
REVIEW CRITERIA FOR RESEARCH MANUSCRIPTS

A Tool for Reviewers: “Review Criteria for Research Manuscripts”

Georges Bordage, MD, PhD, and Addeane S. Caellegh

Peer review lies at the core of science and academic life. In one of its most pervasive forms, peer review for the scientific literature is the main mechanism that research journals use to assess quality. Editors rely on their review systems to inform the choices they must make from among the many manuscripts competing for the few places available for published papers. In the past 50 years, the use of peer review has become the “gold standard” by which journals are judged, just as journals use it to judge papers. And whereas journals in all branches of science share the core ethos and values of peer review, it has evolved in diverse ways to best fit the environments and circumstances of the various sciences and disciplines.

However, our purpose in this task force report, “Review Criteria for Research Manuscripts,” is not to discuss the general nature and permutations of peer review, as important as those topics are. Others have already done this thoughtfully and well.¹⁻⁴ Their work has focused on the tensions inherent in the peer review process, the state of peer review and major changes in it, particularly over the last 20 years, the development of data derived from research on peer review, and specific areas of contention and ethics raised by the conduct of peer review. Our intention, in contrast, is to contribute to the practice of review and develop a scholarly resource for reviewers to use as they review manuscripts.

Both review and reviewers are often misunderstood by authors and the reviewers themselves. Authors often feel that decisions about their manuscripts are based on mysterious criteria and standards, in a largely secretive process run by editors and unknown reviewers. Their concerns about the opacity of review processes are confirmed; Colaianni found that fewer than half of the journals in her sample of journals from four subject fields actually included clear statements about their peer review practices.⁵ Reviewers, too, are handcuffed by a lack of information; they are usually told little if anything about their role in how decisions are made about journal articles or about what is expected of them.⁶

“Review Criteria for Research Manuscripts” grew from an effort to address the need for more information about review systems and reviewing in the medical education research community. By forming a task force to concentrate on the needs of reviewers, we hoped to develop, sort, and present information that would, in turn, help to increase the quality of peer review that members of this community provide to journals and to one another. To meet this need, the task force focused on the core issues: Who needs information most, and what information do they most need? The trajectories of our answers to these questions (developed through a normative group process) crossed at reviewers and criteria, and what we have produced is a reference tool for reviewers to use when they receive research manuscripts that they have been asked to review.

BACKGROUND

When grappling with what information was needed and who needed it, we could not ignore how perceptions of and attitudes toward peer review have changed over recent decades and among different research communities. Further, these changes have varied from discipline to discipline, field to field, science to science. Peer review was originally conceived to provide advice for the editor, the equivalent of asking the knowledgeable colleague down the hall for an opinion. By the 1960s and 1970s, however, it had come to be the measure of quality for journals—high-quality journals use strong peer-review systems. When the National Library of Medicine created *Index Medicus* in the 1960s, peer review was not a requirement for a journal's inclusion, but it was a highly weighted factor, as remains the case today. As scholarly publication flourished, particularly in the sciences, and hundreds of new journals emerged, the expectation was that these journals would be founded on the practice of peer review, and the practice was solidified.

The spread of peer review and its adoption as the standard

of quality brought with them, however, ethical and other problems that challenge the conduct normal in peer review systems. A few widely known cases of fraud and misconduct (particularly the Darsee⁷ and Slutsky⁸ ones) that came to light in the 1980s illustrated starkly the problems with authorship, duplicate publication, and other publication misconduct that many editors had been concerned and frustrated about for years. In 1978 the editors of ten internationally prominent medical journals formed a group to begin cooperative work on common problems that affected journals. Originally called the Vancouver Group (after the site of their first meeting), the group soon took more formal shape and status, becoming the International Committee of Medical Journal Editors (ICMJE). The group has become increasingly important over the past 20 years, meeting each year and periodically issuing consensus statements, which hundreds of other journals voluntarily sign on to. Several of the statements deal indirectly, and some directly, with peer review.⁹

In seeking to understand and improve peer review, editors in biomedicine had more questions than answers, however. Stephen Lock's pivotal book, *A Delicate Balance: Editorial Peer Review in Medicine*,² presented a systematic look at peer review, bringing together the whole body of relevant research across the sciences. Then in 1989, the American Medical Association sponsored the First International Congress on Peer Review in Biomedical Publication, and JAMA published the proceedings in a special issue with the evocative title, "Guarding the Guardians: Research on Editorial Peer Review."¹⁰ Two other conferences followed, in Chicago in 1993 and in Prague in 1997, each with a proceedings in JAMA,^{11,12} and a third conference is scheduled for Barcelona in September 2001. The emphasis of these meetings is research on peer review and other issues important in bioscience journals; the importance of creating a community and forum for the presentation of research on peer review can not be overstated.

