

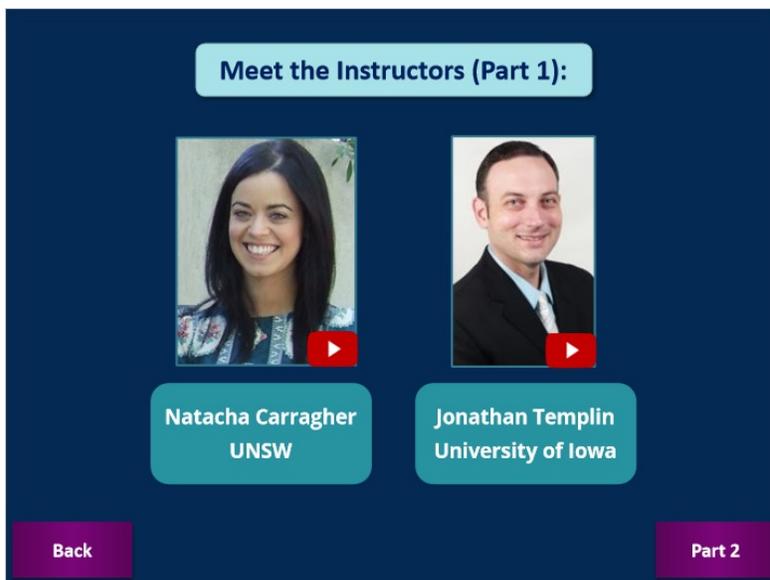
DM04 SLIDES (DCM Checklists, Version 1.4)

1. Module Overview

1.1 Module Cover (START)



1.2 Instructors (I)



1.3 Instructors (II)

Meet the Instructors (Part 2):



Philip Jones
UNSW

Boaz Shulruf
UNSW

Gary Velan
UNSW

Back

Part 1

1.4 Designers

Meet the designers:



André A. Rupp
Mindful Measurement

Xi Lu
Florida State University

Back

1.5 Welcome



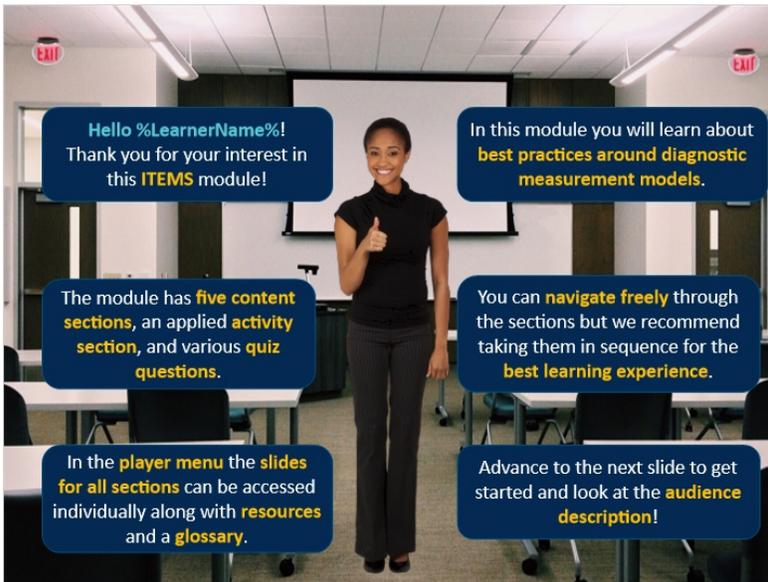
**Welcome to the
ITEMS Module!**

The woman to the left is Laura!

Along with the instructors
she will be guiding you
through the module content.

Please type your name here:

1.6 Overview



Hello %LearnerName%!
Thank you for your interest in
this ITEMS module!

The module has **five content sections**, an applied **activity section**, and various **quiz questions**.

In the **player menu** the **slides for all sections** can be accessed individually along with **resources** and a **glossary**.

In this module you will learn about **best practices around diagnostic measurement models**.

You can **navigate freely** through the sections but we recommend taking them in sequence for the **best learning experience**.

