CANADIAN MEDICAL EDUCATION JOURNAL | REVUE CANADIENNE DE L'ÉDUCATION MÉDICALE

2026;17(5): https://doi.org/10.36834/7e1tgj57

BRIEF REPORTS

A qualitative study of validation practices in classroom-based assessment in health professions education: what do faculty members actually do?

Mélanie Marceau,[1] *Kathleen Ouellet,*[2] *Aliki Thomas,*[3] *Jean-Sébastien Renaud,*[4] *Christina St-Onge*[5]

*Author information is provided in the back matter of this manuscript

Abstract

Background: Faculty in health professions education (HPE) play a key role in assessing student learning through classroom-based assessments. While researchers have studied admission and licensure assessments, we know little about how classroom assessments are validated. Through validation practices, faculty help ensure that decisions about student learning are well-supported. This study explores how HPE faculty validate classroom-based assessments.

Methods: The authors conducted a qualitative interpretive descriptive study. Twenty faculty members in HPE from four Canadian universities were recruited for individual interviews using quota sampling. Two team members analyzed the qualitative data through inductive thematic analysis and held discussions with the team to reach a consensus.

Results: Although they do not refer to recommended practices by name, participants used many means to ensure the quality of their classroom-based assessments. When preparing their assessments, they consider elements such as authenticity, content representativity and adequate preparation of students. During assessment, they pay attention to elements such as plagiarism and support. Following the assessment, they make decisions on scoring strategy and process and use assessment data to adjust assessments.

Conclusion: Faculty members in HPE recognize the importance of validation and used multiple strategies to document validity evidence. To better support faculty in their validation practices, it is valuable to understand current approaches and provide adequate resources to build on existing validation practices.

© MARCEAU, OUELLET, THOMAS, RENAUD, ST-ONGE

Étude qualitative des pratiques de validation des évaluations en contexte de classe dans la formation aux professions de santé : ce que font concrètement les enseignants

Résumé

Contexte : Les enseignants de programmes de professions de la santé jouent un rôle clé dans l'évaluation des acquis des étudiants au moyen d'évaluations en classe. Si les chercheurs se sont penchés sur les évaluations d'admission et de certification, on en sait peu sur la manière dont les évaluations en classe sont validées. Grâce à ces pratiques de validation, les enseignants contribuent à garantir que les décisions relatives aux acquis des étudiants reposent sur des bases solides. Cette étude explore la manière dont les enseignants de programmes de professions de la santé valident les évaluations en classe.

Méthodes : Les auteurs ont mené une étude qualitative avec une approche descriptive interprétative. Vingt enseignants de programme de professions de la santé issus de quatre universités canadiennes ont été recrutés pour des entretiens individuels à l'aide d'un échantillonnage par quotas. Deux membres de l'équipe ont analysé les données qualitatives par le biais d'une analyse thématique inductive et ont tenu des discussions avec l'équipe afin de parvenir à un consensus.

Résultats : Bien qu'ils ne mentionnent pas explicitement les pratiques recommandées, les participants ont utilisé de nombreux moyens pour garantir la qualité de leurs évaluations en classe. Lors de la préparation de leurs évaluations, ils prennent en compte des éléments tels que l'authenticité, la représentativité du contenu et la préparation adéquate des étudiants. Pendant l'évaluation, ils prêtent attention à des éléments tels que le plagiat et le soutien. Après l'évaluation, ils prennent des décisions concernant la stratégie et le processus de notation et utilisent les données d'évaluation pour ajuster les évaluations.

Conclusion : Les enseignants de programmes de professions de la santé reconnaissent l'importance de la validation et ont utilisé de multiples stratégies pour documenter les preuves de validité. Afin de mieux soutenir les enseignants dans leurs pratiques de validation, il est utile de comprendre les approches actuelles et de fournir des ressources adéquates pour s'appuyer sur les pratiques de validation existantes.

Introduction

While there is ample literature documenting the validation processes of admission, certification and licensure assessments in health professions education (HPE),[1] we know less about how validation is approached in classroom-based assessments by everyday teachers. Yet, these assessments are essential for monitoring students' progress,[2] including the acquisition of knowledge, skills, and attitudes aligned with key milestones in HPE. Classroom-based assessments play a crucial role in documenting student performance over time, using a variety of assessment methods across different contexts,[3] and providing feedback to students.[4] They offer a continuous, contextualized view of student development and contribute meaningfully to decision-making processes about readiness, progression, and competence. However, how teachers validate their assessment work remain largely unknown.

