

Evaluation
StandardsFocusing the
EvaluationStakeholder
Blueprint→
Focus & EvidenceDesign,
Implementation,
and AnalysisResults→
Scholarly
Activity

A Standards-Based Checklist for Program Evaluation as Scholarly Activity

Dorene F. Balmer, PhD (@dorenebalmer)

Jennifer A. Rama, MD, MEd (@jenramacastro)

Deborah Simpson^{ID}, PhD (@debsimpson3)

The Challenge

Educators in graduate medical education (GME) regularly make decisions about adding, deleting, or revising some aspect of a curriculum, rotation, or other program activity. At the core of each decision is the activity's value to the learners, faculty, program, and sponsoring institution, and/or to meeting accreditation/licensure requirements. When these decisions are informed by systematic program evaluation, the results may be generalizable and have value to others in GME beyond their local worth. Yet how does an educator know if their work might be relevant to others or fulfill scholarship criteria? These are particularly vexing challenges for evaluation, as every program is local and has unique settings, trainees, faculty, and resources. Thus, disseminating evaluation as scholarship is a balancing act: staying true to the local nature of the program while addressing concerns of skeptics who may say, "That's not relevant. . . it wouldn't work here."

What Is Known

Applying standards of program evaluation can tip the balance toward becoming a scholarly activity. Other Rip Outs^{1,2} in this series outlined the American Evaluation Association (AEA) standards (accuracy, feasibility, integrity [ie, propriety], and utility)³ and suggested their application to program planning and implementation. Viewing these same standards in a checklist format can guide evaluation as a scholarly activity for authors and reviewers. Reminder: standards apply to evaluation, not to the project (eg, asking if the evaluation is feasible, not if the project is feasible). Standards and sample items include:

- **Accuracy:** conveys trustworthy, reliable data in logical flow
 - Program outcomes go beyond acceptability and reaction.

Rip Out Action Items

1. Use the evaluation standards checklist for writing and reviewing program evaluation articles.
2. Incorporate the evaluation standards throughout your manuscript to sustain the storyline and logical flow.
3. Incorporate an evaluation model or program theory into your next project design.

- Evaluation findings (both intended and unintended) are sufficiently described.
- **Feasibility:** is doable—efficient, realistic, and cost-effective
 - The program and program context are sufficiently described so that others can determine relevance.
 - Evaluation methods are realistic given typical constraints of time, finances, and personnel in GME.
- **Integrity:** is fair and ethical with due regard to people involved in or responsible for the program; findings are honest and balanced.
 - Evaluation data give voice to multiple stakeholders. Apart from learners, what other stakeholders were included (eg, instructors, families, GME, and/or system leadership)?
 - Institutional review board/external review is addressed.
- **Utility:** is useful for decision-making, addresses the needs of people involved in or responsible for the program; findings are clear, concise, and on time.
 - Evaluation addresses a local problem at minimum and may speak to a larger pressing problem.
 - Evaluators share practical, transferable lessons learned.

As evaluators, readers, and editors, we like checklists ordered to match the standard introduction/aim, methods, results, discussion/conclusions (IMRD) framework. However, evaluation standards underpin the entire program evaluation endeavor; thus, one standard can—and often should—appear in multiple sections. See online

DOI: <http://dx.doi.org/10.4300/JGME-D-22-00603.1>

Editor's Note: The online version of this article contains the complete evaluation standards checklist.

supplementary data for the complete Evaluation as Scholarly Activity Checklist by AEA Standards.

How You Can Start TODAY

1. **Move your evaluation toward scholarship.** Use the checklist to determine if your evaluation would be of interest to others outside your organization.
 - Did you evaluate a local problem that is of interest to others in your field? Is it clear what a broader readership can learn about this problem from your evaluation? (Utility)
 - Could your evaluation be reasonably conducted in other contexts, even if the findings would differ? (Feasible)
 - Have you incorporated multiple stakeholders' perspectives, rather than a select group? (Integrity)
 - Can you understand what the program evaluators did and why? (Accuracy)
2. **Write an abstract before you start writing the article.** Share it with colleagues beyond your local setting and ask if the evaluation's utility, feasibility, integrity, and accuracy are clear.
3. **Draft the evaluation manuscript.** Evaluation articles tell a story that starts with a common and important problem. In the introduction, reference the problem, the gap—what others have tried and the limits of their success—and the evidence-supported solution you propose. In the methods, describe what was done to address the problem as well as the evaluation model or program theory you used. Emphasize the role of key stakeholders and the steps taken to compile accurate (if not psychometrically valid) evaluation data. In the results, present findings that are balanced, fair, and that logically flow from your introduction and methods. Conclude your evaluation story in the discussion by reporting what decisions were made based on the data, limitations, and transferable lessons to inform readers.
4. **Remember, no evaluation is perfect!** Addressing every item under each checklist standard is typically not feasible in an evaluation report. If you cannot address a standard, discuss that in the limitations. Apply these same expectations to authors whose work you review.

What You Can Do LONG TERM

1. **Start an evaluation.** Use the evaluation standards checklist to routinely monitor the evaluation process, from planning to write-up. Seek to utilize

an evaluation model or program theory to provide a consistent approach throughout the evaluation's design, implementation, analysis, and write-up. Ideally the proposed solution is based on an established theory of learning or change to provide explanatory power as to how and why the activity worked (or did not).

2. **Select a journal or venue.** Journals vary regarding acceptance of evaluation manuscripts. Seek journals that have previously published evaluation articles. Consider how these articles moved the problem from local to beyond. What evaluation model and/or program theory did the authors utilize to guide their solution and frame their findings? Negotiate with editors to allow the format of your program evaluation paper to be consistent with your purpose.
3. **Practice reviewing evaluation papers.** Volunteer to review and critique evaluation manuscripts for your department or peer-reviewed journals to help improve the quality of your own writing and understanding of the criteria for program evaluation as scholarship.
4. **Learn more about evaluation as scholarship.** Attend a course, workshop, or conference sponsored by the AEA or Centers for Disease Control. Take part in national faculty development programs or graduate certificates/degrees.

References and Resources

1. Balmer DF, Riddle JM, Simpson D. Program evaluation: getting started and standards. *J Grad Med Educ.* 2020;12(3):345-346. doi:10.4300/JGME-D-20-00265.1
2. Balmer DF, Rama JA, Simpson D. Implementing the evaluation plan and analysis: who, what, when, and how. *J Grad Med Educ.* 2021;13(1):129-130. doi:10.4300/JGME-D-20-01523.1
3. Yarbrough DB, Shula LM, Hopson RK, Caruthers FA. *The Program Evaluation Standards: A Guide for Evaluators and Evaluation Users.* 3rd ed. Thousand Oaks, CA: Corwin Press; 2010.



Dorene F. Balmer, PhD, is Professor of Pediatrics, Perelman School of Medicine, University of Pennsylvania, and Associate Editor, *Journal of Graduate Medical Education* (JGME); **Jennifer A. Rama, MD, MEd**, is Associate Professor of Pediatrics, Baylor College of Medicine; and **Deborah Simpson, PhD**, is Director of Education—Academic Affairs, Advocate Aurora Health, Adjunct Clinical Professor of Family Medicine, University of Wisconsin School of Medicine and Public Health and Medical College of Wisconsin, and Deputy Editor, JGME.

Corresponding author: Deborah Simpson, PhD, Advocate Aurora Health, deb.simpson@aah.org, Twitter @debsimpson3