STATUS OF RESEARCH ON REVIEWING AND REVIEWERS

The research by the bioscience editors is not the only research on peer review, although it has come to dominate in the past decade. Parallel work has been done in psychology, sociology, economics, and other fields. Taken together, this knowledge illuminates many aspects of review and provides increasing evidence that editors need to support their review systems or changes to them. In particular, two decades of research have deepened the understanding of reviewing and reviewers.

- Three overviews at different times and from different perspectives have summarized what is known from research.

The first, based in the social sciences, was Armstrong's 1982 article¹³ that reviewed research on science journals and the editorial policies of leading journals, and then presented the implications and his recommendations for improvements. The second was Lock's 1985 book,² already mentioned. The third, which is very recent, is the systematic review by Overbeke¹⁴; it is the best present summary source and introduction to studies in the area.

- Research into the kinds of reviewers who do better reviews for editors, that is, the types of reviews that editors value, has produced contradictory results so far. A 1993 study¹⁵ found fairly strong evidence that good peer reviewers tended to be under age 40, were from top-ranked academic institutions, were well known to the editor, and were blinded to the identity of the paper's author. A 1998 study,¹⁶ on the other hand, was not able to identify the characteristics of good reviewers. The closest findings, very weak, were that reviewers between ages 40 and 60 did better reviews than did those over age 60, and also that reviewers educated in North America and trained in epidemiology or statistics did better reviews. These two studies were done at medical journals; there is not a parallel body of research for the social science journals.
- Studies in the biomedical sciences and social sciences over the past decade produced mixed findings about using a masked review system, also called a double-blinded system. In a masked system, the reviewer does not know the identity of the author or institution. This is in addition to the customary practice of concealing the identity of the reviewer from the author. Studies in economics journals produced strong support for using masked review.¹⁷⁻¹⁹ But similar studies of review in the bioscience journals have been more mixed, although evidence is firming on some issues. Although earlier research had indicated otherwise, two randomized controlled trials in the 1990s found that masking made no difference in the quality of the reviews of papers at prestigious biomedical journals.^{20,21} Likewise, open peer review, where the identities of the author and reviewer are known to each other, apparently did not affect the quality of reviews.²² Nonetheless, this issue is debated strongly among editors, and more research is needed at the few journals that have open peer review.
- Reviewers have to respond to the widely varied expectations and procedures of the journals that ask them for reviews, because the journals have different ways of obtaining information from them.⁶
- Journals have been able to develop validated assessment instruments to evaluate reviewers' performances.²²
- Men and women may behave somewhat differently as reviewers. For example, a 1990 study reported that women reviewers accepted three times more articles by women authors than by men authors, where male reviewers ac-

cepted equal proportions.²³ And a 1994 study found differences between men and women in several review activities.^{24,25}

- Reviewers may react to papers differently, depending on the content. Again, the findings are mixed. On the one hand, a study found that reviewers seemed to favor results that support the status quo.²⁶ In another study, reviewers did not react differently to content.²⁵

Without doubt, the international congresses have focused attention on peer review and dramatically increased the available research. Editors are working to understand their systems better and to make evidence-based improvements to their peer review.

