

# How We Review a Medical Education Research Manuscript

Margaret M. Hayes<sup>1</sup> and Henry E. Fessler<sup>2</sup>

<sup>1</sup>Division of Pulmonary, Critical Care and Sleep Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts; and <sup>2</sup>Division of Pulmonary and Critical Care Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland

## ABSTRACT

Peer review is a necessary and important component of scholarly publication. When done well, it benefits both the reviewer and authors and improves the science itself. However, the skills of effective peer review are rarely taught. In the adolescent field of medical education research, peer review is especially important to advance the scientific rigor of the field. From our experience reviewing biomedical and medical education research, we have found that a thorough review takes multiple readings and multiple hours. The first reading provides a general overview of the aims and methods. Subsequent readings focus on the details of the methodology, results, and interpretation. The written review should provide firm but gentle feedback that the authors can use to improve their work, even if we have recommended rejection for this submission. We hope that this description of our process for reviewing a medical education research manuscript will assist others and thereby advance the quality of publications in our field.

### Keywords:

education; peer review; scholarship

Peer review has been subject to criticism for the length of time it takes and its potential for bias and inconsistency (1). However, done well, it can ensure clarity, methodological rigor, and scientific

integrity (2). For these reasons, we consider it our duty to review, to enhance, and to advance the science in our field (3–5). Peer review is often considered thankless. However, in addition to helping

(Received in original form July 28, 2021; accepted in final form October 13, 2021)

This article is open access and distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives License 4.0. For commercial usage and reprints, please e-mail Diane Gern.

**Correspondence and requests for reprints should be addressed to** Margaret M. Hayes, M.D., A.T.S.F., Division of Pulmonary, Critical Care, and Sleep Medicine, Department of Medicine, Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Boston, MA 02215.  
E-mail: Mhayes7@bidmc.harvard.edu.

ATS Scholar Vol 3, Iss 1, pp 38–47, 2022  
Copyright © 2022 by the American Thoracic Society  
DOI: 10.34197/ats-scholar.2021-0105PS

the authors and their science, peer review also benefits the reviewer. Reviewing a manuscript, like writing one, requires content expertise and some knowledge of experimental design, statistics, strengths and weaknesses of techniques, and open mindedness to the results. For topics at the fringes of the reviewer's expertise, it can motivate an independent exploration of the literature and help the reviewer learn more (6). Careful attention to the authors' writing craft can improve the reviewer's own. And although reviewers are often anonymous to the authors, they are known to the editors. Thoughtful, constructive reviews are noticed and may lead to invitations to provide an editorial, editorial board membership, or other opportunities. For better or worse, they are almost certain to generate more invitations to review.

Following the growth of clinician-educator training and career tracks (7, 8), medical education research has grown apace (9). The number of medical education research manuscripts both submitted and published has increased, likely owing to, in part, an increase in journals accepting educational research (6). *ATS Scholar*, launched less than 2 years ago, is one such example. However, the lingering reputation of medical education research is that it is less methodologically rigorous than biomedical research. To dispel that misconception, many medical education scholars are working to elevate the quality of educational science (10). Journal editors are setting standards equally high for medical education research as for other forms of biomedical research. Reviewers of this work should do the same.

Despite these benefits and the importance of peer review to the scientific community at large, peer review is not intuitive and is rarely taught during training (11). In addition, many reviewers are less familiar

with medical education methodology than they are with basic or clinical research, making medical education research more challenging to evaluate. We will describe our approach to reviewing an education research paper in the hope that it will encourage and facilitate others to take on this vital and rewarding task. Although peer review has several formats, we will limit our consideration to the single blind format (authors do not know the reviewers' identities) used by the journals of the American Thoracic Society and many others.