Validation is the process of examining the appropriateness of the interpretation and use of assessment results.[3,5–8] As Chan[9] points out, despite divergent definitions of validation and validity, a common

contemporary view is that validation involves accumulating evidence to support the decisions we make based on assessment results. In addition to traditional sources of validity evidence (e.g., content, internal structure, and relations to other variables), more recent frameworks include evidence based on response processes and consequences.

Despite their central role in student learning, classroom-based assessments often receive less attention when it comes to validation than high-stakes assessments. Faculty members responsible for these assessments[10] generally have access to fewer resources (e.g., time, psychometricians) than those involved in high-stakes assessments. Previous research has described the many demands placed on faculty, such as teaching, research, and service.[11,12] As a result, validation practices might be less comprehensive or less documented than those used in higher-stakes assessments. This matters because faculty must support the interpretations of assessment results and the decisions that stem from them. Therefore, the purpose of our study was to document the validation practices of HPE faculty members in their classroom-based assessments.

Conceptual framework

We drew on the principles of knowledge translation (KT), as these can support the implementation of evidence-informed educational practices and policies. KT is usually defined as "a dynamic and iterative process that includes the synthesis, dissemination, exchange, and ethically sound application of knowledge to improve health [...]"[13,14] Ultimately, we aimed to bridge the gap between theory and practice—specifically, the disconnect between recommended validity and validation practices (e.g., *Standards for Educational and Psychological Testing*[15]) and how these are actually applied and engaged with in classroom-based assessments. Our ultimate goal is to support faculty members more effectively in their validation practices.

Materials and methods

Study design

We conducted a qualitative study to explore how faculty members in HPE ensure the quality of their assessments. We used Thorne's[16] interpretive description approach, a qualitative methodology developed for applied disciplines such as health and education. This approach emphasizes the generation of practical and contextually relevant knowledge.

Participants and recruitment

We recruited faculty members involved in HPE from four Canadian universities. To be included, participants had to be 1) in a tenure-track position or have a teaching contract in medicine, nursing, occupational therapy or physiotherapy, programs from the targeted universities; 2) involved in the design, development, or validation of assessments; and 3) able to communicate in French (French being the language of teaching at all four institutions).

Data collection material and procedure

We developed and piloted an interview guide. In the introduction to the interview, we explicitly presented the research objectives, specifying that the study focused on validation practices in assessment of learning. We informed participants that the interview aimed to explore how they ensure the quality of the assessments they design or use.

We first discussed with participants the assessment strategies they use most frequently. Because we assumed that our participants would not necessarily be familiar with validity theories and validation practices, we chose not to frame the conversation around a specific theoretical model (e.g., the *Standards for Educational and Psychological Testing*[15]) or categories of validity evidence (e.g., content, response process, internal structure, relations to other variables, and consequences).[15] Instead, we adopted an inductive approach and asked participants to describe in detail, and in a chronological manner, how they ensured the quality of their assessments.[17] We

asked them how they developed their assessments, how these assessments were conducted, and what steps they took following the assessments. Our intention was to avoid orienting participants' responses toward a particular normative framework and instead allow them to describe their validation practices as they understand and implement them in their own contexts. The second part of the interview focused on the barriers and facilitators experienced during validation processes. These results are not included in this paper but will be presented in another publication.

We conducted semi-structured individual interviews between July and December 2021. We compiled sociodemographic data using a short survey. An experienced research professional (KO) conducted individual interviews (approximately 60 minutes in length) using the Microsoft Teams video-conferencing platform. We conducted the interviews in French, the native language of our participants.

Ethical considerations

Participation was voluntary. We obtained ethics approval from the Ethical Review Board of Education and Social Sciences of one university included in our study on June 9, 2021 (file number: 2021-2938/Marceau). This was followed by approval from the three other universities. We followed the ethics standards and the principles of the Declaration of Helsinki.

