

# And Who Will Review the Review(er)s?

*"The referee is the lynchpin about which the whole business of Science is pivoted."<sup>1</sup>*

*"As a human enterprise, peer review is inherently ideological: no amount of scientific training will completely mask the human impulses to partisanship."<sup>2</sup>*

Peer review maintains an implacable presence in the collaborative enterprise of scholarly production. Widely viewed as the "gold standard," it is considered a requirement for affirming validity and quality, as well as for codifying disciplinary boundaries. As journal editor Richard Smith vividly recalled: "It is the method by which grants are allocated, papers published, academics promoted, and Nobel prizes won. ... When something is peer reviewed it is in some sense blessed."<sup>3</sup>

But peer review is not without its detractors. Some note that peer review's monopoly on validation is not supported by research testifying to its efficacy (or lack thereof); others emphasize that, in their experience, the method is deeply flawed. Mario Biagioli observes a "... remarkable epistemological and symbolic burden placed on peer review" despite a deficiency of empirical or philosophical examination of the practice.<sup>4</sup> As an oft-quoted piece in *The Lancet* posited: "[W]e know that the system of peer review is biased, unjust, unaccountable, incomplete, easily fixed, often insulting, usually ignorant, occasionally foolish, and frequently wrong."<sup>5</sup> In spite of voluminous commentary on the flaws of the process and concerns over potential abuses, the weight and preference assigned to peer-reviewed work persists.<sup>6</sup>

How do peer reviewers—and the processes designed to channel their input or influence—figure into the essential, imperfect peer review method? This editorial will consider the peculiar role of referees in peer review and offer examples of systems and practices that have been implemented with the goal of guiding, developing, or honing the selection of reviewers.<sup>7</sup> Confusingly, "peer review" often refers to an overarching method as well as to the many permutations of systems and approaches installed to enact the method: the term elides real distinctions in application. Given the range of practices encompassed by the peer review method, and the varied disciplinary cultures and norms that additionally shape these practices, there is no expectation of coherence across peer review systems.

Attention to referees requires attention to mechanisms designed to channel, influence, or assess their efforts. Peer reviewers are typically constrained by the peer review systems in which they play a part. They are also guided by explicit requirements and unwritten norms of scholarly subcultures. The advent of electronic publishing and shifts related to open access, digital scholarship, and library publishing have prompted reflection on and adjustment of review practices. Looking at examples of these adjustments and experiments reveals mechanisms that have been enabled or simply systematized through digital means. Some of these experiments come from conferences or grant panels rather than journals. Some are related to open peer review. Some have been forged in interdisciplinary spaces that both assign special responsibilities for inclusive, experimental practices and offer a freedom from discipline-entrenched approaches. Some are necessitated by emerging forms of scholarship that don't align with traditional imprimaturs.

As many authors of peer review studies have observed, it is difficult to find compelling evidence or to make systematic or generalizable assertions about the efficacy of the peer review method. Peer review is a broadly defined, anchored approach that is

implemented in many small, localized ways. Attempts to assess peer review at scale often wither for lack of accessible data or nonspecific practices.

Librarians have a duty to scrutinize both the method and the implementation of peer review, given our canonization of peer review in our own systems of scholarly production and assessment and in our work to organize and ensure access to scholarly corpora. In our own systems, we have perpetuated peer review through institutional promotion and tenure policies that favor peer-reviewed research and through applying and supporting the process of peer review in our journals and conferences. In our management of scholarly corpora, we have even implemented discovery layers and designed instruction programs that preference peer-reviewed publications. We have developed library publishing programs that replicate and extend peer review. And, interestingly, we have built or supported open access preprint repositories that provide venues for publishing work prior to peer review—and that offer a comparative basis for analyzing the effects of peer review across a swath of pre- and post-review versions of publications.<sup>8</sup>

Awareness of these local implementations, and of the relationship between peer review as a method, writ large, and peer review as specifically implemented, may help inform several realms of our professional practice. But beyond greater awareness, we need better data on both existing practices and the results of experiments or shifts in peer review. As Lutz Bornmann observes: “The entire process [of peer review] eludes ethnographic observations.”<sup>9</sup> Critiques of peer review, Biagioli notes, unfold in “private conversations... [or] in the context of personal complaints about the perceived incompetence (or other unflattering traits) of editors or referees.”<sup>10</sup> To better assess implementations of peer review—including the influence of referees—these conversations need to move into a public realm.

