

## Moving on From Self-assessment

Christopher P. Morley, PhD

PRiMER. 2024;8:5.

Published: 1/23/2024 | DOI: 10.22454/PRiMER.2024.624901

### Abstract

Self-assessment of knowledge and confidence is common in medical education, and there are both philosophical and practical justifications for it. However, many attempts to establish a correlation between self-assessments of knowledge or confidence and objective measures of knowledge or skill acquisition have failed. While in some circumstances the inclusion or reliance of self-assessment may be warranted, for example when a study is specifically measuring traits or outcomes that rely upon meta-cognition or increases in confidence, it is more often the case that self-assessment is used as a substitute for more objective measures. This is demonstrably flawed, and *PRiMER* as a journal will be moving away from publishing reports that inappropriately rely upon self-assessed knowledge or confidence as the only study outcomes.

The use of self-assessment is prevalent in medical education, particularly in the context of single-institution, smaller, unfunded evaluations. Self-assessment can include a variety of question types that asks what a learner thinks they know about a topic (knowledge), or how confident the learner feels about doing a procedure, prescribing a medication, treating a specific condition, or working with a certain type of patient or population (confidence). Together, self-assessed knowledge and confidence are often used to stand in for objectively measured or observed knowledge or skill.

In education, self-assessment is philosophically tied to the concept of lifelong learning and to problem-based learning. An excellent and quick overview of the evolution of both peer- and self-assessment is included in the introduction to a paper by Papinczak and colleagues.<sup>1</sup> The ability to accurately evaluate one's own acquisition of new knowledge and skills is laudable, and the concept is supported by cognitive theory as well.<sup>2</sup> However, in practice, self-assessment has repeatedly been proven to be poorly correlated with other external or objective assessments of knowledge or skill. To give just a few examples, this has proven true when comparing self-assessment against peer- and tutor-assessment<sup>1</sup> in problem-based learning scenarios<sup>1</sup>; in resident, medical student, and other learner predictions of their own test performance<sup>3-5</sup>; and repeatedly in comparisons of confidence versus knowledge, such as in statistical literacy among clinicians,<sup>6</sup> or ability versus confidence of residents to accurately diagnose dementia.<sup>7</sup> The mismatch between confidence, self-assessed knowledge or skill, and objectively measured knowledge and skills can be tied to the Dunning-Kruger Effect,<sup>8</sup> which describes a nonlinear relationship between how much learners believe they know, versus how much they actually know about a topic, with the mismatch being more pronounced in those who know the least about a topic.<sup>2,9</sup>

Unfortunately, self-assessments of both knowledge and confidence are frequently used to evaluate curricular elements or training modalities, not because of a theoretically-based commitment to lifelong learning, but

rather because they are, frankly, easy to implement. Much like a preponderance of the literature that has reported on comparisons between self- versus objective assessment, a student-led study at my own institution recently found the same thing: we found very poor correlations between fourth-year medical student self-assessments of confidence in their own knowledge and ability to manage diabetes, and objective assessment using diabetes questions from standardized examinations.<sup>10</sup>

At *PRiMER*, we have been progressively applying stricter thresholds for publishing research reports that rely upon self-assessed knowledge or confidence as the only outcome measures. As described in our quality guidelines,<sup>11</sup> *PRiMER* utilizes the Kirkpatrick Model of Assessment in our evaluation of submitted manuscripts. Within the context of the Kirkpatrick framework, we view self-assessment as essentially a level-1 “reaction” measure. Given the poor performance of self-assessment, crossing many domains of knowledge and skill acquisition, and applying to both self-assessment of knowledge and confidence, we believe it is important to hold this line.

Of course, there are times when self-assessment of knowledge or confidence are appropriate. A few examples include:

- The inclusion of a self-assessment alongside other quantitative and objective measures of knowledge or skill acquisition, or alongside qualitative, phenomenological descriptions of student experiences.
- Studies that actually examine meta-cognition, or that evaluate educational activities that are specifically designed to increase self-awareness, confidence, or similar domains. For example, many studies that observe confidence in treating patients or populations that differ from the clinician (in gender, race, culture, or socio-economic status or other traits) can impact who engages in care, who enters specific specialties or practice contexts, etc. In cases like these, where the entire point of an intervention is to increase confidence, it may be intellectually honest to ask the learners about their own confidence levels.
- There are also circumstances where no validated measures that exist, and/or confidence is a necessary precursor to action. For example, an activity may be designed specifically to impart confidence in taking an action, such as “engaging in advocacy work.” It makes sense to ask learners if they feel more confident and interested to engage in advocacy work, if that confidence is required to take the next step.