Advance to the next slide to get started and look at the **audience description!**

1.7 Target Audience: General

Target Audience: General

Anyone who would like a gentle statistical introduction to this topic:

- graduate students in Master's , Ph.D. or certificate programs
- psychometricians and other measurement professionals
- data scientists / analysts
- research assistants or research scientists
- technical project directors
- assessment developers

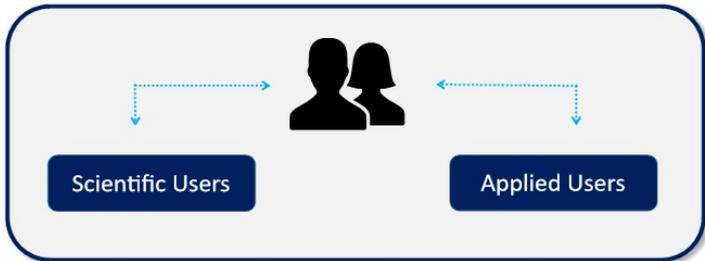


We hope that you find this module useful no matter what your title or role in a professional organization or institution is!

1.8 Target Audience: Specific

Target Audience: Specific

There are two broad groups of users who we think will be interested in and benefit from the checklists presented in this module



Click on the buttons above to learn more or, alternatively, advance to the next slide to continue with the introduction

1.9 Target Audience: Scientific Users

Target Audience: Scientific Users

Estimating DCMs requires the use of **specialized software** (e.g., **Mplus, SAS, R**), **significant time investment**, as well as **psychometric expertise / experience**

Therefore, we envision that the **'Estimation'**, **'Model Fit Evaluation'** and **'Interpretation'** checklists will mostly likely appeal to **scientific users**

These include:

- data analysts / scientists
- research assistants or research scientists
- psychometricians and other measurement professionals
- graduate students in Master's, Ph.D. or certificate programs

1.10 Target Audience: Applied Users

Target Audience: Applied Users

The **'Reporting'** checklist is general and provides a **template for reviewing the quality of the methodology and reporting** of DCM studies

Therefore we envisage that this checklist will likely appeal to **applied users**:

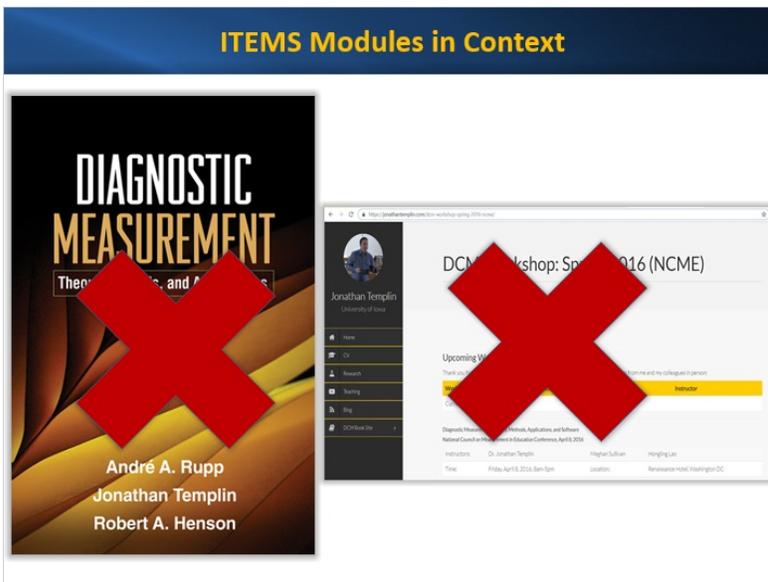
- journal editors
- peer reviewers
- funding agencies
- readers



1.11 Scope (I)



1.12 Scope (II)



1.13 Scope (III)

Scope of Module

This module guides the learner through the **application of Diagnostic Classification Models (DCMs)** to real-world data using **four succinct and easy-to-follow checklists**:

- 1 Estimation Checklist
- 2 Model Fit Evaluation Checklist
- 3 Interpretation Checklist
- 4 Reporting Checklist

1.14 Prerequisites

Prerequisites

In order to **maximize your learning experience** within this module, you should ideally have:

- Basic experience using **SAS**
- Basic experience using **Mplus**
- Basic understanding of the core concepts of **structural equation modeling (SEM)**

However, **no previous knowledge or experience with DCMs is necessary**



[SAS Portal](#) [Mplus Portal](#) [SEM Module](#)

1.15 Objectives (I)

Learning Objectives: Background

Background

- 1 Understand the benefits of DCMs over other classification models
- 2 Understand key statistical properties and foundations of DCMs
- 3 Understand the type of outputs generated by DCMs
- 4 Learn various labels used to describe DCMs
- 5 Understand the basic goals for and key terminology around DCMs

1.16 Objectives (II)

Learning Objectives: Framework & Checklists

LCDM Framework

- 6 Understand the key statistical properties of the LCDM
- 7 Understand differences between the LCDM and ANOVA
- 8 Understand the motivation and purpose of DCM checklists