---

Peer review persists despite what Ann C. Weller, the author of an exhaustive study on the topic, admits is an absence of generalizable practices, guidelines, or standards. As Weller elaborates: “Editors have great flexibility in their implementation of the [peer review] process in their journals. Peer review has been demonstrated to be different for different disciplines, different journals, and different editorships. There is not one solid, accepted definition of what constitutes ‘peer review.’”<sup>11</sup>

Perhaps peer review’s dominance in academia is, in fact, rooted in what Weller describes as a lack of definitional precision or consensus about its application. Such looseness allows us to bundle disparate practices together, glossing over real distinctions in what the term actually entails. This lack of specificity lays the groundwork for clumsy critiques of peer review, which generalize from specific applications and then make the case not for an overhaul of the system, but for adjustment and even extension of particular practices or implementations—that is, for more or different peer review.

Scholars of peer review, faced with evaluating its ambitions and failures, are fond of invoking the famous Winston Churchill quotation about democracy as “the worst form of government except for all those other forms...” The implication, of course, is that, for all of its perceived and actual faults, peer review is simply the best option available and will remain firmly entrenched in the academy. But another comparison presents itself: like democracy, peer review references a broad overarching approach, with opportunities to design and install systems of governance that balance out competing interests. And, like democracy, peer review incurs bias, abuse, and ethical lapses in its application and implementation.

Exemplifying this inclusive framing of peer review and espousing an approach to assessment as “a fluid genre of scholarship,” Korey Jackson argues that “Ultimately,

peer review's fluid past is a way of reframing the notion of peer review's seemingly revolutionary future." He writes:

Developments like the pre-print archive (notable examples include arXiv and bioRxiv) or the emergence of mega-journals like PeerJ and the Open Library of Humanities that emphasize open and participatory review, are not (or not simply) radical reactions to a failed review enterprise. They are instead permutations, bellwethers of the increasingly open and collaborative ways good scholarship gets done and wants to be counted.<sup>12</sup>

In a review of the literature, Lawrence Souder discerns that peer review standards are undergirded by the Mertonian norms of science: communalism, universalism, disinterestedness, originality, and skepticism.<sup>13</sup> The chronicled abuses of peer reviewers are framed as violations of these norms. Jackson's emphasis on continuity and improvement, his call for scholarly assessment practices to be subjected to "constant critique and continual updating," recalls arguments that ground peer review as an ideal, community enterprise, with reviewers playing an important, disinterested role in social knowledge creation.<sup>14</sup>

---

As the named agent in a process that is only defined locally, peer reviewers occupy roles that are impossible to universalize. But in a closed, editor-driven peer-reviewed scholarly publishing process, peer reviewers' roles are clearly constrained. Shielded by anonymity, their recommendations and comments are considered advisory to the editor, who has the discretion and authority to pass the reviews directly to authors, summarize comments, or simply discard them. Editors might also design requirements to serve the goals or values of the editorial peer review system espoused by the journal: they may encourage reviewers to provide constructive comments, set expectations that reviews recommending Revise and Resubmit be involved in subsequent assessments of resubmitted articles, or provide clear and formal guidelines focused on areas of concern.

Unsurprisingly, in this model, authority is clearly vested in the editor. As Weller qualifies, this constitutes a controlled role for referees, who operate within a defined governance structure: "... these opinions of reviewers are just that, and it is the editor who then adjudicates between the author's manuscript and the reviewers' opinions and makes a decision, thereby establishing a system of checks and balances."<sup>15</sup> Of course, editors, governed as they might be by editorial boards or existing journal standards as well as their own preferences, might ascribe more or less authority to reviewers.

