

EDUCATIONAL DILEMMA OR QUESTION

Laura is a PGY2 resident in Emergency Medicine. She is meeting with her Academic Advisor in advance of her 6-month performance review. Of the 55 EPAs she has received, the majority of the comments are brief, non-descriptive and do not provide recommendations on how to get better. Laura and her advisor are frustrated that they do not have more specific information to develop learning goals to focus on. They share their frustration with the Program Director who agrees it is an issue they need to work on with their front-line faculty.

Article Reference

Teresa M. Chan, Stefanie S. Sebok-Syer, Christopher Sampson & Sandra Monteiro (2020): The Quality of Assessment of Learning (Qual) Score: Validity Evidence for a Scoring System Aimed at Rating Short, Workplace-Based Comments on Trainee Performance, Teaching and Learning in Medicine

DOI: 10.1080/10401334.2019.1708365

Why Is This Paper Relevant to Emergency Medicine Education?

Competency Based Medical Education (CBME) in Emergency Medicine is structured around Entrustable Professional Activity (EPA) assessments. As with all forms of feedback in medicine, the quality is variable. Learners want useful feedback, faculty are keen to learn how to best provide this feedback and programs are invested in ensuring that their systems of assessment are of high quality. Generating validity evidence for a practical tool derived to evaluate brief narrative comments in work-based assessments would be welcome by front line faculty and residency programs.

Level of Evidence/Level of Learning

Level III or IV / Undergraduate Medical Education

Funding Sources

Study was supported by the W. Watson Buchanan Clinician Educator Internal Career Award received by Dr. Chan.

Study Design

Randomized multicenter meta-rating survey was conducted where 25 participants from 20 sites were asked to evaluate 50 short qualitative comments each associated with a single clinical performance score.

Setting

The Internet – Surveys were distributed and completed online.

Synopsis

A novel metric named Quality of Assessment for Learning (QuAL) score was developed using Messick's validity framework and was utilized by half the participants to evaluate commentary. It contained three questions:

- Evidence: Does the rater provide sufficient evidence about resident performance? (No comment, No, Somewhat, Yes)
- Suggestion: Does the rater provide a suggestion for improvement? (No, Yes)
- Connection: Is the rater's suggestion linked to the behavior described? (No, Yes)

Synopsis (continued)

They were compared to the group utilizing the more comprehensive Completed Clinical Evaluation Report Rating (CCERR). The CCERR has validity evidence for rating In-Training Evaluation Reports (ITERS) and consists of 9 questions, each marked along a 5-point scale. Both groups also utilized a “utility rating rubric” consisting of two standardized questions regarding the perceived utility of the EPA comments.

The results showed that both the CCERR (n=13) QuAL group (n=12) rated the comments with high reliability (Phi = 0.97). Furthermore, the QuAL score required only two raters to reach an acceptable target reliability of >0.80, while the CCERR required three. Although equally robust in terms in reliability, the QuAL score has the advantage of being simpler and more user-friendly to rate shorter comments generated by workplace-based assessments.

A thematic analysis also revealed that the most useful comments were those that provide specific information about a trainee and their progression, the clinical context, complexity of the case, or offer some opportunity or suggestion for future improvement.

BOTTOM LINE

The Quality of Assessment for Learning (QuAL) tool has validity evidence for rating short narrative comments in workplace-based assessment. This 3-question tool is a potential resource for coaching front-line faculty on how to fill out direct observation assessments and for programs looking to measure quality of their assessment programs.