Academy of Orthopedic Surgeons) from prestigious institutions. When reviewed in the traditional single-blind fashion, which included the identity of the authors, the manuscript was accepted by 87% of the reviewers. By contrast, when the identity of the authors was unknown, the manuscript was accepted by 68% of the reviewers ($P = 0.02$) (Blank, 1991). A study conducted at *The American Economic Review* found that authors at near-top-ranked universities experienced lower acceptance rates when authorship was anonymized (Blank, 1991). Of interest, they also found that for women, there was no difference in the acceptance rate between the double-anonymized and single-blinded reviews; however, for men, the acceptance rate was lower with double-anonymized reviews.

These studies provide strong evidence that knowledge of who and where the study was performed can impact the acceptance of abstracts and manuscripts. This conflicts with the goal of the review process to base our judgments on the quality of what the results demonstrate. It is difficult to estimate how much this may affect the fate of a manuscript at *JID Innovations*. We do not have evidence that our review process has been impacted by bias as is reported in the studies discussed. However, neither can we state with certainty that such bias is not a factor in the reviews we receive. One of the goals of *JID Innovations* is to be a truly open-access journal available to all investigators in skin science from around the world. We have sought to be an outlet for studies that challenge existing paradigms or that may report negative results. We want to be seen as providing fair and objective reviews for all authors, regardless of where they work or who they are. If we are to achieve this goal, it is imperative that the who and where of a specific manuscript do not negatively impact the evaluation of the what. We want young investigators, investigators at less prestigious institutions or from less well-known laboratories, and investigators from any country around the world to be confident that their work will be judged by what they report and not by the who and the where.

To be true to this mission, *JID Innovations* will be initiating DAPR starting in October 2022. This is not being done because we are aware of any issues of bias with our current process of peer review but because we realize that the absence of proof is not proof of absence. As a part of this process, authors will be asked to remove identifying material from manuscripts at the time of submission in preparation for the review process (https://www.jidinnovations.org/content/authorinfo). As a result, primary reviewers will see only the what of the manuscript. We realize that this process involves extra work for both the authors and our staff, but we feel the benefits will outweigh this small cost. Indeed, in other journals that have taken this step, surveys have shown that both authors and reviewers ultimately prefer double-anonymized reviews (Bennett et al, 2018; Moylan et al, 2014). We realize that achieving 100% anonymization of a manuscript is nearly impossible. Studies have shown that the rate of successful anonymizing, where the reviewers cannot discern the authorship of a manuscript, ranged from 47 to 73%. It is however interesting that even with this rate of success in the anonymizing process, a meta-analysis of trials of double- versus that of single-blind peer review has suggested an impact, with lower acceptance rates with double-anonymized peer review (Ucci et al, 2022). More work clearly needs to be done to assess the value of the DAPR process, and we will be monitoring our results carefully.

The institution of DAPR in *JID Innovations* will assure our authors that the what of their manuscript is our focus. It does not matter who you are or where you are from. It will also emphasize to our reviewers that our focus is on the what. We will be carefully monitoring the results of this new policy and plan to report back on our experience. We also welcome your feedback on your experience as a reviewer and author for *JID Innovations*; send your comments to us at InnovationsEditor@sidnet.org.

Finally, this decision should be seen not as the end of our efforts to improve the peer review process but merely as a first step. We will continue to work to improve all aspects of the peer review process for *JID Innovations*. We firmly believe that the use of double-blind -anonymized peer review will bring us closer to ensuring to our authors and readers that the work that is published by *JID Innovations* has been selected

on the basis of what the paper reports and not on who performed the studies or where they were located.

CONFLICT OF INTEREST

The author states no conflicts of interest.

Russell P. Hall III

Editor[*]

[*]*Correspondence: Russell P. Hall, III. e-mail: InnovationsEditor@sidnet.org*

Cite this article as: *JID Innovations* 2022;2:100162.