## GET THE BIG PICTURE

Reviews take time and deep thought. Once you have accepted a review, commit the time to it. You owe it to yourself, to the authors, and to the editors. When approaching a new review, we first just read the paper from start to finish in a quiet place free from distractions. On the first pass, we don't take notes, just read. We liken this to immersing ourselves in the data (12) when analyzing a qualitative project. For this first reading, we try to absorb the main message without being distracted by details. We may consider whether the manuscript meets the minimum requirement for scholarship as described by the five questions below (9, 13):

1. Is there a clearly defined research question?
2. Is there an adequate literature review that puts the question into context?
3. Do the chosen methods match the question?
4. Are the results meaningful?
5. Are the results translated and contextualized?

Many papers will generate affective responses on this first reading that can impair our objectivity. Studies that seem to refute accepted concepts or contradict our own work or biases may evoke a negative reaction. Those that confirm our prejudices or build from our own work

can have the opposite effect. We try to be sensitive to these responses and to remain open-minded to the facts of the paper under review. Our knowledge of the authors, our opinion of their other publications, and unconscious biases based on the demographics or identities of the authors are factors that we try to make conscious to minimize their impact on the current review. Another potential source of bias during this first pass is the writing style. Poor writing takes many forms but can place the authors at a great disadvantage. A poorly written manuscript can be frustrating to read and can obscure the authors' message. Although the writing can be fixed in revision, the science cannot. We make every attempt to extract the core message and key information in this big-picture reading despite an author's literary style. Then we return for a deeper look.

### **SWEAT THE DETAILS**

After getting the big picture, we read the manuscript again—often several times. Although this may be a generational relic, printing the manuscript and reading it on paper improves our focus and has been shown to improve comprehension (14). This time, we take notes. We may pause to look up a reference or independently check some literature. During this second look, we deconstruct the paper and consider it section by section, paragraph by paragraph, and even sentence by sentence.

#### **Introduction/Background**

The introduction should be concise but compelling: concise because the introduction is just the overture and compelling because we want to be excited about the symphony to follow. The authors need to present their elevator pitch—tell us why their work is important,

what is known on the topic, what the gaps are, and how they will fill those gaps. The best papers are both interesting and important. Papers may be publishable if they are just one or the other.

(Importance may be hard to prognosticate.) However, if, by the end of the introduction, the authors cannot convince us that their work is either, they have a steep climb ahead.

Unlike most biomedical research, educational research falls within the social and behavioral sciences. The introduction should note whether the project is grounded in or linked to an underlying educational conceptual framework, a theory that describes social and educational phenomena (15). Such conceptual frameworks, such as Knowles' self-directed learning (16), Ericsson's deliberate practice (17), and Kolb's experiential learning (18), provide context and can highlight the importance of the study (15). The lack of a conceptual framework has been cited as a barrier to medical education research publications in top tier journals (19, 20). The lack of a framework may not be a fatal flaw, but its absence warrants a suggestion in our review.

#### **Methods**

All research studies will live or die on their methods, no matter how interesting or important they seem in their introduction. The chosen methods must answer the research question. They should be valid, reliable, and plainly described (11). It is crucial that authors follow the appropriate guidelines when conducting and writing up their study. The Equator Network (<https://www.equator-network.org/>) is a helpful website that includes reporting standards for different types of studies. Many medical education research manuscripts use mixed methods—the

collection and analysis of both quantitative and qualitative data (21, 22). We consider whether the mixed methods are warranted or just trendy. The authors should describe why they used mixed methods, and there should be a comment on how the methods were used—in an explanatory sequential manner, exploratory sequential, in parallel, or imbedded (23, 24). It is important to note if one method dominated or led to confounding of the other.

For quantitative methodology, we consider whether appropriate statistics were chosen and are sufficiently well described. If we are unsure and concerned, we note that in our review and may request secondary review by a statistician. Often, the number of subjects is a convenience sample based on the size of a class or group, and thus a seemingly small “*N*” is frequently justified in medical education research. However, if a preplanned number of subjects were recruited to a study, a power calculation should be included.