Data analysis

We summarized participants' sociodemographic characteristics using descriptive statistics and inductively analyzed the interview data following Braun and Clarke's[18] six-step approach to thematic analysis. We managed and analyzed the qualitative data using Dedoose software (Version 9.0.17). The analysis started with two team members (MM and KO) familiarizing themselves with the data (step 1). After generating a preliminary set of codes (step 2), the same team members each independently coded two transcripts. The resulting initial code tree was revised by all team members. MM and KO then identified themes by linking related codes (step 3). We unpacked and organized validation practices to clearly outline the chronological steps involved. We did not classify validation practices according to existing validity frameworks (e.g., *Standards for Educational and Psychological Testing*,[15] the frameworks of Messick[8,19,20] or Kane[21]) because we wanted to remain closer to the faculty members' verbatim. We refined the themes and subthemes through an iterative process (step 4) and discussed the final structure and description of the themes as a team (step 5). The themes and supporting excerpts are presented below (step 6).

Results

We conducted 20 semi-structured individual interviews. Participants were involved in medical (n = 7), nursing (n = 7), and occupational therapy or physiotherapy (n = 6) programs. Of the 20 participants, 70% (n = 14) were women and 65% (n = 13) had completed training in education or health professional education. More specifically, among those who had received such training, four had completed a graduate certificate (microprogram), two had completed a master's or a Ph.D. degree, and seven had attended one-time workshops or training sessions.

Assessment strategies used

Our analysis of the 20 interviews revealed a variety of assessment strategies (one participant could use multiple assessment types): multiple-choice questions (MCQ) and/or short answer questions (SAQ) (n = 12); short essay questions (SEQ) or long essay questions (n = 10); simulation activities (including objective structured clinical examination [OSCE]; n = 11); workplace-based, clerkship or internship assessments (n = 6), and longitudinal assessment of non-clinical skills (n = 1).

Validation processes used

Although participants did not refer to recommended practices by name, they used many means to ensure the quality of their classroom-based assessments. When preparing their assessments, they considered elements such as authenticity. For example, one participant mentioned that students would go on clinical placements after his course. He therefore aimed to create clinical vignettes that were "as

close as possible to clinical reality" (Interview 10). Participants also considered content representativity, as expressed by this participant: "every time I finish a course, I have key messages, and then the key messages are my checklist of items that I will assess (...) It's my way of making sure that everything is aligned" (Interview 5).

During assessments, they paid attention to elements such as plagiarism (e.g., by changing the order of questions from one group to another) and to supporting students (e.g., by providing access to assessment grids and assessment criteria).

After assessments, they made decisions regarding scoring strategies and processes. Participants reported strategies such as grading in small groups or grading one question at a time rather than one exam at a time. Participants also emphasized the importance of justifying what needed adjustment in future assessments: "is it just a few individuals or does it concern the entire cohort, meaning that there is something going on in the teaching?" (Interview 18). Table 1 presents a comprehensive overview of the elements considered before, during, and after assessments, with supporting examples.

Nevertheless, although participants described numerous strategies aimed at ensuring assessment quality, their accounts suggested that validation was not consistently conceptualized as an explicit and iterative process. For example, one participant reported relying on existing assessments without critically re-examining their continued relevance: "Sometimes you go for the easy option. If you have an assessment that's there, fine, you use it" (Interview 2). A second example illustrates how grading decisions were described as final once completed, suggesting limited reconsideration of assessment judgments after marking: "There are elements in my exams that are more subjective (...) but once it's finished, once I've graded it and followed my rubric, well, there's really no way to say, 'Well, I should have given more points'" (Interview 15).

Table 1. Considerations identified by participants to ensure assessment quality during three key moments

Key moments	Considerations	Examples
Before the assessment	Selection of the assessment strategy	Alignment/consistency between teaching strategies and assessment strategies Feasibility Authenticity Decisions grounded in evidence-based practices
	Content assessed	Content representativity (e.g. blueprinting) Clarity of questions
	Preparation of students and assessors	Clear expectations for assessors and students Practice assessment Preparation of materials Standardization (e.g., standardized patients, trained assessors)
During the assessment	Administration of the assessment	Providing necessary materials
	Plagiarism and academic integrity	Exam invigilation Question order Internet access policies
	Support offered	Consultation periods (e.g. for writing assignment) Feedback during clinical assessments
After the assessment	Scoring decisions	Scoring strategy (e.g., scoring as a team vs. individually) Scoring process (e.g., item-by-item vs. full-assessment scoring)
	Monitoring quality	Statistical analysis Post exam surveys
	Use of assessment data	Adjusting administrated assessment (e.g., weighting) Revise future assessments Revise teaching strategies Implementing processes to manage challenges or failures

Discussion

In this study, we aimed to explore HPE faculty members' validation processes for their classroom-based assessments. Our results show that participants reflect and take action at different key moments in the process of validating classroom-based assessments.