Checklist Development

- 9 Understand the purpose and benefits of checklists in general
- 10 Learn how checklists are used in applied settings
- 11 Understand the structure of the four DCM Checklists

1.17 Objectives (III)

Learning Objectives: Usage & Activity

Checklist Usage	12	Understand how to use the DCM checklists in order to estimate, evaluate, interpret and report the results of a DCM analysis
	13	Specify and estimate an LCDM and C-RUM using real-life data and Mplus
Activity	14	Perform relative model fit analyses and choose a best-fitting model
	15	Interpret the resulting parameter estimates and perform basic computations
	16	Translate the findings into a brief diagnostic score report

1.18 Resources

Resources

Carragher N., Templin J., Jones P., Shulruf B., & Velan G. M. (2019). Diagnostic measurement: Modeling checklists for practitioners (Digital ITEMS Module 04). *Educational Measurement: Issues and Practice*, 38, 89-90. Available online at <https://ncme.elevate.commpartners.com/>

Module Citation

 <p>Additional References</p>	 <p>Jon Templin's Channel</p>
--	--

References (I) (Slide Layer)

Some Useful Resources (I)

Bradshaw L. P. (2016). Diagnostic classification models. In A. A. Rupp & J. P. Leighton (Eds.), *The Handbook of Cognition and Assessment: Theory, Methods, and Applications* (pp. 297-327). New York, NY: Wiley.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd edition). New York: Guilford Press.

Gawande, A. (2011). *The checklist manifesto: How to get things right*. London: Profile Books.

Haynes, A. B., Weiser, T. G., Berry, W. R., Lipsitz, S. R., Breizat, A. H., Dellinger, E. P., Herbosa, T., Joseph, S., Kibatala, P. L., Lapitan, M. C., Merry, A. F., Moorthy, K., Reznick, R. K., Taylor, B., Gawande, A. A. (2009). A surgical safety checklist to reduce morbidity and mortality in a global population. *The New England Journal of Medicine*, *360*, 491-499.

Merenda, P. F. (1997). A guide to the proper use of factor analysis in the conduct and reporting of research: Pitfalls to avoid. *Measurement and Evaluation in Counseling and Development*, *30*, 156-164.

More
Resources

References (II) (Slide Layer)

Some Useful Resources (II)

Rupp A. A., Templin J., & Henson R. A. (2010). *Diagnostic measurement: Theory, methods and applications*. New York, NY: Guilford Press.

Schumacker, R. E., Lomax, R. G. (1996). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Erlbaum.

Shook, C. L., Ketchen, D. J., Hult, G. T. M., Kacmar, K. M. (2004). An assessment of the use of structural equation modeling in strategic management research. *Strategic Management Journal*, *25*, 397-404.

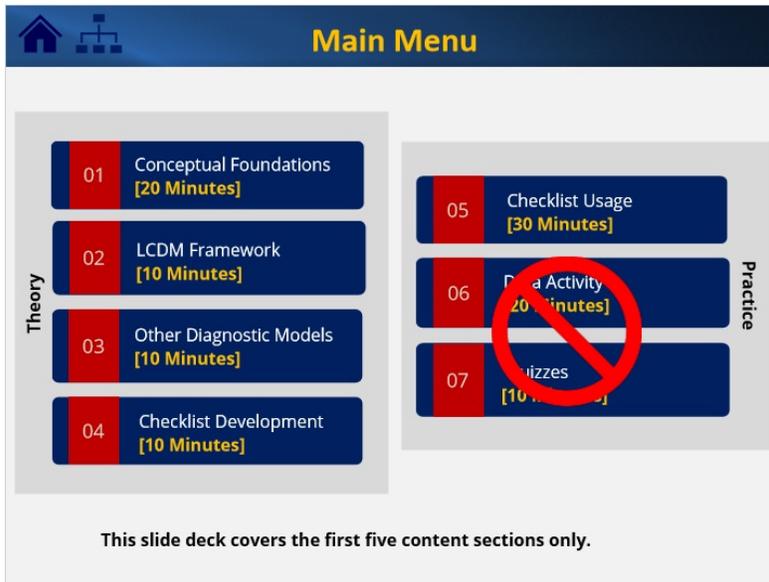
Templin J., & Hoffman L. (2013). Obtaining diagnostic classification model estimates using Mplus. *Educational Measurement: Issues and Practice*, *32*, 37-50.

van de Schoot, R., Lugtig, P., Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, *9*, 37-41.