An example that encompasses most of these points appears in a 2023 paper by Robinson and Mishori.<sup>12</sup> In that report, self-assessed knowledge and confidence questions were combined with a deeper qualitative study phase to give an overall picture of the outcomes from a medical student advocacy workshop. There are few widely-recognized instruments or formalized assessments of advocacy skills, there is a need for a person to be interested and engaged in order to move forward with advocacy activities, and the manuscript combined self-assessment with other modalities of evaluation. This is a good recent example of where we have moved forward with a publication that incorporated self-assessment.

Where we will take a much stronger stance is with instances where self-assessed knowledge or comfort are simply presented as a proxy for objective measurement of knowledge or skill acquisition. The model evaluation will contain:

- Measurement of skill or knowledge against a baseline
- Preferably with either a reference population, statistical control for covariates, or both
- Assessment that is based upon objective measurement (eg, quiz items, tests, outcomes on standardized examinations) or external (nonself) observation
- Analysis conducted within a reasonable time interval and an appropriate level of rigor.

We believe that enforcing these standards, which are in keeping with the quality guidelines that have been in place since *PRiMER* began, will serve two purposes: increasing the overall rigor of the journal and its contributions to the literature, and aiding our educational mission to nurture new scholars as they initiate their

research careers. Holding a higher bar will establish good habits at the outset of those careers, and contribute to an overall improvement in the quality of medical education research, as an aspirational goal for *PRiMER*.

## Corresponding Author

Christopher P. Morley, PhD

Department of Public Health & Preventive Medicine, Upstate Medical University, 750 E Adams St, Weiskotten Hall 2262, Syracuse, NY 13210. 315-464-1520

[morleycp@upstate.edu](mailto:morleycp@upstate.edu)

## Author Affiliations

Christopher P. Morley, PhD - Departments of Public Health & Preventive Medicine and Family Medicine, SUNY Upstate Medical University, Syracuse, NY

## References

1. Papinczak T, Young L, Groves M, Haynes M. An analysis of peer, self, and tutor assessment in problem-based learning tutorials. *Med Teach*. 2007;29(5):e122-e132. doi:10.1080/01421590701294323
2. Fraundorf SH, Caddick ZA, Nokes-Malach TJ, Rottman BM. Cognitive perspectives on maintaining physicians' medical expertise: III. Strengths and weaknesses of self-assessment. *Cogn Res Princ Implic*. 2023;8(1):58. doi:10.1186/s41235-023-00511-z
3. Parker RW, Alford C, Passmore C. Can family medicine residents predict their performance on the in-training examination? *Fam Med*. 2004;36(10):705-709.
4. von Hoyer J, Bientzle M, Cress U, Grosser J, Kimmerle J, Peter Holtz. False certainty in the acquisition of anatomical and physiotherapeutic knowledge. *BMC Med Educ*. 2022;22(1):765. doi:10.1186/s12909-022-03820-x
5. Jones R, Panda M, Desbiens N. Internal medicine residents do not accurately assess their medical knowledge. *Adv Health Sci Educ Theory Pract*. 2008;13(4):463-468. doi:10.1007/s10459-007-9058-2
6. Lakhlifi C, Lejeune FX, Rouault M, Khamassi M, Rohaut B. Illusion of knowledge in statistics among clinicians: evaluating the alignment between objective accuracy and subjective confidence, an online survey. *Cogn Res Princ Implic*. 2023;8(1):23. doi:10.1186/s41235-023-00474-1
7. Lerner BS, Kalish V, Ledford CJW. Exploring residents' skills in diagnosing dementia: the unexpected dissonance between ability and confidence. *Fam Med*. 2017;49(6):460-463.
8. Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J Pers Soc Psychol*. 1999;77(6):1121-1134. doi:10.1037/0022-3514.77.6.1121
9. Rahmani M. Medical trainees and the Dunning-Kruger effect: when they don't know what they don't know. *J Grad Med Educ*. 2020;12(5):532-534. doi:10.4300/JGME-D-20-00134.1
10. Pitter D, Indelicato AM, Morley CP, Feuerstein B, Weinstock RS. US fourth-year medical students: diabetes knowledge & confidence dissonance. *PRiMER Peer-Rev Rep Med Educ Res*. 2024;8:4. doi:10.22454/PRiMER.2024.497586
11. *PRiMER*. Quality Guidelines for Authors and Reviewers. Accessed January 15, 2024. [https://journals.stfm.org/primer/authors/#Quality\\_Guidelines](https://journals.stfm.org/primer/authors/#Quality_Guidelines)
12. Robinson R, Mishori R. The efficacy of short, skills-based workshops in teaching advocacy to medical students: a pilot study. *PRiMER Peer-Rev Rep Med Educ Res*. 2023;7:21. doi:10.22454/PRiMER.2023.427789