For qualitative methods, the specific methodology and its justification must be present (25). A discussion of why the methodology can be trusted, the concept of trustworthiness (26), is also needed and should be commented on in the methods. The concepts of internal validity, generalizability, and reliability for quantitative methodology are intuitive for most, but their counterparts in qualitative methods, credibility, transferability, dependability, confirmability, and reflexivity should be commented on in the methods (27).

Many education research papers describe creation and evaluation of a curriculum. We look to ensure that this was done according to best practices of curriculum design (28). One of the most important steps is a rigorous needs assessment, which

has been defined as a “systemic process to collect and analyze information on what a target group needs to learn” (29). This process often includes surveys and focus groups. Another important step is an effort to evaluate the effectiveness of the curriculum, as described next.

We consider the study’s chosen outcomes, which often will determine the overall importance of the study. Kirkpatrick’s levels of learning evaluation (30) provide a useful yardstick, somewhat analogous to the various phases of clinical trials. Kirkpatrick’s lowest level is reaction, that is, did the learners like or dislike the curriculum, how did it make them feel, and did it improve their confidence or comfort level? This level is usually based on learner self-assessments, which can be less accurate. The next level is learning, which tests knowledge acquisition. Level three is a behavior change: do the learners do something differently based on their new knowledge? Level four is “results” or actual patient outcomes and is rarely achieved in educational research. If the study shows only the learner’s “reaction” to the curriculum, their belief that they learned something, this typically will not justify publication unless the curriculum or the study is truly novel. The higher the level on Kirkpatrick’s pyramid (30), the more impactful the outcome.

Surveys and pre-/postexams are frequently used in medical education research. Proper design is essential for their results to be valid (31). For both surveys and questionnaires, we like to see that the questions were developed rigorously according to best practices, including conducting literature reviews, creating questions through expert consensus process or based on previously validated examples, and conducting pretesting and pilot testing (32, 33).

Surveys should ideally be included as supplemental material, and if not, we may request them, especially if we have concerns about validity. Questionnaires given before and after a teaching session or curriculum should not use the same questions, but the questions should be of equal difficulty. Ideally, the exams will have had some attempt at external validation, such as showing that experienced learners do better than novices. Formal calculation of exam reliability statistics is another plus. We note the duration of time between the teaching of new material and its testing to distinguish between learning (immediate) and retention (longer).

The methods should be reported in sufficient detail to be reproducible should a reader desire, and methods that do not require specialized equipment or expertise available only at the authors' institution improve the value of the findings for the educator community.

### Results

This section should just present the findings without editorializing. Data can be provided in narrative text, tables, or figures as appropriate, but the same data should be presented just once. The experimental subjects should be described in sufficient detail so that we can understand to what types of learners the findings may or may not apply. Results that are not mentioned in the methods or do not address the experimental questions proposed in the introduction should not be included in the body of the paper. They could be included in supplemental materials. If they are not, a reviewer could suggest that they are and ask to see them. Figures should be visually concise—complete but without uninformative decorative flourishes such as three-

dimensional bars or gradient fills, also known as “chart junk” (34). We verify that the results haven't been published elsewhere with a quick literature search on the authors. It is perfectly acceptable to submit a study of the authors' previously described curriculum as long as it is appropriately cited.

### Discussion

The discussion should put the results into context and explain to the reader why this study matters and how it advances the field. It may open with a concise summary of the results (11) but should not duplicate them. The discussion should explain how the study answered, or failed to answer, the questions posed in the introduction. It should have no more than minimal speculation beyond the actual findings nor draw inferences from differences that did not meet statistical significance. The reviewer must ensure that statements are justified and supported by the study results in the context of the current literature.

If a conceptual framework was described in the introduction, the results should be fit to this scaffold. If there is published literature with opposite findings, the authors should acknowledge them and suggest reasons for the discordance. A comment on why this particular study is novel or useful is also important, especially if the results are low Kirkpatrick levels. All methods have shortcomings. The discussion must address them, preferably in a clearly labeled “limitations” paragraph.