Online seminars and workshops, exercises on DCMs, and more:
<http://jonathantemplin.com/teaching/academic-courses/dcm/>
<https://jonathantemplin.com/sas-macro-estimation-lcdm-mplus/>
<https://jmu.njvid.net/show.php?pid=njcore:138598>

Back

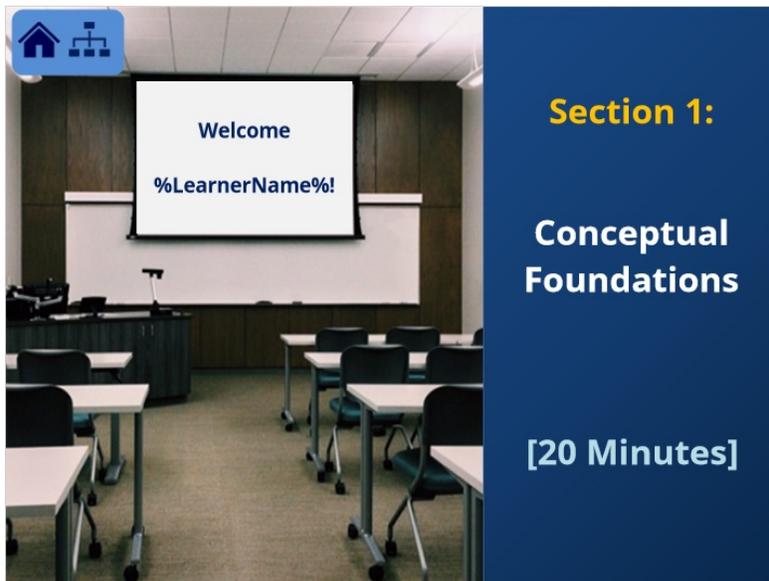
1.19 Main Menu



The Main Menu slide deck interface features a dark blue header with a home icon and a tree diagram icon on the left, and the text "Main Menu" in yellow on the right. Below the header, there are two columns of content. The left column is labeled "Theory" and contains four items: 01 Conceptual Foundations [20 Minutes], 02 LCDM Framework [10 Minutes], 03 Other Diagnostic Models [10 Minutes], and 04 Checklist Development [10 Minutes]. The right column is labeled "Practice" and contains three items: 05 Checklist Usage [30 Minutes], 06 Data Activity [20 Minutes], and 07 Quizzes [10 Minutes]. A red circle with a diagonal slash is drawn over the "06 Data Activity" item. At the bottom of the slide deck, there is a text box that reads "This slide deck covers the first five content sections only."

2. Section 1: Conceptual Foundations

2.1 Cover: Section 1



The slide for Section 1: Conceptual Foundations is split into two parts. On the left, there is a photograph of a classroom with a projector screen displaying "Welcome" and "%LearnerName%!". On the right, there is a dark blue vertical panel with the text "Section 1: Conceptual Foundations" in yellow and white, and "[20 Minutes]" in white at the bottom. A small home icon and tree diagram icon are visible in the top left corner of the slide.

2.2 Learning Objectives: Section 1



Learning Objectives



1. Understand the benefits of DCMs over traditional classification models
2. Understand key statistical properties and basic foundations of DCMs
3. Understand the type of outputs generated by DCMs
4. Learn various labels used to describe DCMs
5. Understand the basic goals for and key terminology around DCMs

2.3 Topic Selection



**General Motivation:
Diagnostic Measurement**

**Modeling Approaches:
IRT vs. DCMs**

Section End

2.4 Bookmark: General Motivation



2.5 General Motivation (I)

General Motivation (I)

Assessments are ubiquitous:

- primary, secondary and tertiary education
- screen or diagnose a medical or psychological condition
- music examinations
- eye-sight tests
- DNA or genetic testing
- IQ test
- personality tests
- driving test . . .

You take an assessment, you get a score, which is used for decision-making (e.g., to provide educational or clinical classifications using a cut-point).

But what if tests gave you more than just a single score...?

