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, Meredith Young 📌, Frances Gallagher et Christina St-Onge

Volume 15, numéro 3, 2024

URI : https://id.erudit.org/iderudit/1112778ar
DOI : https://doi.org/10.36834/cmej.77727

Résumé de l’article
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.
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, 1 Meredith Young, 2 Frances Gallagher, 1 Christina St-Onge 1

1 Faculty of Medicine and Health Sciences, Université de Sherbrooke, Québec, Canada; 2 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

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.

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. 1 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. 1, 2

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 1 (detailed in Table 1). Validity as a social imperative was identified in a discourse analysis 2 of the HPE literature, was later refined through a concept analysis, 1 and examined for feasibility and acceptability by stakeholders. 1
While identified inductively from the HPE literature, validity as a social imperative shares some characteristics with other conceptualizations of validity (e.g., Mislevy\(^4\) or Messick\(^5\)). While validity as a social imperative has been well described,\(^4,5\) 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

<table>
<thead>
<tr>
<th>Validity as a social imperative</th>
<th>Ways to get a grip</th>
<th>Target stages</th>
<th>Attribute</th>
<th>Definition of the attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planning the documentation framework for collating validation data</td>
<td>1. Build a repository or documentation framework for collating validation data</td>
<td>Planning the assessment</td>
<td>Validity built into the assessment process</td>
<td>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.</td>
</tr>
<tr>
<td>Explicitly document the purpose and stakes of the assessment</td>
<td>2. Explicitly document the purpose and stakes of the assessment</td>
<td>Validity evidence includes both quantitative and qualitative data</td>
<td>Validity evidence must be collected using rigorous approaches, and attention should be paid to quantitative and qualitative data sources as legitimate validity evidence.</td>
<td></td>
</tr>
<tr>
<td>Situate individual assessments in the larger educational and assessment plan</td>
<td>3. Situate individual assessments in the larger educational and assessment plan</td>
<td>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 student’s academic progress and their level of competence for embarking on an independent professional career.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anticipate potential consequences and implementing mitigation strategies</td>
<td>4. Anticipate potential consequences and implementing mitigation strategies</td>
<td>Validity evidence seen as credible by society</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identify resources available for validation.</td>
<td>5. Identify resources available for validation.</td>
<td>Planning the documentation of validity evidence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Choose validity evidence judiciously</td>
<td>6. Choose validity evidence judiciously</td>
<td>Planning the interpretation of validity evidence and assessment scores</td>
<td>Decisions depend on interpretation of the combination of assessment findings</td>
<td>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.</td>
</tr>
<tr>
<td>Ensure the quality of scores collected support the intended score use</td>
<td>7. Ensure the quality of scores collected support the intended score use</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Critically review the quality of evidence documented through the validation</td>
<td>8. Critically review the quality of evidence documented through the validation</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 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\(^4\) 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.6

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.7 Exam administration
should be carefully planned and monitored to reduce
validity threads, 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).6,9 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,8 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.9 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.6 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.1,10 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


