

## Eight ways to get a grip on validity as a social imperative Huit façons d'appréhender la validité en tant qu'impératif social

Mélanie Marceau,<sup>1</sup> Meredith Young,<sup>2</sup> Frances Gallagher,<sup>1</sup> Christina St-Onge<sup>1</sup>

<sup>1</sup>Faculty of Medicine and Health Sciences, Université de Sherbrooke, Québec, Canada; <sup>2</sup>Institute of Health Sciences Education, Faculty of Medicine and Health Sciences, McGill University, Québec, Canada.

Correspondence to: Mélanie Marceau, Faculty of Medicine and Health Sciences, Université de Sherbrooke. 3001 12th Avenue North, Sherbrooke, Québec, Canada, QC J1H 5N4; phone: 1 819 821 8000; 73162; email: [melanie.marceau@usherbrooke.ca](mailto:melanie.marceau@usherbrooke.ca)

Published ahead of issue: Jun 24, 2024; published: Jul 12, 2024. CMEJ 2024, 15(3) Available at <https://doi.org/10.36834/cmej.77727>

© 2024 Marceau, Young, Gallagher, 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.

### Abstract

*Validity as a social imperative* foregrounds the social consequences of assessment and highlights the importance of building quality into the assessment development and monitoring processes. *Validity as a social imperative* is informed by current assessment trends such as programmatic-, longitudinal-, and rater-based assessment, and is one of the conceptualizations of validity currently at play in the Health Professions Education (HPE) literature. This Black Ice is intended to help readers to get a grip on how to embed principles of *validity as a social imperative* in the development and quality monitoring of an assessment. This piece draws on a program of work investigating *validity as a social imperative*, key HPE literature, and data generated through stakeholder interviews. We describe eight ways to implement validation practices that align with *validity as a social imperative*.

### Résumé

La *validité en tant qu'impératif social* met de l'avant les conséquences de l'évaluation des apprentissages sur la société et souligne l'importance d'intégrer la qualité dans le développement et le monitoring de l'évaluation des apprentissages. La *validité en tant qu'impératif social* est influencée par les tendances actuelles en matière d'évaluation, telles que l'évaluation programmatique, longitudinale et l'évaluation par des évaluateurs. La validité en tant qu'impératif social fait partie des conceptualisations actuellement présentes dans les écrits scientifiques dans le contexte de la pédagogie des sciences de la santé. Ce texte a pour but d'aider les lecteurs à comprendre comment intégrer les principes de *validité en tant qu'impératif social* dans le développement et le suivi de qualité d'une évaluation. Cet article s'appuie sur un programme de recherche qui examine la *validité en tant qu'impératif social*, sur les écrits scientifiques en pédagogie des sciences de la santé et des données provenant d'entrevues avec différentes parties prenantes. Nous décrivons huit façons de mettre en œuvre des pratiques de validation qui sont en accord avec la *validité en tant qu'impératif social*.

### Introduction

Validation practices traditionally rely heavily on psychometrics, however these may fall short for assessment strategies such as programmatic, authentic, or narrative assessments. The complexities of validation practices that align with more recent assessment developments have led the Health Professions Education (HPE) community to consider the social and consequential aspects of validity—reflected in the concept of validity as a social imperative.<sup>1</sup> This emerging conceptualization of validity in the HPE literature includes the notion of purposefully building validation into the assessment development and monitoring processes at the outset.<sup>1,2</sup>

The concept of validity as a social imperative is composed of four attributes or characteristics: 1) validity evidence is seen as credible by society; 2) validity is built into the assessment process; 3) decisions depend on interpretation of the combination of assessment findings and 4) validity evidence includes both quantitative and qualitative data<sup>1</sup> (detailed in Table 1). Validity as a social imperative was identified in a discourse analysis<sup>2</sup> of the HPE literature, was later refined through a concept analysis,<sup>1</sup> and examined for feasibility and acceptability by stakeholders.<sup>3</sup>

While identified inductively from the HPE literature, validity as a social imperative shares some characteristics with other conceptualizations of validity (e.g., Mislevy<sup>4</sup> or Messick<sup>5</sup>). While validity as a social imperative has been well described,<sup>1,2</sup> there has been little guidance for how to implement validation practices that align with validity as a social imperative.

Here, we present eight ways to get a grip on validity as a social imperative throughout assessment development and quality monitoring (Table 1; expanded below). For each way to get a grip, we describe when they should be considered in assessment development and validation and map how to get a grip on the attributes of validity as a social imperative.

