



education review // reseñas educativas

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

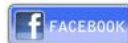
a multi-lingual journal of book reviews

May 5, 2011

ISSN 1094-5296

Education Review/Reseñas Educativas is a project of the National Education Policy Center <http://nepc.colorado.edu>

Follow Education Review on Facebook



Ekbatani, Glayol (2010) *Measurement and Evaluation in Post-secondary ESL*. NY: Routledge, Taylor and Francis Group.

Pp. 112 ISBN 978-0-8058-6126-6

Reviewed by Mary Speer
University of Texas, Austin

Developing qualitative assessments for language learners is a formidable challenge. For this reason, production of an optimal test and the interpretation of the results must be the culmination of a systematic and balanced approach guided by frameworks that highlight principles of reliability, validity, and test usefulness (see Bachman and Palmer, 1996). Glayol Ekbatani explores the field of testing in a simplified and accessible manner by offering practical application advice and examples to assist in creating test material consistent with the constructs of assessment in her book entitled, *Measurement and Evaluation in Post-secondary ESL*.

The intended purpose of this text is to provide a brief practical guide to assessment- related topics in the field of second language education by defining technical terminology and providing practical examples, discussing

Citation: Speer, Mary. (2011 May 5) Review of *Measurement and Evaluation in Post-secondary ESL* by Glayol Ekbatani. *Education Review*, 14. Retrieved [Date] from <http://www.edrev.info/reviews/rev1077.pdf>

effective methods in assessing the reading, writing, and oral proficiency of post-secondary ESL students, and addressing lacunae in the existing literature by offering innovative methods of testing. The author hopes to achieve the outcome of the production of well-informed testers and testing practices.

Measurement and Evaluation in Post-secondary ESL is meant to be an introduction for teachers and administrators interested in testing college bound international students in TESOL programs in the US and around the world. Due to its introductory and informative nature, it could also be valuable to teachers who have no previous technical knowledge of assessment, and its easy readability would even allow for an expanded audience of graduate and undergraduate education classes. This audience would then be able to evaluate, interpret, and choose an appropriate test for placing students in their classrooms or institutions.

With the goal to provide examples of the application of testing constructs in mind, the author divides the book into three sections. The first section briefly outlines and defines basic qualities of test usefulness, the second section provides an overview of relevant statistical methods for scoring and interpretation, and the final section of the book offers suggestions that can help test-makers write test questions for reading, writing, and oral proficiency. Throughout the text, the author formats key concepts in bold to serve as quick references for readers. She also includes an appendix that describes and updates the new techniques being implemented in the three most widely used standardized test for measuring post-secondary ESL performance: MELAB, IELTS, and TOEFL.

The author is providing a fresh look at old concepts; though for specialists who are already familiar with technical terms, the explanations she offers prove to be redundant. There is already a comprehensive explanation of validity, reliability, construct definition, wash back, reliable scoring, and rater training in Bachman and Palmer among others. However, from a non-specialist point of view this is actually a strength of the book because it simplifies this terminology and provides examples from outside of the testing field to draw upon in aiding reader comprehension. For example, the term validity as related to the test-taking field refers to the extent to which a test

measures what it purports to measure (Bachman and Palmer, 1996). To simplify this concept, Ekbatani defines the term validity by explaining its Latin origin and its historical roots in the law. These contexts allow the reader to associate the term validity with definitions like efficacious and well-grounded so that they have more of an idea of how the term can be applied when evaluating a test. She also provides the reader with a concrete example of a grammar test which claims to examine a student's writing ability and explains that this kind of test would not be valid as the measure of the ability to write clear, grammatically accurate prose because it tests grammar not writing. Clarifications and examples like these make testing terminology more accessible to readers with no prior knowledge of the testing field.