WHAT IS NEEDED

Regrettably, the increase in research on peer review has not been accompanied by more teaching of peer review. Reviewers receive very little preparation for performing reviews as part of their formal education. Nor do they receive it one-on-one from mentors, overworked attendings, dissertation supervisors, or lab chiefs. For, despite faculty's and trainees' continuing belief that a mentor working one-on-one with a student provides the best education, the reality is that few trainees receive any help from faculty (mentors or otherwise) in such areas as reviewing, ethics, and writing for publication. Eastwood points out that "many such relationships have deteriorated to the extent that faculty regard responsibilities such as peer review and the writing of book chapters as independent opportunities for trainees" and that the situation has serious implications for the development of reviewers who are cognizant of the responsibilities of review.²⁷

New reviewers may be experts in particular fields, but each one is at some time a complete novice as a reviewer. Most new reviewers have seen reviews because they have submitted their own papers to journals. Therefore, in a left-handed way, they know about good reviews and bad reviews, but from the author's viewpoint. To review a paper for a journal requires a different viewpoint, however. The reviewer is expected to apply a set of outside criteria and standards to the paper, to write constructive criticism for the author, to write a critique and make critical judgments that will aid the editor in making decisions, and to accept a set of ethical responsibilities in relation to these activities. And unless this newly invited reviewer is fortunate enough to have a helpful mentor to turn to for guidance, these tasks must be undertaken without training or advice.

"Review Criteria for Research Manuscripts," then, serves two purposes and two audiences: to reaffirm the rules for reviewing by developing criteria for reviewers who are early in their careers, and to help more experienced reviewers re-

fresh their memories or unlearn bad habits. "Review Criteria for Research Manuscripts" was written to bring together a set of criteria for reviewing research manuscripts. If reviewers feel uncomfortable—because they are new to reviewing, or because of concern that they are not doing reviewing well—"Review Criteria" can be a resource. Further, "Review Criteria for Research Manuscripts" orients the novice to how journals work, the review process, and relevant ethical issues.

Regardless of the reviewer's experience, training in review benefits the journals and the scientific community. Just as the ICMJE took shape in the 1980s to confront common problems then, other groups began to form in the latter 1990s to deal with different as well as familiar issues. The World Association of Medical Editors was formed in 1995 to improve the quality of medical journals, particularly ones with limited resources and away from the centers of medical publishing. It conducts its activities over the Internet and now has approximately 500 members internationally. Among its early tasks was a statement of principles of professionalism and responsibilities of editors, including principles for the review process and reviewers.²⁸ The most recently formed is COPE (Committee on Publication Ethics), which began in London in the spring of 1997 as a small group of editors who met to discuss ethical problems the editors faced and must resolve, including those inherent in peer review.²⁹

It is also important to consider the effect that reducing reviewer bias would have not only on the careers of individual researchers but also on groups of researchers or their fields of study, such as theories or methods of research. As Beyer summarized the situation 25 years ago, even a small proportion of biased decisions would over time give some groups or individuals a large cumulative advantage. This is so because the bias would affect publication, which is used as the measure of merit upon which further promotion and advantage are based.³⁰

Some publications have dealt with the training of reviewers. Over the years, many editors have written short articles in their journals offering advice about reviewing. These can be especially helpful because they offer both general guidance and advice specific to particular research communities. The National Research Council of Canada, which publishes 15 journals, developed a document (to our knowledge the first) that summarized the responsibilities of authors, editors, and reviewers.³¹ It remains an excellent resource but is little known. Most recently, Godlee and Jefferson's book⁴ contains instructional essays on reviewing, such as Moher and Jadad's "How to Peer Review a Manuscript,"³² Altman and Schulz's "Statistical Peer Review,"³³ and Demicheli and Hutton's "Peer Review of Economic Submissions."³⁴ The first is framed as tips that are "the result of our combined experience as peer reviewers for some 30 journals." It is indeed a

useful tool, and it summarizes relevant research well; however, it focuses on generic aspects and has limited background information. The others use lists of criteria, a format much the same as the one in the "Review Criteria." Because statistics are an important part of analysis in many areas of science, the Altman and Schultz lists have many of the same criteria as does the "Review Criteria." (This is reassuring for all, since the task force derived its criteria lists independently through a normative group process before the Godlee and Jefferson book was available to them.) The Demicheli and Hutton list also has overlaps with the "Review Criteria"'s lists, although, as would be expected, it also has items specific to economics.