2.6 General Motivation (II)



General Motivation (II)

What if the assessment gave you:

- fine-grained information on multiple skills for a content area
- classification of learners into groups with similar skill profiles
- adaptive assessment that is fine-tuned to skill profiles

-> **diagnostic measurement & diagnostic classification models (DCMs)**



2.7 Applications of DCMs



Applications of DCMs

Psychiatry

- DSM symptoms that a person endorses
-> **leads to a differentiated clinical diagnosis / tailored treatment**

Organizational Psychology

- Skills of an applicant that are matched to job requirements
-> **leads to more informed acceptance / rejection decision**
- employee's or firm's performance to diagnose where deficits occur
-> **more tailored professional development**

Education

- Competencies or skills that a person has /has not mastered
-> **leads to instructional interventions / learning support**

2.8 Illustration

  **Illustration**

Using CTT and IRT models (norm-referenced):	Using DCMs (criterion-referenced):
 <ul style="list-style-type: none">▪ Jake obtained a score of 20▪ Jake is at the 60th percentile▪ Jake has an 'intermediate' level of performance▪ Jake passed the exam	 <ul style="list-style-type: none">▪ Jake is <i>proficient</i> in addition▪ Jake is <i>proficient</i> in subtraction▪ Jake <i>should work</i> on multiplication▪ Jake <i>should work</i> on division

Some DCMs also allow for higher-order summaries of performance on higher-order attributes / dimensions

2.9 Terminology: Model Labels

  **Terminology: Model Labels**

▪ Cognitive diagnosis models	
▪ Cognitive psychometric models	
▪ Skills assessment models	
▪ Latent response models	
▪ Restricted (constrained) latent class models	
▪ Multiple classification models	
▪ Structured located latent class models	
▪ Structured item response theory models	

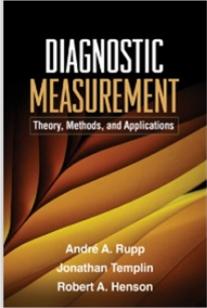
2.10 Terminology (II)

  **Terminology: Resources**

The next slide provides some **basic terminology** for the ITEMS module

Additional terms are explained in the **'Glossary'** section of the player interface for this ITEMS module

An **extensive glossary** is also provided in the book *'Diagnostic Measurement: Theory, Methods, and Applications'*



2.11 Basic Terminology

  **Terminology: Mini Glossary**

Term	Conceptually	Psychometrically
Attribute	Unobservable features	Statistical variable
Diagnosis	Clinical evaluation	Statistical classification
Dimension	Construct aspect	Statistical variable
Item	Stimulus	Statistical variable
Latent variable	Dimension	Statistical variable
Latent class	Group of learners	Set of observations
Q-matrix	Design mapping	Matrix with numbers

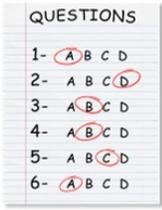
Click on each row to see an additional explanation.  **Part End**

2.12 Attribute



Attributes

- **Conceptually**, the term attributes refers to skills, dispositions, or any other constructs that are related to the behavioral procedures or cognitive processes that a learner must engage in to solve an assessment item
- **Psychometrically**, attributes refer to unobserved (latent) variables in a statistical model (DCM), which are measured through the assessment items and encoded in a Q-matrix [see definition]



QUESTIONS	
1-	A B C D
2-	A B C D
3-	A B C D
4-	A B C D
5-	A B C D
6-	A B C D

Back

2.13 Diagnosis



Diagnosis

- **Conceptually**, a diagnosis can refer to a comprehensive clinical evaluation of a person, sometimes with reference to a clinical manual such as the DSM
- **Psychometrically**, a diagnosis refers to the classification of a learner into one of several unobserved (latent) classes via a statistical model (DCM)



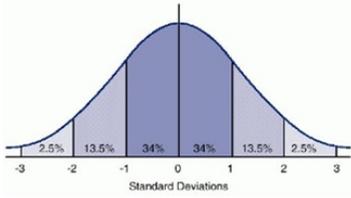
Back

2.14 Dimension



Dimension

- **Conceptually**, a dimension can refer to an aspect or facet of a cognitive response process that a learner engages in while responding to an assessment item
- **Psychometrically**, a dimension is a statistical variable on which learners can be ordered from low to high to facilitate rank ordering of learners (IRT) or classify learners (DCM). Dimensions are also known as attributes [see definition] and encoded in a Q-matrix [see definition]



Standard Deviations

Back

2.15 Item



Item

- **Conceptually**, an item is a physical or digital stimulus presented to a learner to which they respond by selecting a response from various options presented or by providing a constructed response
- **Psychometrically**, an item is related to a statistical variable, which is the score that the learner received for their response to the item (dichotomous or polytomous)



Data File Example

Back

2.16 Data File: Example

Home icon

Data File: Example

Columns = Items

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
Individual 1	1	0	1	0	1	1
Individual 2	1	0	0	0	0	0
Individual 3	0	1	1	1	0	1
Individual 4	1	0	0	1	0	0
Individual 5	1	0	0	0	0	0
Individual 6	0	1	1	1	0	1
Individual 7	1	0	1	0	1	0
Individual 8	1	0	0	1	0	1