Another strength of *Measurement and Evaluation in Post-Secondary ESL* is its specific instruction on how to create task-based test questions to demonstrate the construct being tested. For example, in the chapter on reading comprehension and vocabulary, Ekbatani refers to lists of reading skills one could assess, introduces reading models presented by theorists, and offers techniques for writing multiple-choice, short response, and extended response questions. Each question type is accompanied by an example question and a list of tips for writing the question. Although many of the tips are common sense, some of them can help test-makers fall more accurately within the boundaries of validity, reliability, and authenticity by considering appropriateness of level, using original wording, and following a rubric to score the test. More importantly a teacher or administrator can choose a test that best exhibits these qualities in order to make the best decision about placement of students in the appropriate level or program.

Ekbatani also offers advice about what combination of question types would produce a balance of validity and reliability in a given test by offering an explanation about why multiple choice questions tend to be more reliable while constructed response questions tend to be more valid. Advice like this would help an administrator trying to choose a test because it would enable her to choose based on the quality that is more important in the context in which she would be testing. It might also lead her to

choose a test that elicits both kinds of responses from test-takers.

Finally the author provides a comprehensive and specific rubric to aid in the scoring of a student's level and suggests that beginning test scorers agree on criteria for assessing constructed response items to maximize reliability. The rubrics are comprehensive and specifically worded to help the raters in accurately scoring the test. This part of the book would be particularly useful for raters who work for a testing company or teachers who must rate their students' performance on a particular test. One addition that would strengthen this chapter of the book would be the inclusion of poorly written question examples that do not follow the guidelines presented by the author. This would allow the test-makers to make a clear distinction between an acceptable and an unacceptable test question which would clarify the boundaries of appropriateness when writing new questions.

Though the book is not as comprehensive as Bachman and Palmer (1994) treatment of test usefulness and development, its explanation of theoretical concepts in the field of testing does achieve the author's intended objective of informing testers and professionalizing test practices. The rubrics are detailed and specific, and the examples can be used as models for testers who would like to create their own tests. While some of the advice about how to create a test question may seem to be common sense, following the formulaic recipe could make novice test makers more confident and professional when writing their own questions.

One strength of the book is its discussion of the innovations of the new task-based and integrative approaches applied to recent standardized tests. Ekbatani explains that the integrative approach is a response to the recent quest for higher rates of authenticity- that is that test tasks best reflect real life situations that students must be able to manage in order to succeed in a college setting. She also advises readers to use task based and constructive responses rather than selected responses because it increases the authenticity of the test. Academics' recent preoccupation with authenticity is cutting edge information that should be brought to the attention of all

teaching professionals because it has changed certain areas on the TOEFL and other tests and could lead to more changes in the future to assessment methods like portfolios and self-assessment.

Alternately, one weakness of the book is that the alternative forms of assessment were introduced in a limited manner. Discussion of alternative assessments occupied only two pages of the book although the cover claims that special emphasis is placed on them. In these two pages, Ekbatani briefly introduces writing portfolios and self-assessment as new methods and then cites other books which investigate these practices. Greater detail regarding these forms of assessments may have allowed readers to benefit from this knowledge without having to refer to another text. Overall, the author does accomplish her goals of informing readers of professional testing practices and making them accessible to anyone interested in testing. The fresh treatment of concepts mentioned on the cover was characterized in the text by simplification and usage of non-technical explanations. The methods presented appear effective and innovative because of the new emphasis placed on higher levels of authenticity, and the appendix updates the changes made on the most common tests allowing readers to make informed decisions about what tests could be used to place students in their classrooms. The value of this book lies in its ability to introduce new trends in testing and make testing terminology accessible to non-experts who want to make or choose an appropriate assessment for their incoming college students.

Reference

Bachman, L. & Palmer, A. (1996) *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford: Oxford University Press.

About the Reviewer

Mary Speer is a Masters student in foreign language education at the University of Texas. Her interests include Spanish literature, traditional Chinese, English literature, foreign language pedagogy, second language acquisition, socio-linguistics, and language assessment.

Copyright is retained by the first or sole author, who grants right of first publication to the *Education Review*.

Education Review/Reseñas Educativas is a project of the National Education Policy Center <http://nepc.colorado.edu>



Editors

Gene V Glass

glass@edrev.info

Gustavo Fischman

fischman@edrev.info

Melissa Cast-Brede

cast-brede@edrev.info