We observed that faculty members took multiple considerations into account to document the quality of their assessment, without necessarily using the terminology associated with recommendations such as those found in the *Standards for Educational and Psychological Testing*.[15] Participants appear to be guided by an intuitive commitment to design, administer, and monitor high-quality assessments. At the same time, their validation practices may also reflect prior training in assessment.

However, a gap remains between the informal, experience-based strategies that faculty members report using and the validation frameworks promoted in the literature (e.g., the *Standards for Educational and Psychological Testing*[15]). This gap does not reflect an absence of quality-oriented practices. Rather, it lies in how validation is conceptualized and articulated. Although faculty described practices that align with recommended sources of validity evidence, these were rarely framed within an explicit validation framework, and validation was not consistently described as a continuing process. This highlights the need for greater alignment and support to bridge this divide.

The gap between how faculty members describe their practices and the expected validation practices could stem from two main factors. First, our findings suggest that current guidelines, such as the *Standards for Educational and Psychological Testing*,[15] may not be widely known or sufficiently tailored to the realities of classroom-based assessment, making them difficult for faculty members to apply in practice. Second, most faculty members are clinicians and although 65% of them have received training in learning assessment, it is possible that they have no formal training in validation practices or the concept of validity itself. Addressing this gap requires strategies to increase their knowledge and skills with expected validation practices. To better support faculty in their validation efforts, it is essential to understand current practices, propose practical solutions, provide adequate resources, and invest in faculty development.

Limitations

Our study involved only 20 participants, however, they were recruited from four Canadian universities and represented a range of academic career stages and health professions programs. In line with the aims of qualitative research, our goal was transferability rather than statistical generalizability, and we believe that these findings may resonate with faculty members engaged in classroom-based assessment.

Other limitations should also be acknowledged. Participants were recruited using quota sampling and volunteered to participate, which may have introduced selection bias. Faculty members with a particular interest in the assessment of learning may have been more likely to participate, potentially leading to an over-representation of reflective or assessment-engaged practices. In addition, the data were based on self-reported practices obtained through interviews, and participants' descriptions may therefore have been influenced by recall bias or social desirability bias.

Conclusion

All faculty members in our sample demonstrate an awareness of the importance of validation in assessment practices. Although, they may not consistently employ formal validation terminology, they nonetheless engage in a variety of strategies to gather and document the quality of their classroom-based assessments. These findings underscore the need to better understand and support these intuitive validation practices, including through targeted faculty development and training initiatives that build on the wide range of validation practices that faculty are already using. This study represents an important initial step toward our long-term goal of developing and implementing effective strategies to assist faculty in conducting rigorous, high-quality validation processes.

Author information:

1- School of Nursing, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Quebec, Canada

2- Health Sciences Education Centre, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Quebec, Canada

3- School of Physical and Occupational Therapy, Faculty of Medicine, McGill University, Quebec, Canada

4- Department of Family Medicine and Emergency Medicine, Faculty of Medicine, Université Laval, Quebec, Canada

5- Department of Medicine, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Quebec, Canada

Correspondence to:

Mélanie Marceau

email: melanie.marceau@usherbrooke.ca

Published ahead of issue:

May 26, 2026

© 2026 MARCEAU, OUELLET, THOMAS, RENAUD, ST-ONGE; licensee Synergies Partners. This is an Open Journal Systems article distributed under the terms of the Creative Commons Attribution License. (https://creativecommons.org/licenses/by-nc-nd/4.0) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is cited.

Conflict of Interest:

All authors declare that they have no conflict of interest.

Funding:

This study was supported by startup funds for professors offered by the School of Nursing and the Faculty of Medicine and Health Sciences of the Université de Sherbrooke.

Acknowledgements:

The authors would like to thank the participants for their generosity and availability.