The immediate answer to the question of "who will review the reviewers," in such a system, is evident: the editor will. In addition to responding to reviews, editors have often been known to maintain lists of reliable reviewers with relevant expertise.<sup>16</sup> With the advent of fully featured electronic journal management systems, these lists have, in some cases, become automated and rooted in tracking data, which record whether reviewers have submitted their reports on time, and may include ratings or notes from editors and, in some cases, authors on the quality of their reviews. Gary Marchionini has argued that editors' adoption of electronic journal management systems raises questions about utility, confidentiality, communal versus individual attribution. These systems have the potential to affect reviewers' behaviors, as well as to pose new concerns for the management of reviewer data:

The critical element of these systems is that the reviews themselves as well as these ratings are persistent, outliving the terms and memories of individual edi-

tors, thus reducing community memory to simple scales that persist beyond the memories and perceptions of individuals in a community. As they become more uniformly adopted and more sophisticated in scaling review contribution, the ratings and the reviews themselves become the basis for evaluation of scholarly productivity.<sup>17</sup>

As befits the anonymity and confidentiality that characterize this permutation of peer review, editors are typically the keepers of this data. Indeed, reviewers in such closed systems may never see the reports submitted by their peers, nor be folded into any assessment process that extends beyond the boundaries of their reports. Only editors—and, potentially, authors—are positioned to access perfect information about the review process.

Recently, compelling experiments with the openness of the review process have emerged. For example, the proposal process for the international Digital Humanities conference no longer maintains the expectation that only those making final determinations to accept, reject, or require revisions to a work have access to the evaluated work and its reviews. Beginning in 2012, program committee chairs Bethany Nowviskie and Melissa Terras spearheaded a set of reforms, many of which were aimed at improving the overall review process.<sup>18</sup> Some reforms had the effect of producing data for the program committee, to be used as the basis for weighing reviews; others further structured the review process for authors and reviewers alike, providing opportunities for authors to respond to reviews and for reviewers to report conflicts of interest, indicate the proposals they'd prefer to (or not to) review, and view other reviewers' comments. These experiments, largely aimed at the social changes of facilitating more evaluation and exchange, were also implemented electronically, through the conference management platform. The innovation allowing reviewers to view other reviewers' comments remains in place as of 2017. This feature presents referees who have submitted their reports with the option of accessing other referees' reports on the same proposals. Identities remain anonymized—only the content of the reports is accessible. Having viewed others' comments on the same proposals, referees are permitted to adjust their own reviews. As Nowviskie explained when publicly debuting the reforms:

We hope the sharing of good examples of thoughtful and constructive critique will increase reviewers' quality of engagement with the proposals and their cordiality to authors, and contribute to the fellow-feeling with which we all undertake the service of reviewing. To minimize any danger of group-think, we will ask reviewers who augment their comments after seeing others' to offer a thorough justification.<sup>19</sup>

Remarkably, the reforms to the Digital Humanities review process redistribute the authority to "review the reviewers." Suddenly, we've gone from conference organizers having the sole opportunity to review referees to this privilege being extended to authors and other referees. Some of these reforms formalize feedback from authors as accepted practice: authors may have previously had the opportunity to flag unhelpful, biased, mean-spirited, or ambiguous reviews, but prompting for such feedback, and designating a channel for ongoing communication, validates that exchange. Similarly, reviewers are typically free to indicate the strength of their familiarity with an area to editors, committee members, or grant officers, but prompting for such a self-assessment encourages and quantifies it. Reforms aimed at giving reviewers the option to engage with other reviewers' reports open up another avenue of access that initiate conversation across a group rather than up or down a hierarchy.