The literature cited should be relevant and mostly recent. If there are controversies in the field, the authors should acknowledge them and their supporting publications. If there are relevant papers missing, we will point this out to the authors and suggest they are

included. We sometimes read or reread the cited literature ourselves and have found that they don't always support the authors' statement.

### Conclusion

The conclusion should be three sentences or less without extrapolation beyond the results. Because papers that answer one question often raise new ones, one of these sentences may suggest specific directions for future studies. However, papers whose major conclusion is a generalization that more studies are needed will generate little enthusiasm.

### Abstract

Abstracts are downloaded much more frequently than full-text papers. We read the abstract last, after we fully understand the purpose, methods, and results of the paper. Many journals have a required structure for abstracts. The editor may point that out, but we may save an additional cycle of revision by noting it in our review. The important qualities of an abstract are that it must convey information about the study and findings, not just be a teaser, and, in abbreviated form, it must convey the same information and conclusions as the manuscript.

## WRITING AND SUBMITTING THE REVIEW

When completing the write up, we follow the Golden Rule: "do to others what you would have them do to you." Simply put, good reviewers are professional, empathic, and helpful (2). That does not mean unflinchingly positive. No manuscript will meet all the high expectations we described, and we try to align our comments to the standards of the journal. We encourage empathy in ourselves by envisioning the recipients of our reviews.

They may have spent several years planning and conducting their experiments, analyzing data, and writing their manuscript. They believe it is interesting and important, even if we did not, and that it was written with crystalline clarity, even if it was not. They may be junior investigators, insecure about their future as scientists and prospects for promotion. English may not be their native language. Our role is to be helpful and instructive, like all feedback should be. We don't downplay flaws, which would not be instructive. But rejection hurts. We try not to crush the authors, which would not be helpful.

We organize our review in the manner that most journals request: summary, major comments, and minor (or specific) comments. Although the submission is done through the journal's online submission portal, we recommend writing it with your preferred word processing software and then cutting and pasting into the portal. Some journals will ask a series of questions regarding the manuscript. Most will have separate sections for you to enter confidential comments to the editor and open comments to the authors, the body of your review, and your publication recommendation.

The "Summary" should be your synopsis of the manuscript in three to four sentences (4). This is important to show the authors and editors that you understand the point of the manuscript before you get into the specific details. We consider it an act of kindness to use this space to note a few positive aspects of the paper before diving into its deficiencies.

The "Major Comments" section should indicate any major deficiencies in the paper. We consider the five questions we described in Get the Big Picture. For example, if there is an inadequate

literature review or inappropriate methods, this belongs in the “Major Comments” section. It is crucial to not just point out problems but also explain why it is a problem. We try to offer solutions or wording suggestions. We view our role as a partner with the author with the shared goals to improve the manuscript, improve the science, and advance the field. If the manuscript suffers from poor writing style, this should be included as a major comment but the details of that provided later.

We prefer the term “specific comments” to describe the next section because many of them are not minor. Line by line, this section should list all the other issues that we think the authors need to correct. Suggestions may include deletion of sentences, paragraphs, or whole sections. Reordering of paragraphs may improve flow and readability. Concerns about methods, statistical analysis, display of results, or interpretation that do not rise to the level of a major comment or that elaborate on a major comment belong here. We may note spelling and grammar here as well. For nonnative English speakers, we may just recommend that they consult with an English-fluent editor rather than returning a demoralizing long list of grammatical errors.

#### **Confidential Comments to the Editor**

This is brief but helpful to the editor. Do not omit it. It should summarize the major reasons for your recommendation to publish or not. Editors often receive conflicting recommendations from the several reviewers. This is the opportunity to make your case to them or to summarize the most important revisions that would allow the paper to move ahead. Be sure that your recommendation to the editor is concordant with your

review. In the comments to the authors, do not fawn over a paper that you do not like or provide minimal suggestions for improvement. Few things are more confusing and less helpful to authors than to get reviews that seem enthusiastic accompanied by a rejection. Do not include your publication recommendation in comments to the author. This decision is the purview of the editor.

