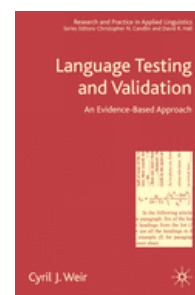


## Book Review

### Language Testing and Validation: An evidence-based approach

by Cyril J. Weir (1996) ISBN: 1-4039-1189-4

Houndgrave, Hampshire, UK: Palgrave-Macmillan (2005)



In his introduction, Weir states, "The core of this book is concerned with exploring a framework for establishing the validity of the interpretation of scores on tests produced by Exam Boards or by teachers for use in their classrooms." The "evidence-based approach" in the subtitle alludes to the view that the process of validating the use of scores on any given test is much like a courtroom trial – requiring evidence to support arguments for or against a favorable verdict - that the test was "fair".

In first of the book's four parts, Weir introduces five ways to collect the evidence needed to make a convincing case; the first two are collected a priori, or in the design phase of a test, and consist of defining what abilities the test is supposed to measure as well as how the sample of tasks in the test represents the abilities in 'the real world' (outside the test itself) that test users are looking for. The remaining three types of evidence are empirical, collected a posteriori as statistical procedures to estimate or enhance the reliability of the test scores, studies of how well test scores correlate with external criteria such as other tests purported to measure the same abilities or actual performance in the real world, and lastly, the study of the backwash or social consequences of test use for the stakeholders: teachers, students, parents, administrators, and "the marketplace".

Part two begins a detailed survey of frameworks for tests of reading, listening, speaking and writing. Each framework is presented as a flow-chart of boxes which detail the a priori considerations – test taker characteristics, test characteristics, theories of internal processes and resources – to the a posteriori considerations for investigating scoring characteristics, criterion-related evidence of score value and the impact of score interpretation. Six chapters present examples from actual research to illustrate how evidence of validity was obtained "in action".

Part three could serve as a syllabus for a practicum on test validation methodology. It contains pointers for sound research procedures, checklists and questionnaires to help researchers collect each of the five types of evidence for valid test score use. This part alone would justify buying the book, as it can certainly encourage teachers to not only have more confidence in their own assessment practices, but would probably result in some pretty interesting presentations or publications.

Part four of the book is entitled "Further resources in language testing", and is actually an up-to-date and comprehensive list of textbooks, journals, professional organizations, professional conferences, e-mail lists, bulletin boards and websites, databases and statistical packages – a list probably equal to a few years of word-of-mouth searches for answers to the questions that plague anyone who gets involved with testing and assessment.

In conclusion, this book is highly recommended. It is the most readable and comprehensive treatment of validation that I have ever come across. It is as free of technical jargon as can be; important points are presented clearly through Weir's choice to enclose poignant quotes and concepts in boxes within the text which saves on magic markers or underlining, in addition to the other features of this book that have already been reported. Its 284 pages contain not a single mathematical equation or algorithm, so it's not a cookbook for number crunching, but it'll tell you where to go for that if you need it – plus much, much more.

- reviewed by Jeff Hubbell

HTML: [http://jalt.org/test/hub\\_2.htm](http://jalt.org/test/hub_2.htm)

/

PDF: <http://jalt.org/test/PDF/Hubbell2.pdf>