



education review // reseñas educativas

editors: gene v glass gustavo e. fischman melissa cast-brede

a multi-lingual journal of book reviews

September 1, 2010

ISSN 1094-5296

Lissitz, Robert W. (Ed.) *The Concept of Validity: Revisions, New Directions and Applications*. Charlotte, NC: Information Age Publishing.

Pp. 270 ISBN 978-1607522270

Reviewed by Joel P. Wiesen
Scarsdale, New York

This fascinating, highly organized and readable, in-depth treatment of the concept known as validity presumes the reader has basic knowledge of the topic. The editor selected authors representing traditional, innovative and even radical views of the theory of test validity, as well as authors who report on four, diverse, large-scale testing programs, with emphasis on the application of validity theory. If you use or develop tests, this book may widen your horizons and introduce you to new conceptual views of validity, some quite provocative, and some that may be disquieting. This little book was meant to be provocative, and it is.

Each of the eleven chapters of this book is clear, crisp, authoritative, and thorough. The book reflects 10 presentations at a conference on validity sponsored by the Maryland State Department of Education. The first chapter

Citation: Wiesen, Joel P. (2010 September 1) Review of *The Concept of Validity* by Robert W. Lissitz (Ed.). *Education Review*, 13. Retrieved [Date] from <http://www.edrev.info/reviews/rev965.pdf>

explains that the purpose of the book is to present conflicting views on validity: mainstream and divergent. It provides an overview of the book that could, in itself, serve as a book review. Nevertheless, I summarize many of the major ideas below. The book is organized in three sections.

The first section, titled, “Relatively Mainstream,” contains four chapters, one each by Stephen Sireci, Michael Kane, Bruno Zumbo, and Robert Mislevy. Sireci presents a historical review of the concept of validity, arguing that validity long has been defined in terms of the inferences based on test scores rather than a property of the test itself, and making suggestions of ways the field can advance within this current, accepted view of validity. Kane, too, presents a historical review of test validity. He then presents three examples of the argument-based approach to validity, in which the test developer or user first identifies the assumptions and inferences of the proposed test use and then assembles evidence to evaluate their plausibility, discussing tests of observable attributes, operationally defined attributes, and trait interpretations, and the limitations of content validity. Zumbo presents the four elements of a contextualized, explanatory-focused, integrative cognitive judgement of validity: validity, utility, psychometric, and social consequences. He explores this model in light of historical views of validity, and goes on to define a multilevel construct as one that may be meaningful both for individual test takers and for various higher levels of aggregation, such as schools or school districts. Mislevy discusses psychometrics as model based reasoning, recognizing that all models are only approximations to the real world. He sees model based reasoning as embodying the four way interrelationships of a model, system, user and purpose. Together these chapters provide examples of the range of approaches to validity that accept and build on mainstream thinking concerning validity.

The section titled “Relatively Divergent” consists of two chapters, one by Joel Michell and the other by Denny Borsboom and colleagues. Michell suggests that an “attribute is measurable only if it possesses both ordinal and additive structure.” He holds that additive or quantitative

structure is often (always?) lacking, so test validation is a myth. Borsboom and colleagues critique and debunk construct validity, showing by example that it is a logically flawed concept, and presenting an alternative that is consistent with the view that a test is valid if it measures what it **should** measure. They make a strong case for the abandonment of the concept of construct validity.

The final section is titled “Application Oriented” and presents four, varied examples of large-scale testing programs. The chapters are authored by William Shafer and colleagues, Robert Linn, Krista Mattern and colleagues, and Michelle Chalhoub-Deville. Schafer reviews validity evidence provided by state testing programs in support of their Final Assessment System under the federal NCLB program, summarizes the validity information provided by five successful states in support of their peer reviews, and makes recommendations for improved validity evidence for state assessments. Linn discusses validity as it relates to NCLB and particularly validity evidence that can support determinations concerning Adequate Yearly Progress. Mattern and colleagues describe how they went out about validating the new (2005) SAT and how they communicated their findings to diverse user audiences. Finally, Chalhoub-Deville considers content validity in light of standards used to develop foreign language tests for non-native speakers of English.

In sum, this book treats the concept of validity both from theoretical and applied perspectives. Each chapter is rich in ideas, information, and references.

About the Reviewer

Joel P. Wiesen PhD of Applied Personnel Research is a psychometrician and consultant in Scarsdale, New York. He is a published test author. The tests he has developed include three diagnostic math tests for grade, middle, and high school students.

Copyright is retained by the first or sole author,
who grants right of first publication to the *Education Review*.
<http://edrev.asu.edu>



Editors

Gene V Glass

glass@edrev.info

Gustavo Fischman

fischman@edrev.info

Melissa Cast-Brede

cast-brede@edrev.info