It is notable that such reforms were instituted for an international, interdisciplinary conference, whose organizers have sought to extend its reviewer pool to reflect this diversity of disciplinary affiliations and national cultures, and to manage assessments that sometimes veer into (inter)disciplinary boundary policing.<sup>20</sup> In her study of academic evaluation, focused on peer review panels for grants and awards, Michèle Lamont observes: “American higher education brings together disciplines that are remarkably different in their evaluative cultures, intellectual traditions, and professional language. Disciplinary norms are stronger in some fields than in others, because American academia is also multidimensional, traversed by networks and literatures that are not always bounded by disciplines.”<sup>21</sup> She discerns a competition among these disciplinary norms in debates over excellence between peer reviewers on multidisciplinary panels, where scholars ask “... whose criteria gets universalized as disciplinary criteria.”<sup>22</sup>

When evaluating multidisciplinary work, who is a “peer” qualified to assess?<sup>23</sup> Digital scholarship furnishes additional examples of efforts to match an emerging, multidisciplinary area of scholarly production with systems and practices for evaluation and validation.<sup>24</sup> These efforts have taken on greater urgency as scholars engaged with digital work have come up for promotion or tenure in academic departments accustomed to relying on publication venue—whether a university press monograph or an article in a well-regarded journal—as indicators of excellence. The absence of such imprimaturs for digital scholars has prompted a need to analyze and design assessment structures. As Susan Schreibman, Laura Mandell, and Stephen Olsen argue in their introduction to a special section of the Modern Language Association’s *Profession* journal: “... digital scholarship requires review by experts who can bring to bear not only field knowledge to evaluate the intellectual content of a project but also the technical experience to understand the intertwined theoretical and technical choices in a project’s design.”<sup>25</sup> Schreibman, Mandell, and Olsen reference Jerome McGann’s Networked Infrastructure for Nineteenth-Century Electronic Scholarship (NINES) project as pioneering a trend towards “area-specific peer-reviewing organizations” that deploy appropriate reviewers.<sup>26</sup>

---

The fluid definition of peer review and its dominance in research position the method to adapt and grow, spurred by editorial reflection, technological capabilities, and shifting forms of scholarship that interrogate or replicate existing practices. Peer review practices can adjust—incorporating, for example, peer review of data or open peer review—without repudiating the overarching method and necessarily disrupting its status.<sup>27</sup>

Even as a lack of specificity about peer review—and an accompanying flexibility in how it is administered—gives librarians license to experiment with its implementation, we have an obligation to scrutinize, meta-analyze, rebuke, and interrogate. Weller’s study, which I have referenced so frequently in this editorial, serves as an example of such an approach to thorough analysis and documentation, even as it largely endorses the practice of peer review. By promoting a more encompassing definition of peer review, Jackson provides a framework for incorporating alternative metrics and other systems of establishing the public worth of a work, approaches that can serve as a counterballast to more entrenched peer review practices.

The experiments and shifts highlighted here also suggest questions about openness, accountability, confidentiality, expertise, and the collaborative nature of scholarship. By enacting the potential flexibility of peer review, we can consider and evaluate what has often been presumed to be its essential character. Experimentations with peer

review, particularly those that turn the lens of assessment and validation to reviewers themselves, ultimately signal a commitment to perfectibility and extensibility in scholarly communication.

Sarah Potvin

Digital Scholarship Librarian and Associate Professor  
Texas A&M University Libraries

## Notes

1. John Michael Ziman, *Public Knowledge: The Social Dimension of Science* (Cambridge University Press, 1966), 148. Quoted in Harriet Zuckerman and Robert K. Merton, "Patterns of Evaluation in Science: Institutionalisation, Structure and Functions of the Referee System," *Minerva* 9, no. 1 (January 1971): 66–100, available online at <http://www.jstor.org/stable/41827004> [accessed 8 August 2017].

2. Lawrence Souder, "The Ethics of Scholarly Peer Review: A Review of the Literature," *Learned Publishing* 24, no. 1 (January 2011): 55–74, doi:10.1087/20110109.

3. Richard Smith, "Peer Review: A Flawed Process at the Heart of Science and Journals," *Journal of the Royal Society of Medicine* 99, no. 4 (April 2006): 178–182, available online at <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1420798/> [accessed 8 August 2017] doi:10.1258/jrsm.99.4.178.