Rows = Individuals

Cells = Observed Item Response
(0 = incorrect / not endorsed, 1 = correct / endorsed)

2.17 Latent Variable

Home icon

Latent Variable

- **Conceptually**, a latent variable is a dimension along which we compare learners whose meaning is anchored in our definition of the construct that the assessment is measuring
- **Psychometrically**, a latent variable is an unobserved statistical quantity that we estimate using the observed scores from learners along with the associated item parameters for the assessment items



Back

2.18 Latent Class



Latent Class

- **Conceptually**, a latent class is an unobserved grouping of learners that share similar characteristics – in the case of DCMs, this is the same set of attribute mastery statuses / attribute profiles
- **Psychometrically**, a latent class is an unobserved classification state into which learners are sorted by estimating model parameters. The number of latent classes is pre-determined by the number of attributes and the number of mastery levels. The latent classes are mutually exclusive



[Back](#)

2.19 Q-matrix



Q-matrix

- **Conceptually**, a Q-matrix shows the association between items and dimensions for an assessment or indicates which items provide measurement information about which dimensions
- **Psychometrically**, a Q-matrix is a two-dimensional table (rows represent items, columns represent attributes) with numeric entries (typically 0s and 1s) that is used to compute the response probabilities for individual items and latent classes under a given DCM

	Attribute 1	Attribute 2	Attribute 3
Item 1	0	1	0
Item 2	1	0	0
Item 3	1	1	0
Item 4	0	1	1

[Q-matrix Example](#)

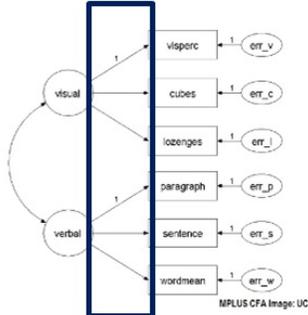
[Back](#)

2.20 Q-matrix: Item-attribute Alignment

🏠 📊 Q-matrix: Item-attribute Alignment

- A Q-matrix is a central feature of DCMs and **specified a priori** to indicate which latent attributes are measured by each item (i.e., **item-attribute alignment**)

- A Q-matrix is identical to the **factor pattern matrix in a confirmatory factor analysis**



MPLUS CFA Image: UCLA

2.21 Q-matrix: Example

🏠 📊 Q-matrix: Example

Item	Attribute 1	Attribute 2	Highest level of interaction
1	1	0	1
2	1	1	2
3	0	1	1
4	1	0	1
5	1	0	1
6	1	0	1
7	1	0	1
8	1	0	1
9	1	1	2
10	1	0	1
11	1	1	2
12	1	1	2
13	1	1	2

Attribute columns: 1 = attribute is measured, 0 = attribute is not measured

Level of interaction: 1 = main effect, 2 = 2-way interaction

2.22 Q-matrix: Specification



Q-matrix: Specification

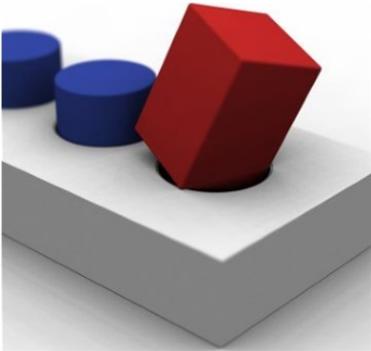
- The specification of a Q-matrix **underlies all inferences** resulting from a DCM analysis. Therefore, accurate specification of the Q-matrix is **critical as it affects model fit and interpretation**
- In education, the initial specification of a **Q-matrix typically requires content-specific learning theory and expert insights** from teachers, researchers and psychometricians
- Verifying the **item-attribute alignment** is a complex process that requires **substantial qualitative research** to identify the specific mental processes that examinees use to respond to items

2.23 Q-matrix: Retrofitting



Q-matrix: Retrofitting

- Since **large-scale applications** of DCMs are still in their infancy, there are **few studies** that have **prospectively created tests** designed to be modeled with a DCM
- Consequently, **most studies** unfortunately use **retrofitted** data from assessments that were **not originally designed** for analyses with DCMs



2.24 Bookmark: Modeling Approaches



2.25 Overview

Overview	
Multidimensional IRT	DCMs
Multiple continuous variables	Multiple discrete variables
Continuous score profiles	Classifications into latent classes
Item difficulty and discrimination	Item guessing and slipping
Parametric estimation	Parametric estimation
Fit assessment possible	Fit assessment possible
Score reliability / precision	Classification reliability
Monotonicity	Partial monotonicity