REFERENCES

Alam M, Kim NA, Havey J, Rademaker A, Ratner D, Tregre B, et al. Blinded vs. unblinded peer review of manuscripts submitted to a dermatology journal: a randomized multi-rater study. Br J Dermatol 2011;165:563–7.

Baldwin M. Scientific autonomy, public accountability, and the rise of "peer review" in the Cold War United States. Isis 2018;109:538–58.

Baxt WG, Waeckerle JF, Berlin JA, Callaham ML. Who reviews the reviewers? Feasibility of using a fictitious manuscript to evaluate peer reviewer performance. Ann Emerg Med 1998;32:310–7.

Bazi T. Peer review: single-blind, double-blind, or all the way-blind? Int Urogynecol J 2020;31:481–3.

Bennett KE, Jagsi R, Zietman A. Radiation oncology authors and reviewers prefer double-blind peer review. Proc Natl Acad Sci USA 2018;115:E1940.

Blank RM. The effects of double-blind versus single-blind reviewing: experimental evidence from the American Economic Review. Am Econ Rev 1991;81:1041–67.

Godlee F, Gale CR, Martyn CN. Effect on the quality of peer review of blinding reviewers and asking them to sign their reports: a randomized controlled trial. JAMA 1998;280237–40.

Hames I. Peer review at the beginning of the 21st century. Sci Ed 2014;1:4–8.

Horrobin DF. Something rotten at the core of science? Trends Pharmacol Sci 2001;22:51–2.

Hwang WS. Can peer review police fraud? Nat Neurosci 2006;9:149.

Justice AC, Cho MK, Winker MA, Berlin JA, Rennie D. Does masking author identity improve peer review quality? A randomized controlled trial. PEER Investigators [published correction appears in JAMA 1998;280:968. JAMA 1998;280:240–2.

Kuehn BM. Rooting out bias. ELife 2017;6:e32014.

Lee CJ, Sugimoto CR, Zhang G, Cronin B. Bias in peer review. JASIST 2013;64:2–17.

Moylan EC, Harold S, O'Neill C, Kowalczuk MK. Open, single-blind, double-blind: which peer review process do you prefer? BMC Pharmacol Toxicol 2014;15:55.

Okike K, Hug KT, Kocher MS, Leopold SS. Single-blind vs double-blind peer review in the setting of author prestige. JAMA 2016;316:1315–6.

Ross JS, Gross CP, Desai MM, Hong Y, Grant AO, Daniels SR, et al. Effect of blinded peer review on abstract acceptance. JAMA 2006;295:1675–80.

Rothwell PM, Martyn CN. Reproducibility of peer review in clinical neuroscience. Is agreement between reviewers any greater than would be expected by chance alone? Brain 2000;123:1964–9.

Scharschmidt BF, DeAmicis A, Bacchetti P, Held MJ. Chance, concurrence, and clustering. Analysis of reviewers' recommendations on 1,000 submissions to the Journal of Clinical Investigation. J Clin Invest 1994;93:1877–80.

Schroter S, Black N, Evans S, Godlee F, Osorio L, Smith R. What errors do peer reviewers detect, and does training improve their ability to detect them? J R Soc Med 2008;101:507–14.

Smith R. Peer review: a flawed process at the heart of science and journals. J R Soc Med 2006;99:178–82.

Tomkins A, Zhang M, Heavlin WD. Reviewer bias in single- versus double-blind peer review. Proc Natl Acad Sci USA 2017;114:12708–13.

Tvina A, Spellecy R, Palatnik A. Bias in the peer review process: can we do better? Obstet Gynecol 2019;133:1081–3.

Ucci MA, D'Antonio F, Berghella V. Double- vs single-blind peer review effect on acceptance rates: a systematic review and meta-analysis of randomized trials. Am J Obstet Gynecol MFM 2022;4:100645.

van Rooyen S, Godlee F, Evans S, Smith R, Black N. Effect of blinding and unmasking on the quality of peer review: a randomized trial. JAMA 1998;280:234–7.

Weissmann G. Science fraud: from patchwork mouse to patchwork data. FASEB J 2006;20:587–90.