A contemporary approach to validity arguments:
a practical guide to Kane’s framework

Reference:
Cook DA¹, Brydges R², Ginsburg S³, Hatala R⁴. A contemporary approach to validity arguments: a practical guide to Kane’s framework. Medical Education. 2015 Jun;49 (6):560-75

Author Institutions:
¹Mayo Clinic Online Learning, Mayo Clinic College of Medicine, Rochester, Minnesota, USA
²Division of General Internal Medicine, Mayo Clinic, Rochester, Minnesota, USA
³Department of Medicine, University of Toronto, Toronto, Canada
⁴Wilson Centre, University Health Network, Toronto, Canada
⁵Department of Medicine, University of British, Vancouver, Canada

Tags
Clinical domain   Educational domain
General            Assessment
                  Undergraduate
                  (medical school)
                  (Post)graduate
                  (Residency training)

Background

One of the key components of CBME is programmatic assessment. Programmatic assessment requires longitudinal sampling of representative and authentic performance of trainees. Of course, hidden within this statement is a challenge facing training programs and certification bodies. How do you defend the validity of the high stakes decisions (i.e. summative assessment) that are the outputs of assessment programs?

Traditional psychometric approaches include:

1. Content – how the assessment items are created;
2. Criterion – how the scores correlated with gold/reference standards; and
3. Construct – how well an abstract concept relates to an concrete trait, where a theoretical link between abstract and concrete exists.

This approach to validity struggles with both the number of required samples to determine “competence” (e.g. there is a large number of competencies within a competency framework) and the measurement error of direct observation assessments typical of work-based assessment.
What to do? This paper offers a counterpoint to a traditional understanding of validity that builds on the work of Messick and Kane to provide a contemporary approach to validity. In essence, how can we know that our judgments are “true”.

**Purpose**

“To offer a practical introduction to the key concepts of Kane’s framework that educators will find accessible and applicable to a wide range of assessment tools and activities.”

**Type of paper**

Theory Paper

**Key Outcomes**

A number is not useful for teachers or learners. Qualitative data is equally valuable. Both are markers of performance that inform a judgment of competence.

Validity is about the system of data that defends the judgment of competence. Analyzing a system can be complex; a validity framework can improve clarity.

When using a validity framework, multiple sources of evidence that link multiple elements in the framework increase the defensibility of the judgment (e.g. assessment program).

A clear statement of how the assessment data is to be used (low v high stakes, formative v. summative, clinical context etc.) will help frame and prioritize the nature and scope of evidence required for a validity argument.

Kane’s framework includes:

1. **Scoring inferences**
   - Construction and implementation of an instrument
   - Data collection / data quality

2. **Generalization inferences**
   - Adequacy of sampling
   - Reproducibility

3. **Extrapolation inferences**
   - Design of instrument reflects authentic performance
   - Strength of relationship between instrument findings and authentic performance

4. **Implication inferences**
   - Interpretation justifies use / action based on instrument
   - Impact of the instrument on learner / program

In other words, how well does the sample data reflect unlimited data? How well does the instrument reflect a perfect instrument? How well does the instrument reflect the real world? Does the instrument have an impact on learning?
Some educators emphasize Messick’s validity framework that uses a similar argument approach but is organized differently.

**Key Conclusions**

The authors conclude "...validation is not an endpoint but a process. Stating that a test has been ‘validated’ merely means that the process has been applied, but does not indicate the intended interpretation, the result of the validation process or the context in which this was done. Secondly, validation ideally begins with a clear statement of the proposed interpretation and use (decision), continues with a carefully planned interpretation/use argument that defines key claims and assumptions, and only then proceeds with the collection and organisation of logical and empirical evidence into a substantiated validity argument. Thirdly, educators should focus on the weakest links (most questionable assumptions) in the chain of inference. Fourthly, in all of the clinical and educational examples cited herein, the Scoring, Generalisation and Extrapolation evidence is fairly strong; only when we attempt to infer actionable Implications, moving from the real world score to specific decisions, do important deficiencies come to light. For this reason, we believe that the Implications and associated decisions are ultimately the most important inferences in the validity argument."

**Spare Keys – other take home points for clinician educators**

This manuscript is a great example of translating theoretical concepts into highly relevant and applicable lessons for Clinician Educators. This author group’s previous work on simulation for assessment revealed a significant gap in the literature around the validity of tools and programs in practice. This paper bridges this gap in the literature by suggesting a workable solution.