4. Mario Biagioli, "From Book Censorship to Academic Peer Review," *Emergences* 12, no. 1 (2002): 11–45. doi:10.1080/104572202200000343 p. 11.

5. Souder, "The Ethics of Scholarly Peer Review."

6. In a review of studies of peer review Tom Jefferson et al. concluded that "the methodological problems in studying peer review are many and complex" and called for a large-scale study of its effects, observing in their limited study "little empirical evidence .... to support use of editorial peer review as a mechanism to ensure quality of empirical research." Souder details complaints about peer reviewers noted (and in some cases substantiated) in the literature: they are biased, violate ethical norms or commit fraud, plagiarize, breach confidentiality, etc. See Tom Jefferson, Melanie Rudin, Suzanne Brodney Folse, and Frank Davidoff, "Editorial Peer Review for Improving the Quality of Reports of Biomedical Studies," *Cochrane Database of Systematic Reviews* 2 (2007), available online at <https://doi.org/10.1002/14651858.MR000016.pub3> [accessed 8 August 2017]; Souder, "The Ethics of Scholarly Peer Review."

7. I use the terms "referee" and "reviewer" interchangeably.

8. Martin Klein et al's work examining papers in ArXiv posits that publications don't change significantly between pre- and post-print versions—an argument that peer review's effects are minimal. See Martin Klein, Peter Broadwell, Sharon E. Farb, and Todd Grappone, "Comparing Published Scientific Journal Articles to their Pre-print Versions," *Proceedings of the 16<sup>th</sup> ACM/IEEE-CS on Joint Conference on Digital Libraries* (Newark, NJ: June 19–23, 2016): 153–162. doi:10.1145/2910896.2910909. For a critique of a preprint version of Klein et al's analysis, see Dalmeet Singh Chawla, "Do Publishers Add Value? Maybe Little, Suggests Preprint Study of Preprints" *Retraction Watch* blog (June 24, 2016), available online at <http://retractionwatch.com/2016/06/24/do-publishers-add-value-maybe-little-suggests-preprint-study-of-preprints/> [accessed 8 August 2017] as well as comments on the blog post.

9. Lutz Bornmann, "Scientific Peer Review: An Analysis of the Peer Review Process from the Perspective of Sociology of Science Theories," *Human Architecture: Journal of the Sociology of Self-Knowledge* 6, no. 2, Article 3 (2008), available online at <http://scholarworks.umb.edu/human-architecture/vol6/iss2/3> [accessed 8 August 2017].

10. Biagioli, "From Book Censorship to Academic Peer Review," 11.

11. Ann C. Weller, *Editorial Peer Review: Its Strengths and Weaknesses* (Medford, NJ: Information Today / ASIS&T Monograph Series, 2001): 308–9.

12. Korey Jackson, "Watching the Detectives: Review's Past and Present," *Ada* 4 (April 2014). doi:10.7264/N38W3BK9.

13. Souder, "The Ethics of Scholarly Peer Review," 57.

14. In further commentary on the role of the reviewer, Harriet Zuckerman and Robert K. Merton frame referees and editors as "significant status-judges." They argue that status judges (a category that can be applied in other contexts to teachers and coaches, among others) "... are integral to any system of social control through their evaluation of role-performance and their allocation of rewards for that performance." There are, of course, other status-judges in the scholarly ecosystem, including authors and readers, whose assessments are formally incorporated via citations, reviews, comments, and even downloads, through scholarly impact metrics that play

a growing role in the continued assessment of scholarship; Harriet Zuckerman and Robert K. Merton, "Patterns of Evaluation in Science: Institutionalisation, Structure and Functions of the Referee System," *Minerva* 9, no. 1 (January 1971): 66–100, available online at <http://www.jstor.org/stable/41827004> [accessed 8 August 2017].