References

1. Young M, St-Onge C, Xiao J, Vachon Lachiver E, Torabi N. Characterizing the literature on validity and assessment in medical education: a bibliometric study. *Perspect Med Educ.* 2018;7(3):182-91. https://doi.org/10.1007/s40037-018-0433-x
2. Touchie C, Kinnear B, Schumacher D, et al. On the validity of summative entrustment decisions. *Med Teach.* 2021;43(7):780-7. https://doi.org/10.1080/0142159x.2021.1925642
3. van der Vleuten CPM, Schuwirth LWT. Assessing professional competence: from methods to programmes. *Med Educ.* 2005;39(3):309-17. https://doi.org/10.1111/j.1365-2929.2005.02094.x
4. Baird JA, Andrich D, Hopfenbeck TN, Stobart G. Assessment and learning: fields apart? *Assess Educ Princ Policy Pract.* 2017;24(3):317-50. https://doi.org/10.1080/0969594X.2017.1319337
5. Cizek GJ, Rosenberg SL, Koons HH. Sources of validity evidence for educational and psychological tests. *Educ Psychol Meas.* 2008;68(3):397-412. https://doi.org/10.1177/0013164407310130
6. DeLuca C. Interpretive validity theory: mapping a methodology for validating educational assessments. *Educ Res.* 2011;53(3):303-20. https://doi.org/10.1080/00131881.2011.598659
7. Downing SM. Validity: on the meaningful interpretation of assessment data. *Med Educ.* 2003;37(9):830-7. https://doi.org/10.1046/j.1365-2923.2003.01594.x
8. Messick S. Standards of validity and the validity of standards in performance assessment. *Educ Meas Issues Pract.* 1995;14(4):5-8. https://doi.org/10.1111/j.1745-3992.1995.tb00881.x
9. Chan EKH. *Standards and guidelines for validation practices: development and evaluation of measurement instruments.* In: Zumbo BD, Chan EKH, editors. Validity and validation in

social, behavioral, and health sciences. Cham: Springer International Publishing; 2014. p. 9-24. https://doi.org/10.1007/978-3-319-07794-9_2
10. Gordon S, Smith E. Who are faculty assessment leaders? *Assess Eval High Educ.* 2022;47(6):928-41. https://doi.org/10.1080/02602938.2021.1983518
11. García-Ramírez Y, Bijelić V. Assessing task prioritization for professors through affinity and satisfaction scores. *Cogent Educ.* 2024;11(1):2321355. https://doi.org/10.1080/2331186X.2024.2321355
12. Link AN, Swann CA, Bozeman B. A time allocation study of university faculty. Econ Educ Rev. 2008;27(4):363-74. https://doi.org/10.1016/j.econedurev.2007.04.002
13. Canadian Institutes of Health Research. *Knowledge translation.* Available from: https://cihr-irsc.gc.ca/e/29418.html
14. Straus SE, Tetroe JM, Graham ID. Knowledge translation is the use of knowledge in health care decision making. *J Clin Epidemiol.* 2011;64(1):6-10. https://doi.org/10.1016/j.jclinepi.2009.08.016
15. American Educational Research Association., *American Psychological Association., National Council on Measurement in Education., Joint Committee on Standards for Educational and Psychological Testing (U.S.). Standards for educational and psychological testing.* Washington, DC: American Educational Research Association; 2014.
16. Thorne SE. *Interpretive description: qualitative research for applied practice.* New York, NY: Routledge; 2016.
17. Marceau M, Gallagher F, Young M, St-Onge C. Validity as a social imperative for assessment in health professions education: a concept analysis. *Med Educ.* 2018;52(6):641-53. https://doi.org/10.1111/medu.13574
18. Braun V, Clarke V. *Thematic analysis: a practical guide.* Los Angeles, CA: SAGE Publications; 2022.
19. Messick S. *Validity.* In: Linn RL, editor. Educational Measurement. 3rd ed. New York, NY: Macmillan; 1989. p. 13-103.
20. Messick S. Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *Am Psychol.* 1995;50(9):741-49. https://psycnet.apa.org/doi/10.1037/0003-066X.50.9.741
21. Kane MT. *Validation.* In: Brennan RL, editor. Educational Measurement. 4th ed. Westport, CT: American Council on Education/Praeger; 2006. p. 17-64.