The final step in a review is the recommendation regarding publication, typically, accept as is, reject, major revision, or minor revision. Although this may feel like a weighty burden, a reviewer is an advisor to the editor, one of several, and by now, your opinion should be well supported by your thoughtful narrative. We cannot recall any first submission for which we have recommended immediate acceptance. On the other hand, we have rejected many, guided by the quality metrics we have described as well as the selectivity of the journal. The boundary between major and minor revision is somewhat fluid. If the recommendations are largely cosmetic or stylistic edits, such as improving the logical order of the discussion, shortening sections, improving clarity of figures and tables, or dialing back the breadth of the conclusions, these may constitute a minor revision. If we have serious but nonfatal concerns about the methods, if substantial rewriting is needed, or if additional data or analysis is requested, these should constitute a major revision. In most journals, most major revision decisions are eventually published, perhaps after more than one revision. Because of that, the reviewer’s choice of major versus minor revision may have little practical consequences if the authors are willing to respond to your feedback.

### After Submission

When all reviews are submitted, the editor will adjudicate the viewpoints, write a short summary of the most important elements from all reviews, and send all reviews to the authors and other reviewers. We always read the other reviews. Because the editor tried to represent a variety of areas of expertise in assigning reviewers, we often learn points about the paper we missed. We also learn how to become better reviewers.

### Revisions

If the authors revise their paper, the journal editors will usually send it back to the same reviewers, although they may sometimes make a final publication decision without requiring another review cycle. The goal of your revisions review is to determine if the authors made the changes you requested or adequately justified their failure to do so. You should scrutinize any new information or data

that were included, but it is unfair to the authors to request new edits on anything that was present but not commented on in your first review.

### CONCLUSIONS

We hope this overview of our approach to peer review will be helpful to those who may be new to the process as well as those who are more experienced. For trainees and junior faculty who want to contribute, we suggest you let your mentors or more senior colleagues know. It is likely that they get more invitations to review than they have time to accept. With the editor's permission, they may be willing to review a paper together with you to gain experience and exposure. Peer review is a responsibility and a privilege to which we can all contribute and from which we all benefit.

**Author disclosures are available with the text of this article at [www.atsjournals.org](http://www.atsjournals.org).**

---

### REFERENCES

1. Smith R. Peer review: a flawed process at the heart of science and journals. *J R Soc Med* 2006;99:178–182.
2. Kelly J, Sadeghieh T, Adeli K. Peer review in scientific publications: benefits, critiques, & a survival guide. *EJIFCC* 2014;25:227–243.
3. Isohanni M. Peer review—still the well-functioning quality control and enhancer in scientific research. *Acta Psychiatr Scand* 2005;112:165–166.
4. Hoppin FG Jr. How I review an original scientific article. *Am J Respir Crit Care Med* 2002;166:1019–1023.
5. Goodman SN, Berlin J, Fletcher SW, Fletcher RH. Manuscript quality before and after peer review and editing at Annals of Internal Medicine. *Ann Intern Med* 1994;121:11–21.
6. Hays R, Jennings B, Gibbs T. How to review a paper on medical education. *MedEdPublish* 2019;8:9.
7. Friedman K, Lester J, Young JQ. Clinician-educator tracks for trainees in graduate medical education: a scoping review. *Acad Med* 2019;94:1599–1609.
8. Coleman MM, Richard GV. Faculty career tracks at U.S. medical schools. *Acad Med* 2011;86:932–937.