15. Weller, *Editorial Peer Review*, 322.

16. Ann C. Weller, "Editorial Peer Review: Research, Current Practices, and Implications for Librarians," *Serials Review* 21, no. 1 (1995): 56, available online at [https://doi.org/10.1016/0098-7913\(95\)90021-7](https://doi.org/10.1016/0098-7913(95)90021-7) [accessed 8 August 2017]. Stevan Harnad writes that "Editors usually have 'stables' of referees (an apt if unflattering term describing the workhorse duties this population performs gratis for the sake of the system of the whole) for each specialty; in active areas, however, these populations may be saturated—a given workhorse may be in the service of numerous stables." See Stevan Harnad, "Implementing Peer Review on the Net: Scientific Quality Control in Scholarly Electronic Journals," in *Scholarly Publishing: The Electronic Frontier*, eds. Robin P. Peek and Gregory B. Newby (Cambridge, MA: MIT Press, 1996): 103–118.

17. Gary Marchionini, "Editorial: Reviewer Merits and Review Control in an Age of Electronic Manuscript Management Systems," *ACM Transactions on Information Systems* 26, no. 4, Article 25 (September 2008), doi:10.1145/1402256.1402264. While writing this editorial, I heard some concern, voiced anecdotally rather than in the literature, that systems that enabling tracking and rating of reviewers further strain those reviewers tagged as strong and reliable. This concern is supported by what Souder describes as the commoditization of reviewers, referencing Tsui and Hollenbeck's work on the "reviewing market." See Souder, 59.

18. By way of disclosure: I served as a member of the program committee for the 2013 and 2014 Digital Humanities conferences.

19. Bethany Nowviskie, "Cats and Ships," *nowviskie.org blog* (November 2, 2012), available online at <http://nowviskie.org/2012/cats-and-ships/> [accessed 8 August 2017].

20. As 2017 program committee chair Diane Jakacki expressed in her proposal to extend the conference reviewer pool, adding new reviewers "better reflects and represents the dimension of scholar-practitioners in [Digital Humanities] whose work is presented at the conference in all inclusive senses (in terms of language, region, race, ethnicity, culture, labor, identity, as well as the ever-expanding types of scholarship, publication, and expression that are associated with [Digital Humanities]." Diane Jakacki, "Recommendation to extend reviewers pool for DH2017/18."

21. Michèle Lamont, *How Professors Think: Inside the Curious World of Academic Judgment* (Harvard University Press, 2009): 102–3.

22. Lamont, 103. Souder reports that "When the peer-review process becomes interdisciplinary, some scholars have discovered epistemological conflicts of interest." See Souder, 62.

23. Research firmly situated in a discipline, too, prompts concern for journal editors seeking to identifier reviewers that qualify as peers: as Smith has asked: "But who is a peer? Somebody doing exactly the same kind of research (in which case he or she is probably a direct competitor)? Somebody in the same discipline? Somebody who is an expert on methodology?" See Smith, n.p.

24. Several disciplinary associations have issued guidelines for evaluating digital scholarship. See the Modern Language Association's *Guidelines for Evaluating Work in Digital Humanities and Digital Media* (Committee on Information Technology, 2012) and the American Historical Association's *Guidelines for the Professional Evaluation of Digital Scholarship in History* (Ad Hoc Committee on the Evaluation of Digital Scholarship by Historians, June 2015).

25. Susan Schreibman, Laura Mandell, and Stephen Olsen, "Evaluating Digital Scholarship," *Profession* (2011): 123–135, available online at <http://www.jstor.org/stable/41714114> [accessed 8 August 2017].

26. *Ibid.*, 124–5.

27. See other editorials in *C&RL's* series on peer review for analysis of open peer review and peer review of data. Morten Wendelbo, "Perspectives on Peer Review of Data: Framing Standards and Questions," *College & Research Libraries* 78, no. 3 (April 2017), available online at <https://doi.org/10.5860/crl.78.3.16585> [accessed 8 August 2017]; Emily Ford, "Advancing an Open Ethos with Open Peer Review," *College & Research Libraries* 78, no. 4 (May 2017), available online at <https://doi.org/10.5860/crl.78.4.406> [accessed 8 August 2017].