Researchers in some fields have used consensus conferences to develop specific guidelines for reporting particular types of research. An example is the CONSORT (consolidated standards of reporting trials) statement and its 21-item list and flow chart for presenting the fundamental information necessary to accurately evaluate the internal and external validity of a randomized controlled trial.^{35,36} Two other examples are known. The 1999 QUOROM (Quality of Reporting of Meta-analyses) statement is similar to the CONSORT in that it has a statement, itemized list, and flow chart,³⁷ and in 2000 a consensus group issued the MOOSE checklist, summarizing recommendations for reporting meta-analyses of observational studies in epidemiology.³⁸

The task force that prepared this document dealt with reviewing rather than reporting, and it chose to present the core of its recommendations as criteria. Because reviewing is about assessing and making judgments, reviewers are applying different criteria and standards at different times and places. In formal protocols or by informal consensus, a research community can agree upon the standards for research in a particular field. But often there is no consensus. And in some fields, especially cross-disciplinary or newly emerging ones, there may be no agreement as to the criteria, let alone the standards, to be applied.

The existence of the task force and these review criteria are a statement that there are basic criteria for research reports that cut across disciplines and fields. Through the review criteria, the task force has taken the position that good review is about making careful, systematic judgments, about assessing strengths and pointing out weaknesses, and about making global assessments if requested by the editor. We want to be clear about an important point, however: laying out these criteria is entirely separate from setting the standards to be used in making decisions about which papers a journal will publish. The criteria are the elements that reviewers should consider in preparing a review. And reviewers will apply their own professional standards as they use the criteria to prepare a review for the editor. But it is the duty of the journal, based on the journal's mission, traditions, and

circumstances, to set the standards to be used for different categories of papers, and perhaps these will differ at different times in the journal's development. Just as any assessment system must determine how good "good enough" must be, journals must set the standards that will allow them not only to separate the good papers from the poor ones, but also to choose which few good papers to publish from among many good papers. This requires setting standards, and that is wholly the province of journals.

WHO CAN BENEFIT?

"Review Criteria for Research Manuscripts" was written as a reference tool for a wide community of reviewers. Because *Academic Medicine* and GEA—RIME are from the health professions, and more specifically the branch for medical education research, the resulting document is set in the research framework most common to its members. More often than not, the examples given come from the social and behavioral sciences that are the framework of the majority of their research. "Review Criteria" has cast its criteria in the most general, widely applicable form precisely so that they can be used by researchers working in the widest possible range of sciences and disciplines. We think they apply in areas beyond the social and behavioral sciences, perhaps in some areas of the biosciences and physical sciences. The greater the distance from these sciences, however, the more the users may need to adapt parts of the criteria or add to them in order to make them work well for the research and traditions of their own fields.

By extension, researchers can also use "Review Criteria for Research Manuscripts" to plan and conduct their studies and to write papers properly prepared to compete for publication in peer-reviewed journals and to truly contribute to the scientific literature. By a further extension, the document can be used as a resource for faculty development to train authors and reviewers. Finally, editors may find it useful, perhaps to help in creating or revising review forms or in working with other editors.

No matter how important they may be, however, all these additional groups are secondary to the original audience—reviewers. The Task Force wrote "Review Criteria for Research Manuscripts" for them.

REFERENCES

1. Freese L. On changing some role relationships in the editorial review process. *Am Sociologist*. 1979;14(Nov):231–8.
2. Lock S. *A Difficult Balance: Editorial Peer Review in Medicine*. Philadelphia, PA: ISI Press, 1985.
3. Chubin DE, Hackett EJ. Peer review and the printed word. In: *Peerless Science: Peer Review and U.S. Science Policy*. Albany, NY: State University of New York Press, 1990.

