Digital Module 07: Subscores: Evaluation & Reporting

Sandip Sinharay, ETS

Available in the ITEMS Portal at https://ncme.elevate.commpartners.com

Module Overview

Test score users often demand the reporting of subscores due to their potential diagnostic, remedial, and instructional benefits. Therefore, there is substantial pressure on testing programs to report subscores. However, professional standards require that subscores have to satisfy minimum quality standards before they can be reported. In this digital ITEMS module, Dr. Sandip Sinharay reviews the status quo on the reporting of subscores. Specifically, he first provides examples of operationally-reported subscores, discusses why subscores are in high demand, and discusses professional quality standards that subscores have to satisfy. He then describes various statistical methods that can be used to evaluate whether subscores satisfy professional standards, which include descriptive statistics, DIMTEST / DETECT, factor analysis, multidimensional item response theory, and the Haberman method. He provides guidance for how to implement these methods on real data using the *R* package 'subscores'. The digital module includes video illustrations using real data, quiz questions with diagnostic feedback, data-based activities with video solutions, curated resources, and a glossary.

Key words: Diagnostic scores, disattenuation, DETECT, DIMTEST, factor analysis, multidimensional item response theory (MIRT), proportional reduction in mean squared error (PRMSE), reliability, subscores

Prerequisite Knowledge

This ITEMS module assumes no prior knowledge of the psychometric issues related to the reporting of subscores. However, to get the most out of this module, it might be beneficial to have a basic understanding of:

- Classical test theory (including familiarity with true scores, parallel forms, and reliability)
- Unidimensional and multidimensional tests
- Unidimensional and multidimensional IRT models

Studying the following NCME ITEMS modules may serve as a useful introduction to the prerequisite knowledge:

- Digital Module 01: Reliability in Classical Test Theory (Lewis & Chajewski, 2018)
- Module 32: Subscores (Sinharay, Puhan & Haberman, 2011)

These modules and others are available for free in the ITEMS portal.

Learning Objectives

Upon completion of this ITEMS module, learners should be able to:

- Understand what subscores are and why they are of interest to test users
- Appreciate the need to assess the quality of subscores
- Know about key methods to assess the quality of subscores
- Interpret results from key methods to real data
- Apply the Haberman method using the *R* package 'subscores'
- Decide whether to report subscores for a test

Module Structure

The digital module is divided into the following sections, which can be reviewed sequentially or independently [*approximate completion times in parentheses*].

- Module Introduction [*5 Minutes*]
- Section 1: Conceptual Foundations [*10 Minutes*]
- Section 2: Multivariate methods [*15 Minutes*]
- Section 3: Haberman method (with worked example) [*30 Minutes*]
- Section 4: Data Activity [*30 Minutes*]
- Section 5: Quizzes [*10 Minutes*]

In the ITEMS portal site, you can also find a video version of the core content as well as a handout with all core slides along with other materials.

Module Components

This ITEMS module includes the following components, which are delivered within a web-delivered unified design shell that is compatible across platforms (i.e., laptops, desktops, tablets, cell phones) and was created with modern course development software (*Articulate 360*):

- integrated content slides that provide a structured walk-through of the content with suitable voice-over
- embedded didactic videos to demonstrate software implementations
- interactive quiz questions with diagnostic feedback
- data activity with sample data, access to *R* code, and video-based solutions
- glossary of key terms
- supplementary digital resources

Additional materials may be added over time so check back periodically!

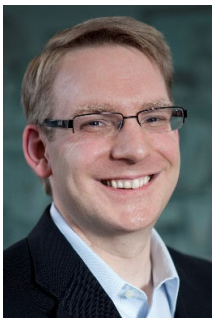
Content Developers

Sandip Sinharay, *Principal Research Scientist, ETS*

Sandip is a Principal Research Scientist at ETS. He received his Ph.D. and M.S. degrees from the Department of Statistics at Iowa State University. He has received five awards from the National Council on Measurement in Education including the Bradley Hanson Award (2018), the Technical or Scientific Contributions to the Field of Educational Measurement (2009 and 2015), the Jason Millman Promising Measurement Scholar Award (2006), and the Alicia Cascallar Award for an Outstanding Paper by an Early Career Scholar (2005). Sandip has coedited two published volumes and authored or coauthored more than 100 articles in peer-reviewed statistics and psychometrics journals and edited books. His research interests include statistical methods for detecting test fraud, reporting of subscores, Bayesian statistical methods, as well as model-data fit and model selection methods. The collaboration with the instructional design team on this project was a unique learning experience for Sandip.

Instructional Design Team

André A. Rupp, *Research Director at the Educational Testing Service (ETS)*



André is a research director in the Psychometrics, Statistics, and Data Sciences area at ETS. He is the co-author and co-editor of two award-winning interdisciplinary books entitled *Diagnostic Measurement: Theory, Methods, and Applications* (2010) and *The Handbook of Cognition and Assessment: Frameworks, Methodologies, and Applications* (2016); he is currently working on the *Handbook of Automated Scoring* (2020). His research synthesis- and framework-oriented work has appeared in a wide variety of prestigious peer-reviewed journals. Among other things, he is passionate about improving processes for interdisciplinary collaborations during the development and implementation of scoring solutions for digitally-delivered assessments. Consequently, he is very excited to serve as the associate editor / lead instructional designer of the ITEMS portal for NCME whose mission is to provide free digital resources to support self-directed learning and professional development.

Xi Lu, *Doctoral Candidate at Florida State University*

Xi is a doctoral candidate in the Instructional Systems and Learning Technologies program at Florida State University. Her current research interest focuses on designing and developing optimal learning supports to facilitate STEM learning in digital interactive environments. She also works as a research assistant with Dr. Val Shute on an NSF project targeted at designing various learning supports for a 2D physics game called *Physics Playground* to help middle school kids learn physics. Before coming to FSU, Xi taught Chinese for six years in Monterey Bay, California.
