
Digital Module 25: Testlet Response Theory

Hong Jiao, University of Maryland, College Park

Manqian Liao, Duolingo, Inc.

Module Overview

In this digital ITEMS module, Dr. Hong Jiao and Dr. Manqian Liao describe the *testlet response theory* for the construction and evaluation of new measures and scales. Nowadays, testlets are common test construction units in practice. The use of testlets help to create an authentic context to assess higher-order thinking skills. In innovative assessment such as game-based or simulation-based assessment, testlets are the building blocks for creating a situational context for an assessment purpose. In this ITEMS module, Drs. Jiao and Liao start with an introduction to the needs of testlets when local item dependence is present. They then introduce the basic testlet response models proposed from different theoretical frameworks that are built upon standard item response theory models. Furthermore, they introduce different methods for model parameter estimation and related software programs that implement them. Finally, they introduce further extensions to showcase extended testlet response models for more complex local item dependence issues in innovative assessment. The module is designed for students, researchers, and data scientists in various disciplines such as psychology, sociology, education, business, health and other social sciences in developing testlet-based assessment. It contains audio-narrated slides, sample data, syntax files, diagnostic quiz questions, data-based activities, curated resources, and a glossary.

Key words: Bayesian estimation, innovative assessment, item response theory, local item dependence, multi-part items, paired passages, testlets, testlet response theory

Prerequisite Knowledge

This ITEMS module assumes no prior knowledge of testlet response theory. However, it is probably helpful to have a working knowledge of foundational assessment and statistical concepts such as:

- Foundations of unidimensional item response theory
- Basic mathematical transformation functions (e.g., logarithms, exponentiations)
- Foundations of Bayesian estimation

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

- Print Module 7: Comparison of 1-, 2-, and 3-Parameter IRT Models
- Print Module 10: Rasch Measurement Theory
- Print Module 16: Comparison of Classical Test Theory and Item Response Theory
- Digital Module 06: Posterior Predictive Model Checking
- Digital Module 11: Bayesian Psychometrics

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:

- Describe key components of testlet response theory
 - Understand key perspectives to conceptualize the testlet effects
 - Apply testlet response theory to scale construction
 - Perform testlet response model analysis using select computer programs
 - Interpret the nature of testlet effects and the associated model parameters
 - Understand strategies for developing new testlet models
-

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 [*15 Minutes*]
- Section 2: Basic Testlet Response Theory [*30 Minutes*]
- Section 3: Model Parameter Estimation [*20 Minutes*]
- Section 4: Software Programs [*20 Minutes*]
- Section 5: Advanced Testlet Models [*20 Minutes*]
- Section 6: Quizzes [*20 Minutes*]

In the 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
- embedded didactic videos to demonstrate software implementations
- interactive quiz questions with diagnostic feedback
- sample data, syntax file, and access to the OpenBugs program
- data-based activities with video solutions
- glossary of key terms
- supplementary digital resources

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

Instructors

Hong Jiao, *Professor at the University of Maryland, College Park*



Hong Jiao is currently a professor at the University of Maryland (UMD), College Park specializing in educational measurement and psychometrics in large-scale assessment. She received her doctoral degree from Florida State University. Prior to joining the faculty in Measurement, Statistics, and Evaluation at UMD, she worked as a psychometrician at Harcourt Assessment on different state assessment programs. The overarching goal of her methodological research is to improve the practice in educational and psychological assessment and develop solutions to emerging psychometric challenges. Many of these are due to the use of more complex innovative assessment formats. Two areas of her research include methodological research on local dependence due to the use of testlet and

Bayesian model parameter estimation. Her methodological research has been recognized by a national award, academic work including numerous edited books, book chapters, refereed journal papers, and national and international invited and refereed presentations and different research grants and contracts on which she serve as PI or CO-PI. Hong Jiao proposed a multilevel testlet model for mixed-format tests that won the 2014 Bradley Hanson Award for Contributions to Educational Measurement by the National Council on Measurement in Education.

Manqian Liao, *Psychometrician at Duolingo, Inc.*



Manqian Liao is a psychometrician at Duolingo where she conducts validity and fairness research on the Duolingo English Test. Among other things, she has worked on investigating the differential item functioning in the Duolingo English Test items and evaluating the efficiency of the item selection algorithm. Manqian received her Ph.D. degree in Measurement, Statistics and Evaluation from University of Maryland, College Park. Her research interest focuses on item response theory (IRT) and diagnostic classification models (DCM). Her dissertation is on modeling multiple problem-solving strategies and strategy shift with DCM.

Instructional Design Team

Jonathan Lehrfeld, *Psychometrician at Educational Testing Service (ETS)*



Jon graduated from Fordham University in 2016 with a Ph.D. in psychometrics and quantitative psychology, where his dissertation focused on integrating propensity score methods with structural equation modeling. After graduating, he worked at the Council for Aid to Education (CAE) for three years, serving as their psychometrician and Associate Director of Measurement Science. While at CAE, his operational and research work focused on practical problems in low-stakes testing. He most recently joined ETS where he currently works as a psychometrician on a large-scale state assessment team.

André A. Rupp, *Mindful Measurement*



André is 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: Theory into Practice*. His synthesis- and framework-oriented research 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 developer of the ITEMS portal for NCME whose mission is to provide free digital resources in educational measurement to support self-directed learning and professional development.

This is the pre-peer reviewed version of the following article: Jiao, H., & Liao, M. (2021). Digital Module 25: Testlet response theory. Educational Measurement: Issues and Practice, 40(2). It has been published in final form <https://onlinelibrary.wiley.com/journal/17453992>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.