2.26 IRT Foundations (I)



IRT Foundations (I)

- Performance is based on a **single continuous latent trait θ** (*unidimensional IRT*) or **multiple continuous latent traits $\underline{\theta}$** (trait vector) (*multidimensional IRT*)
- Students with **higher** latent trait values have **higher** probabilities of getting an **item correct** (*dichotomous items*) or obtaining a **higher score** (*polytomous items*)

2.27 IRT Foundations (II)



IRT Foundations (II)

- **Item parameters** can be used to characterize an item's operating characteristics such as difficulty, discrimination, or guessing (*number of parameters depends on model*)
- **Score precision / item information** and **reliability estimates** can be computed for each latent variable (*via information functions and marginal statistics*)

2.28 DCM Foundations (I)



DCM Foundations (I)

- DCMs are a family of **parametric psychometric models** that:
 - ✓ **classify respondents** into mutually exclusive classes
 - ✓ **provide fine-grained feedback** on performance via a multidimensional profile of latent skills based on responses to a set of items
- DCMs are **constrained versions of latent class models** where the number of classes is specified **a priori** (*like confirmatory factor analysis*)

2.29 DCM Foundations (I)



DCM Foundations (II)

- Performance is based on **multiple discrete latent variables**, which are also called attributes $\alpha = (\alpha_1, \dots, \alpha_k)$
(*similar to multidimensional IRT*)
- The discrete variables result in a **multi-faceted classification** of learners organized into latent classes
(*unlike multidimensional IRT*)
- Each item is designed to measure one or more latent attributes with the exact design captured in a **Q-matrix with 0s and 1s**
(*similar to confirmatory factor analysis*)

2.30 DCM Foundations (III)



DCM Foundations (III)

- **Response probabilities are class-specific** and depend on which attributes are mastered in the particular class
(probabilities depend on model)
- **Reliability estimates** for the discrete score variables that yield the classifications can still be computed
(but use a different formula than IRT)

2.31 DCM Foundations (IV)



DCM Foundations (IV)

- **for each learner**, DCMs produce marginal attribute mastery rates (*for each attribute*), probabilities of latent class membership (*for each latent class*), and the most likely latent class to which the examinee belongs (*marginally*)
- **for groups of learners**, DCMs produce the proportion of learners classified into each latent class (*marginally*) and the attribute mastery rates (*marginally for each attribute*)
- **for each item**, DCMs produce item parameter estimates (*number depends on model*) and the estimated response probabilities (*for each latent class*)

Learner Statistics Group Statistics Item Statistics Part End

2.32 Learner Statistics: Probabilities



Learner Statistics: Probabilities

Below is an extract of results from an LCDM based on 13 items, 2 attributes and 4 latent classes. This SAS output displays the:

- ✓ **posterior probabilities** of latent class membership for each class;
- ✓ **latent class to which each respondent has been assigned** based on their highest posterior probability;
- ✓ **marginal attribute probability** for each attribute

	ID	cprob1	cprob2	cprob3	cprob4	class	prob_attribute1	prob_attribute2
1	1	0.02709	0.03823	0.31229	0.6224	4	0.93469	0.66063
2	2	0.0622	0.10936	0.24326	0.58519	4	0.82845	0.69455
3	3	0.32058	0.4524	0.07585	0.15117	2	0.22702	0.60357
4	4	0.97556	0.0216	0.0027	0.00015	1	0.00285	0.02175
5	5	0.00534	0.0144	0.31674	0.66352	4	0.98026	0.67792
6	6	0.58578	0.40042	0.00762	0.00618	1	0.0138	0.4066
7	7	0.0646	0.27956	0.15639	0.49945	4	0.65584	0.77901
8	8	0.5183	0.44144	0.02035	0.01991	1	0.04026	0.46135
9	9	0.96872	0.02764	0.0034	0.00023	1	0.00363	0.02787
10	10	0.90666	0.06655	0.02226	0.00452	1	0.02678	0.07107
11	11	0.29626	0.32435	0.14716	0.23222	2	0.37938	0.55657
12	12	0.30561	0.33459	0.13956	0.22023	2	0.35979	0.55482