4. Godlee F, Jefferson T (eds). *Peer Review in Health Sciences*. London, U.K.: BMJ Books, 1999.
5. Colaizzi LA. Peer review in journals indexed in *Index Medicus*. JAMA. 1994;272:156-8.
6. Frank E. Editors' requests of peer reviewers: a study and a proposal. *Prevent Med*. 1996;25:102-4.
7. Stewart WW, Feder N. The integrity of the scientific literature. *Nature*. 1987;325:207-14.
8. Locke R. Another damned by publications. *Nature*. 1986;324:401.
9. International Committee of Medical Journal Editors. Uniform Requirements for Manuscripts Submitted to Biomedical Journals [and Separate Statements]. *Ann Intern Med*. 1997;126:36-47; (www.acponline.org/journals/annals/01janr97/unifreq) (updated May 1999).
10. Guarding the guardians: research on editorial peer review. JAMA. 1990;263:1309-456.
11. The Second International Congress on Peer Review in Biomedical Publication. JAMA. 1994;272:79-174.
12. The Third Congress on Biomedical Peer Review. JAMA. 1998;280:203-306.
13. Armstrong JS. Research on scientific journals: implications for editors and authors. *J Forecasting*. 1982;1:83-104.
14. Overbeke J. The state of evidence: what we known and what we don't know about journal peer review. In: Godlee F, Jefferson T (eds). *Peer Review in the Health Sciences*. London, U.K.: BMJ Press, 1999:32-44.
15. Evans AT, McNutt RA, Fletcher SW, Fletcher RH. The characteristics of peer reviewers who produce good-quality reviews. *J Gen Intern Med*. 1993;8:422-8.
16. Black N, Van Rooyen, Godlee F, Smith R, Evans S. What makes a good reviewer and a good review for a general medical journal? JAMA. 1998;280:231-3.
17. 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.
18. Lebrand DN, Piette MJ. Does the blindness of peer review influence manuscript selection efficiency? *Southern Econ J*. 1994;60:896-906.
19. Lebrand DN, Piette MJ. A citation analysis of the impact of blinded peer review. JAMA. 1994;272:147-9.
20. Van Rooyen S, Godlee F, Evans S, et al. Effect of blinding and unmasking on the quality of peer review. JAMA. 1998;280:240-2.
21. Justice AC, Cho MK, Winker MA, Berlin JA, Rennie D, PEER investigators. Does masking author identity improve peer review quality? a randomized controlled trial. JAMA. 1998;280:240-2.
22. Fruer ID, Becker GJ, Picus D, Ramirez E, Darcy MD, Hicks ME. Evaluating peer reviews: pilot testing of a grading instrument. JAMA. 1994;272:117-9.
23. Lloyd ME. Gender factors in reviewer recommendations for manuscript publication. *J Appl Behav Anal*. 1990;23:539-43.
24. Gilbert JR, Williams ES, Lundberg GD. Is there gender bias in JAMA's peer review process? JAMA. 1994;272:139-42.
25. Caelleigh AS, Hojat M, Steinecke A, Gonnella JS. Effects of reviewers' gender on assessments of a gender-related standardized manuscript, 2001 [unpublished].
26. Mahoney MJ. Publication prejudices: an experimental study of confirmatory bias in the peer review system. *Cogn Ther Res*. 1977;1:161-75.
27. Eastwood S. Ethical issues in biomedical publication. In: Jones AH, McLellan F (eds). *Ethical Issues in Biomedical Publication*. Baltimore, MD: Johns Hopkins University Press, 2000:251.
28. World Association of Medical Editors. (www.wame.org). Accessed 6/1/01.
29. Guidelines on Good Publication Practice. The COPE Report 1999. (www.bmjpg.com/publicationethics/cope/cope1999). Accessed 6/1/01.
30. Beyer JM. Editorial policies and practices among leading journals in four scientific fields. *Sociol Q*. 1978;19:68-88.
31. National Research Press. Part 4: Responsibilities. In: *Publication Policy* (http://www.monographs.nrc.ca/cgi-bin/cisti/journals/rp/rpz_cust_e?pubpolicy). Accessed 6/5/01.
32. Moher D, Jadad AR. How to peer review a manuscript. In: Godlee F, Jefferson T (eds). *Peer Review in Health Sciences*. London, U.K.: BMJ Books, 1999.
33. Altman DG, Schulz KF. Statistical peer review. In: Godlee F, Jefferson T (eds). *Peer Review in Health Sciences*. London, U.K.: BMJ Books, 1999.
34. Demicheli V, Hutton J. Peer review of economical submissions. In: Godlee F, Jefferson T (eds). *Peer Review in Health Sciences*. London, U.K.: BMJ Books, 1999.
35. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. JAMA. 1996;276:637-9.
36. Rennie D. CONSORT revised—improving the reporting of randomized trials. JAMA. 2001;285:2006-7.
37. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of Observational Studies in Epidemiology (MOOSE) group. JAMA. 2000;283:2008-12.
38. Moher D, Cook DJ, Eastwood S, et al. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of reporting of meta-analyses. *Lancet*. 1999;354(1993):1896-900.