

Recension / Book Review

Robert L. Johnson, James A. Penny, and Belita Gordon. (2009). *Assessing Performance: Designing, Scoring, and Validating Performance Tasks*. New York, NY: The Guilford Press. 354 pages. ISBN: 978-1-59385-988-6

Laurie-ann Hellsten is an associate professor in the Department of Educational Psychology and Special Education, College of Education, University of Saskatchewan.

Assessing Performance: Designing, Scoring, and Validating Performance Tasks is a good addition to my collection. In this book, the authors attempt to “bridge the literature across disciplines to provide a guide for assessment practitioners as they develop performance assessments” (p. v). According to the authors, the intended audience of *Assessing Performance* are staff in testing programs, licensure and certification programs, new content specialists, program evaluators, researchers, and graduate students. Although the authors do state that the book would be useful for a class in performance assessment, due to its almost exclusive focus on large-scale, standardized, and certification/or licensure assessments, I argue that this text would need to be supplemented by a resource with a focus on performance assessment in the K-12 classroom. Furthermore, due to the depth and complexity of material presented, the book is best suited for graduate students with some background in statistics and measurement theory.

In ten chapters, the authors lead us through the performance assessment process from initial design (What is performance assessment? Why would you use performance assessment?) to designing, administering, scoring, and analyzing results. Each chapter includes concrete examples from existing assessments and a synthesis of recent and seminal

research on the topic (e.g., Crocker & Algina (1986) - test theory; Cronbach (1951) – reliability; Lord & Novak (1968) - measurement theory; and Messick (1993) validity). Additional pedagogical features at the end of each chapter include a checklist of activities to complete for each stage of development, summary questions, and additional readings. The text also includes useful tables and figures, a glossary, and more than 300 references.

Chapter 1, “An Overview of Performance Assessment” defines performance assessment and provides a multitude of examples of existing performance assessments. Chapter 1 also includes a section on the origins of performance assessment which serves to demonstrate that performance assessment is not a new phenomenon. Through their examples, the authors are able to show how pervasive performance assessment is in society and how performance assessments touch our daily lives; for example, music auditions (think *Canadian Idol*) are a form of performance assessment. Such examples also illustrate the natural fit of performance assessment in a variety of contexts and serve to “sell” the reader on the need for performance assessment.

Chapter 2, “Designing the Assessment” describes what a performance assessment might look like and includes a review of the purpose of the assessment, the development of an assessment framework, and an introduction to psychometric properties. In this and later chapters, the authors demystify the language of assessment by defining key terms commonly used in assessment such as “knowledge, skills, and abilities (KSA’s), and subject matter experts (SME’s). Furthermore, it is in this chapter where the authors introduce the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). These standards are appropriately referred to throughout the remaining chapters.

Chapter 3, “Construction of a Performance Task” describes six elements of performance task construction (e.g., structure of the task) alongside the decisions developers need to make for each element (such as issues related to sensitivity and fairness). Chapter 4, “An Example of Constructing a Performance Assessment” provides an example of devel-

oping a writing assessment beginning with a review of language arts content standards based on the state of Georgia's Grade 8 Writing Test.

Chapter 5, "Administration" describes how a performance assessment is administered to examinees and includes topics such as training staff to ensure standardization of procedures and test security. This chapter also contains an introduction to *Scott*, a manager at a company that develops and delivers examinations for use in certification and licensure. Scott is referred to throughout the remaining chapters and through the description of his experiences provides readers with an insiders' view of the administration of performance assessments.

Chapter 6, "Developing Tools for Scoring" describes how to develop tools (such as holistic and analytic rubrics) for scoring performance assessment and how to select benchmark performances. Examples of scoring tools are provided as well as an example of rubric development. Chapter 7, "Training Raters and Staff" focuses on the need for reliable raters and includes information on how to design rater training sessions, evidence raters should be looking for when reviewing performances, and examples of training scripts. Together chapter 6 and 7 show how complex performances can be rated consistently.

Chapter 8, "Scoring and Monitoring" describes how to monitor the quality of the ratings received, statistics that can be calculated to examine rater agreement and accuracy, and score resolution models. Chapter 9, "Forming Scores and Item-Level Analyses" describes methods of forming operational scores, examining task difficulty, estimating score reliability, and examining the validity of the task. The Oral Clinical Examination (American Board of Pediatric Dentistry, 2000) is used to illustrate the procedures. Chapter 10, "Test-Level Analyses" describes the statistical methods that can be used to provide information about the reliability, generalizability, and validity of the test as a whole. The authors provide examples in the form of data tables and statistical output from the EXCEL and SPSS computer programs as well as the relatively unfamiliar generalizability program, GENOVA.

The greatest strength of this book is that performance assessment is put into the context of "real life" (e.g., pilots, lawyers, dentists, etc.). Such examples provide alternatives to the typical K-12 education examples found throughout other assessment books. However, despite textual ref-

erences to these “real life” performance assessments, the majority of extended examples are still from the K-12 education system. Another limitation of the book for the Canadian reader is the American focus, and the repetitive use of examples from K-12 state testing programs. The detailed information presented on the administration of performance assessments via the portrayal of Scott is another asset of the book. Although researchers and students are often familiar with the development of assessments they tend to be much less familiar with the actual administration process. Through Scott, readers are provided with an authentic sense of the trials and tribulations of assessment administration, clues to what can go wrong, security measures taken, and the careful planning and organization required for the successful administration of an assessment. [Word count 1018]

REFERENCES

- American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME) (1999). *Standards for educational and psychological testing*. Washington, DC: AERA.
- Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. Fort Worth, TX: Holt, Rinehart, & Winston.
- Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Lord, F. & Novak, M. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Messick, S. (1993). Validity. In R. Linn (Ed.), *Educational measurement*, (#rd ed., pp. 13-103). Washington, DC: American Council on Education.