*Table 1. Description of ways to get a grip on implementing validation practices*

Validity as a social imperative			
Ways to get a grip	Target stages	Attribute	Definition of the attributes
1. Build a repository or documentation framework for collating validation data	<i>Planning the assessment</i>	Validity built into the assessment process	Validity evidence includes the justification of decisions made during the development and administration of an assessment, and the interpretation of assessment data. This evidence includes consideration of the potential consequences that the interpretation of the assessment scores could have on the individual, the institution, and society.
2. Explicitly document the purpose and stakes of the assessment			
3. Situate individual assessments in the larger educational and assessment plan			
4. Anticipate potential consequences and implementing mitigation strategies	<i>Planning the administration of the assessment</i>		
5. Identify resources available for validation.	<i>Planning the documentation of validity evidence</i>	Validity evidence includes both quantitative and qualitative data	Validity evidence must be collected using rigorous approaches, and attention should be paid to quantitative and qualitative data sources as legitimate validity evidence.
6. Choose validity evidence judiciously		Validity evidence seen as credible by society	Professional bodies (teaching institutions and professional orders) must be able to document, in a way that is perceived as credible by society, decisions made regarding a students' academic progress and their level of competence for embarking on an independent professional career.
7. Ensure the quality of scores collected support the intended score use	<i>Planning the interpretation of validity evidence and assessment scores</i>	Decisions depend on interpretation of the combination of assessment findings	Assessment data generated within an assessment program are often combined to make a final judgment. Greater emphasis should be given to validity evidence gathered for the overall assessment data interpretation – than evidences for individual tools or assessment episodes when the overall assessment data informs decisions about learners.
8. Critically review the quality of evidence documented through the validation			

## How to get a grip on validity as a social imperative

### 1. Build a repository or documentation framework for collating validation data

Document validation approaches and practices clearly and explicitly, with enough supporting documentation for later review for internal or external audiences. For example, use logbooks to record decisions about validation practices and validity evidence necessary to support the later interpretation and use of assessment-generated data.

### 2. Explicitly document the purpose and stakes of the assessment

Clearly and explicitly state and communicate the purpose and stakes of the assessment to all individuals involved in assessment and validation including learners, assessors,

institutions, and professional bodies. Validation practices<sup>6</sup> and the interpretation of assessment data are shaped by the assessment's purpose and stakes. Hence, evidence should match the intended score interpretation, whether for high-stakes (e.g., licensure) or lower-stakes (e.g., end of rotation) assessments.

### 3. Situate individual assessments in the larger educational and assessment plan

The logic for an assessment, including its place in the program, purpose, and use should be made explicit and archived. This deliberate positioning of an assessment within a larger educational and assessment context is particularly relevant for programs of assessment that favor multiple snapshots of learners' performance. It is important to explicitly document the justification for how these snapshots are taken, when these snapshots occur,

and how assessments and scores will be combined in a way that complements their inherent strengths and weaknesses.<sup>6</sup>

#### 4. Anticipate potential consequences and implementing mitigation strategies

Anticipate and minimize potential negative unintended consequences that can influence the quality of the data generated through assessment.<sup>7</sup> Exam administration should be carefully planned and monitored to reduce validity threats, for example, cheating compromises score validity by undermining the representation of students' abilities. Lack of transparency can also threaten validity of score interpretation, and transparency can be increased by being explicit about the assessment purpose and intended score use to all stakeholders, especially assessors and learners.

#### 5. Identify resources available for validation

Decisions about how to build and execute validation practices include considerations for financial or human resources (e.g., access to psychometricians, peers, or expertise), available time, and technical support (human or automatic scoring).<sup>6(p.35)</sup> Identify resources available for validation early in the assessment and validation planning process to ensure a purposeful, comprehensive, achievable, and responsible validation approach within the constraints of available resources.

#### 6. Choose validity evidence judiciously

Use available resources to help prioritize relevant validity evidence and in selecting a validity framework if appropriate. Downing,<sup>8</sup> for example, describes options for data that could be collected that align with each of the five evidences of validity included in the Standards for Educational and Psychological Testing.<sup>9</sup> Ensuring the quality of assessment often requires going beyond psychometric analyses as the sole measure of quality to ensure a broad representation of validity evidence. We do not suggest abandoning psychometric analyses—but rather suggest that evidence of validity may be quantitative (e.g., psychometric or statistical analysis), qualitative (e.g., stakeholder feedback), or a combination of both.<sup>6</sup> Be sure to be explicit, document, and justify the choice of data collected for each evidence of validity, including their complementarity.