9. Sullivan GM, Simpson D, Cook DA, DeIorio NM, Andolsek K, Opas L, *et al*. Redefining quality in medical education research: a consumer's view. *J Grad Med Educ* 2014;6:424–429.
10. Reed DA, Cook DA, Beckman TJ, Levine RB, Kern DE, Wright SM. Association between funding and quality of published medical education research. *JAMA* 2007;298:1002–1009.
11. Azer SA, Ramani S, Peterson R. Becoming a peer reviewer to medical education journals. *Med Teach* 2012;34:698–704.
12. Green J, Willis K, Hughes E, Small R, Welch N, Gibbs L, *et al*. Generating best evidence from qualitative research: the role of data analysis. *Aust N Z J Public Health* 2007;31:545–550.
13. Glassick CE. Boyer's expanded definitions of scholarship, the standards for assessing scholarship, and the elusiveness of the scholarship of teaching. *Acad Med* 2000;75:877–880.
14. Kong Y, Seo YS, Zhai L. Comparison of reading performance on screen and on paper: a meta-analysis. *Comput Educ* 2018;123:138–149.
15. Zackoff MW, Real FJ, Abramson EL, Li ST, Klein MD, Gusic ME. Enhancing educational scholarship through conceptual frameworks: a challenge and roadmap for medical educators. *Acad Pediatr* 2019;19:135–141.
16. Knowles MS. Self-directed learning: a guide for learners and teachers. New York: Association Press; 1975.
17. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med* 2004;79:S70–S81.
18. Kolb DA. Experiential learning: experience as the source of learning and development. Hoboken, NJ: Prentice Hall; 1984.
19. Bordage G. Moving the field forward: going beyond quantitative-qualitative. *Acad Med* 2007;82: S126–S128.
20. Meyer HS, Durning SJ, Sklar DP, Maggio LA. Making the first cut: an analysis of academic medicine editors' reasons for not sending manuscripts out for external peer review. *Acad Med* 2018; 93:464–470.
21. Bowers B, Cohen LW, Elliot AE, Grabowski DC, Fishman NW, Sharkey SS, *et al*. Creating and supporting a mixed methods health services research team. *Health Serv Res* 2013;48:2157–2180.
22. Creswell JW, Clark VLP. Designing and conducting mixed methods research. Thousand Oaks, CA: SAGE Publications; 2011.
23. Halcomb E, Hickman L. Mixed methods research. *Nurs Stand* 2015;29:41–47.
24. Shorten A, Smith J. Mixed methods research: expanding the evidence base. *Evid Based Nurs* 2017; 20:74–75.
25. Moulton C. 10 tips for reviewing a qualitative paper. AM Rounds: Beyond the pages of Academic Medicine blog; 2014 [accessed 2021 Jul 28]. Available from: <https://academicmedicineblog.org/peer-reviewer-resources-10-tips-for-reviewing-a-qualitative-paper/>.
26. Shenton AK. Strategies for ensuring trustworthiness in qualitative research projects. *Educ Inf* 2004; 22:63–75.
27. Korstjens I, Moser A. Series: practical guidance to qualitative research. Part 4: trustworthiness and publishing. *Eur J Gen Pract* 2018;24:120–124.
28. Kern DE, Thomas PA, Hughes MT, editors. Curriculum development for medical education: a six step approach. Baltimore: Johns Hopkins University Press; 2016.

29. Ratnapalan S, Hilliard RI. Needs assessment in postgraduate medical education: a review. *Med Educ Online* 2002;7:4542.
30. Kirkpatrick DL. The four levels of evaluation. In: Brown SM, Seidner CJ, editors. *Evaluating corporate training: models and issues*. Dordrecht: Springer Netherlands; 1998. pp. 95–112.
31. Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB. *Designing clinical research*, third edition. Philadelphia: Lippincott William & Wilkins; 2006.
32. Artino AR Jr, Gehlbach H, Durning SJ. AM last page: avoiding five common pitfalls of survey design. *Acad Med* 2011;86:1327.
33. Artino AR Jr, La Rochelle JS, Dezee KJ, Gehlbach H. Developing questionnaires for educational research: AMEE Guide No. 87. *Med Teach* 2014;36:463–474.
34. Tufte ER. *The visual display of quantitative information*. Cheshire: Graphics Press; 2001.