_____

Digital Module 27: Hierarchical Rater Models
Jodi M. Casabianca, Educational Testing Service
Available in the ITEMS Portal at https://ncme.elevate.commpartners.com

_____

Module Overview

In this digital ITEMS module, Dr. Jodi M. Casabianca provides a primer on the *hierarchical rater model* (HRM) and the recent expansions to the model for analyzing raters and ratings of constructed responses. In the first part of the module, she establishes an understanding of the nature of constructed responses, the rating process and the common rater errors or *rater effects* originating from the rating process. In the second part of the module, she compares traditional measures for analyzing raters to *item response theory* (IRT) measures and discusses various IRT rater models. In the third section, she discusses longitudinal and multidimensional extensions and their foundations in different combinations of IRT and signal detection models. The module contains audio-narrated slides, quizzes with feedback, and additional resources such as a glossary and reference documents.

Keywords: hierarchical rater model, item response theory, longitudinal models, rater effects, multidimensional models, signal detection models

Prerequisite Knowledge

In order to maximally benefit from the introductory section of this digital ITEMS module, learners should have familiarity with approaches to constructed response scoring, rater reliability, and agreement. In order to maximally benefit from the more technical descriptions of the HRM and its extension in certain places, learners should have ideally:

- completed a two-semester graduate level applied statistics course series (or equivalent)
- completed a two-semester graduate level psychometric theory course series (or equivalent)
- completed a graduate level item response theory course (or equivalent)

However, multiple representations and explanations are given whenever possible to facilitate learning.

Learning Objectives

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

- describe the characteristics of rater effects and their impact on test taker measures
- understand the main advantage of the HRM over other IRT rater models
- understand scenarios for which the HRM framework is appropriate and inappropriate
- understand the HRM framework components and be able to conceptualize special cases relevant to specific datasets.
- discuss the HRM-based rater parameters and describe how they behave under different rater behavior

Module Structure

The digital ITEMS 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: Introduction to Analysis of Raters and Ratings [*20 Minutes*]
- Section 2: IRT Models for Rating Data [*15 Minutes*]
- Section 3: The Basics of the HRM Framework [*30 Minutes*]
- Section 4: Advanced Cases of the HRM Framework [*35 Minutes*]

Those interested in advanced cases of the HRM framework should review Section 3 before proceeding to Section 4 unless they already have basic knowledge of the HRM.

Module Components

This digital 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, and cell phones) and was created with modern course development software (Articulate 360):

- integrated content slides that provide a structured walkthrough of the content with voiceover
- quiz questions with feedback
- glossary of key terms
- reference documents

Additional materials may be added over time, so please check back periodically.

Instructor

Jodi M. Casabianca, *Educational Testing Service*

Jodi M. Casabianca is a measurement scientist at ETS. Jodi received a B.A. in both Statistics and Psychology (2001) and an M.S. in Applied and Mathematical Statistics (2004) from Rutgers University as well as an M.A. (2008) and a Ph.D. (2011) in Psychometrics and Quantitative Psychology from Fordham University. She was the 2009 recipient of the Harold Gulliksen Dissertation Research Fellow awarded by ETS. Before joining ETS in 2016, she was an assistant professor at the University of Texas at Austin's Department of Educational Psychology where she taught graduate courses and led a research lab in the Quantitative Methods program. Dr. Casabianca was awarded a National Science Foundation grant in 2013 (as co-PI, with Professor Junker) to expand the HRM into a framework that is more flexible for different assessment scenarios and has provided two conference workshops on the HRM in 2017 and 2018. Her current research focuses on statistical and psychometric modeling for constructed responses, with a focus on the evaluation of automated scoring models. Her research has been published in several notable peer reviewed journals, including the *Journal of Educational and Behavioral Statistics* and the *Journal of Educational Measurement*.

―――

Instructional Designer

André A. Rupp, *Mindful Measurement*

André 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) and has just published the *Handbook of Automated Scoring: Theory into Practice* (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.

―――

―――