ITEMS
Instructional Topics in
Educational Measurement Series

NCME national council on measurement in education

Digital Module 03: Nonparametric Item Response Theory
Stefanie A. Wind, University of Alabama
Available in the ITEMS Portal at https://ncme.elevate.commpartners.com

Module Overview

In this digital ITEMS module we introduce the framework of nonparametric item response theory (IRT), in particular Mokken scaling, which can be used to evaluate fundamental measurement properties with less strict assumptions than parametric IRT models. We walk through the key distinction between parametric and nonparametric models, introduce the two key nonparametric models under Mokken scaling – the monotone homogeneity and double monotonicity model – and discuss modern extensions of the basic models. We also describe how researchers and practitioners can use key nonparametric statistics and graphical visualization tools to evaluate the properties of an assessment from a nonparametric perspective. Finally, we illustrate the key reasoning steps and associated best practices using video-based worked examples using the *mokken* package in *R*.

Key words: Nonparametric IRT, Mokken scaling, monotone homogeneity model, double monotonicity model, rater effects, multilevel modeling, mokken package, R

Prerequisite Knowledge

This ITEMS module assumes very little a priori technical knowledge of item response theory. However, it is probably helpful to have a working knowledge of foundational assessment and statistical concepts such as:

- basic components (e.g., assessments / tests, items / tasks, total scores / scale scores)
- basic variable types and scales (e.g., continuous, discrete, dichotomous, polytomous)
- basic summary statistics (e.g., mean, variance, standard deviation, correlation)
- basic model parameters (e.g., latent variables, item parameters)

Specific prior experience with specialized estimation software is not required. However, we use the *mokken* package in *R* for worked examples but provide introductory videos for how to install this package and prepare for data analyses.

Learning Objectives

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

- Understand the relationship between parametric and nonparametric IRT models
- Understand the key differences between the two core nonparametric models
- Understand the key ideas of extensions of nonparametric models
- Perform real-data analyses with computational routines in *R*
- Estimate scaling coefficients and conduct hypothesis tests with them
- Apply best practices for item analysis within a nonparametric framework

After completion of this module, learners might wish to take additional ITEMS modules on parametric item response theory, classical test theory, or related scaling approaches. Check out the ITEMS Portal webpage for up-to-date information on available ITEMS modules!

Module Structure

The 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: Scaling Principles [*10 Minutes*]
- Section 2: Mokken Models [*10 Minutes*]
- Section 3: Model Extensions [*10 Minutes*]
- Section 4: Data Analysis [*15 Minutes*]
- Section 5: Worked Examples [*30 Minutes*]
- Section 6: Quizzes & Data Activity [*30 Minutes*]

Module Components

This ITEMS module includes the following components, which are delivered within a 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 and video components;
- worked examples using the *mokken* package in *R* with accompanying video;
- interactive quiz questions with diagnostic feedback covering the key aspects of the module;
- performance-based exercises with video solutions;
- a glossary of key terms; and
- a small library of online resources that have been vetted for general scientific accuracy and instructional utility.

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

Content Developer

Stefanie Wind is an Assistant Professor of Educational Measurement in the Department of Educational Research at The University of Alabama. She conducts methodological and applied research on educational assessments, with an emphasis on issues related to raters, rating scales, Rasch models, nonparametric item response theory, and parametric item response theory.

Instructional Design Team

André A. Rupp is a Research Director at the Educational Testing Service (ETS) in Princeton, New Jersey. 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). 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 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's 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.