



CAMBRIDGE ASSESSMENT

A proposed framework for evidencing assessment validity in large-scale, high-stakes international examinations

Stuart Shaw, Victoria Crisp and Nat Johnson
Cambridge Assessment
Email: crisp.v@cambridgeassessment.org.uk

Abstract

It is important for educational assessment bodies to demonstrate how they are seeking to meet the demands of validity in their tests. The approach to validity taken here assumes a 'consequentialist' view where the appropriacy of the inferences made on the basis of assessment results is seen as an essential feature of validity (Messick, 1989b). This paper describes the development of a comprehensive approach to the collection of evidence that can support claims about the validity of the interpretation of examination results and their uses. The framework was developed drawing on the validity literature of the past 50 years. Common elements were identified between existing models for validation and used to provide a list of components of validity to be evidenced. In the proposed framework these are indicated by a number of questions:

- What is (or are) the main declared purpose(s) of the assessment and are they clearly communicated?
- What are the constructs that we intend to assess and are the tasks appropriately designed to elicit these constructs?
- Do the tasks elicit performances that reflect the intended constructs?
- Do the tasks adequately sample the constructs that are important to the domain?
- Are the scores/grades dependable measures of the intended constructs?
- Is guidance in place so that teachers know how to prepare students for the assessments such that negative effects on classroom practice are avoided?
- Is guidance in place so that teachers and others know what scores/grades mean and how the outcomes should be used?
- Does the assessment achieve the main declared purpose(s)?

For each element various data would be gathered to provide 'evidence for validity' and to identify any 'threats to validity', including areas for improvement or further work. This paper describes the development of the proposed framework and how it would be used, including the types of methods that could be used to gather relevant evidence in order to provide a comprehensive evaluation of validity.



UNIVERSITY of CAMBRIDGE
Local Examinations Syndicate