#### 7. Ensure the quality of scores collected support the intended score use

Ensure that the decisions based on the interpretation of assessment data (e.g. continue in practice or not; identification of areas for improvement) align with the purpose and stake of the assessment (e.g. high stake for certification; low stakes for feedback on the progression of the learner). Be explicit about how scores will be used to support decisions and actions, or how they will be combined and interpreted. For example, questions such as “What should an administrator do when scores from different assessors or assessments diverge?” should inform data collection, approaches to validation, and validity evidence to be collected—all in order to support the score interpretation and use.

#### 8. Critically review the quality of evidence documented through the validation

Be transparent in documenting the decisions, justifications, and evidence collected to add credibility and defensibility to the entire assessment process, interpretation of scores, and decisions made based on assessment-generated data. Scrutinize and iteratively revise the data collected, as validity is not a property of a given assessment tool or program. Consider the documentation of validity evidence as a ‘living document’ that should be refined, critiqued, and adapted to situate validation within a continual quality improvement process.<sup>1,10</sup> Transparent and clear documentation of individual assessments and assessment programs makes it possible to publicly defend—to the learner and the society—the decisions resulting from the assessment scores.

## Conclusion

We hope that through describing ways to ‘get a grip’ on the concept of *validity as a social imperative*, we’ve provided concrete ways in which practices aligned with validity as a social imperative can be enacted. This piece describes considerations for how to apply *validity as a social imperative* to support those responsible for the planning, administering, and monitoring of assessments in planning for, collecting, and documenting validity evidence that could be perceived as credible by society. We hope that this encourages those working with assessments to reflect on their assessment development and validation practices in order to generate data that better supports decisions dependent on high quality assessment scores, while adapting these ways to get a grip to individual context and available resources.

**Conflicts of Interest:** Christina St-Onge is a member of the CMEJ Editorial team She has adhered to the CMEJ policy regarding authorship.

**Funding:** This study was supported by the *MEES-Universités programme* (in the form of a scholarship obtained by MM), the Social Sciences and Humanities Research Council (in the form of a scholarship provided by CSO and MY from grant no. 435- 2014-2159) and the *Paul Grand'Maison de la Société des médecins* Research Chair in Medical Education (in the form of a scholarship provided by CSO).

**Edited by:** Jane Gair (senior section editor); Marcel D'Eon (editor-in-chief)

## References

1. 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 Jun;52(6):641-53. <https://doi.org/10.1111/medu.13574>
2. St-Onge C, Young M, Eva KW, Hodges B. Validity: One word with a plurality of meanings. *Adv Health Sci Educ*. 2017 Oct 1;22(4):853-67. <https://doi.org/10.1007/s10459-016-9716-3>
3. Marceau M, St-Onge C, Gallagher F, Young M. Validity as a social imperative: Users' and leaders' perceptions. *Can Med Educ J*. 2022;13(3):22-36. <https://doi.org/10.36834/cmej.73518>
4. Mislevy RJ. Validity by design. *Educ Res* 2007;36(8):463-9. <https://doi.org/10.3102/0013189X07311660>
5. Messick S. Standards of validity and the validity of standards in performance assessment. *Educ Measure*. 1995;14(4):5-8. <https://doi.org/10.1111/j.1745-3992.1995.tb00881.x>
6. Cook DA, Brydges R, Ginsburg S, Hatala R. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Med Educ*. 2015 Jun;49(6):560-75. <https://doi.org/10.1111/medu.12678>
7. Ellaway RH, Topps D, Pusic M. Data, big and small: emerging challenges to medical education scholarship. *Acad Med*. 2019. p. 31-6. <https://doi.org/10.1097/ACM.0000000000002465>
8. Downing SM. Validity: On the meaningful interpretation of assessment data. *Med Educ*. 2003 Sep;37(9):830-7. <https://doi.org/10.1046/j.1365-2923.2003.01594.x>
9. AERA, APA, NCME. Validity. In: standards for educational and psychological testing. Washington: American Educational Research Association; 2014. p. 11-31.
10. Kane MT. Validation. In: Brennan R, editor. *Educational Measurement*. 4th ed. Westport: American Council on Education/Praeger.; 2006. p. 17-64.