2.33 Learner Statistics: Feedback

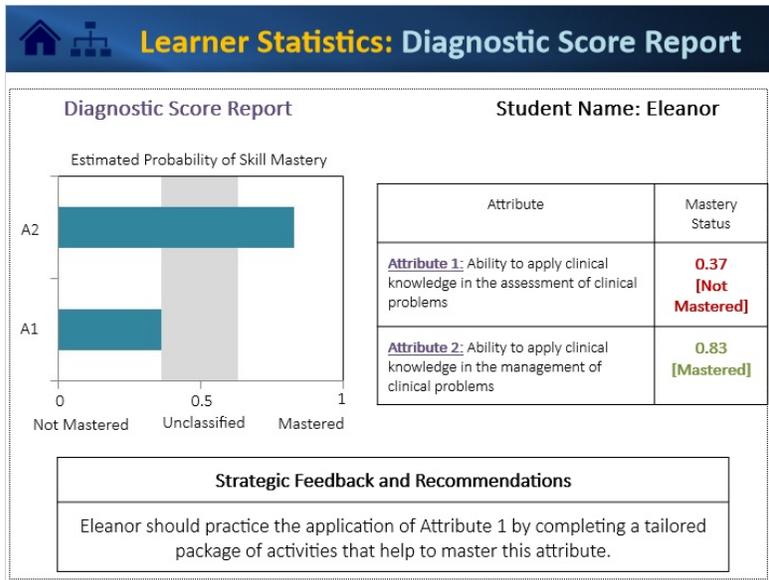


Learner Statistics: Feedback

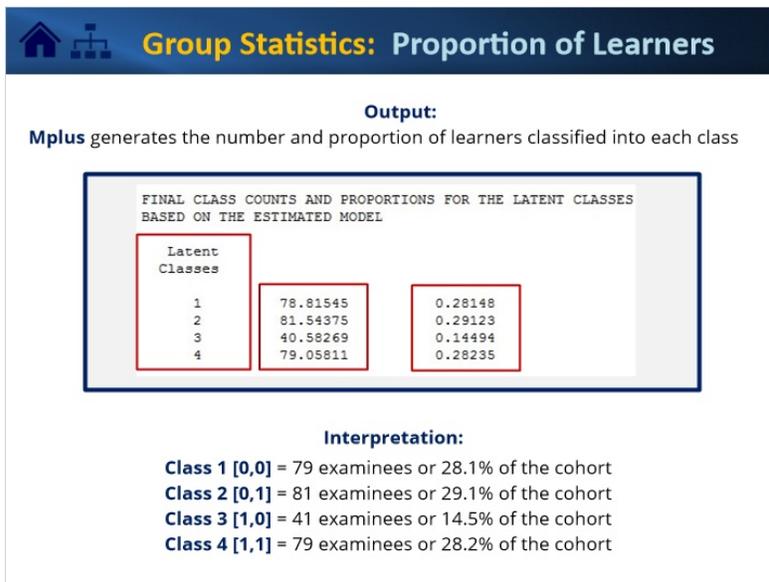
The marginal attribute probabilities can be used to:

- ✓ **develop detailed diagnostic feedback for each respondent** to facilitate identification of remedial instruction and increase their proficiency
- ✓ **summarize the cohort's overall performance** in order to help teachers identify which components of a course they should focus on to increase overall proficiency

2.34 Learner Statistics: Report



2.35 Group Statistics



2.36 Item Statistics

  **Item Statistics: Parameter Estimates**

		New/Additional Parameters	
	G_0		-0.003
	G_11		-0.664
	G_12		0.034
	G_212		0.633
Item 1: Intercept	L1_0		1.191
Item 1: Main effect attribute 1	L1_11		0.501
Item 2: Intercept	L2_0		0.632
Item 2: Main effect attribute 1	L2_11		2.611
Item 2: Main effect attribute 2	L2_12		0.725
Item 2: interaction attributes 1+2	L2_212		0.175
Item 3: Intercept	L3_0		2.099
Item 3: Main effect attribute 2	L3_12		0.733
Item 4: Intercept	L4_0		-0.716
Item 4: Main effect attribute 1	L4_11		0.870
Item 5: Intercept	L5_0		0.189
Item 5: Main effect attribute 1	L5_11		0.200

2.37 Bookend: Section 1



If you are interested in taking a **quiz** on this section click here: [Quiz](#)

If you are interested in **analyzing some sample data** using an applied exercise click here: [Applied Exercise](#)

If you want to return to the **main menu**, click here: [Main Menu](#)

3. Section 2: LCDM Framework

3.1 Cover: Section 2



3.2 Learning Objectives: Section 2

Learning Objectives

1. Understand the key statistical properties of the LCDM
2. Understand similarities and differences between the LCDM and Analysis of Variance (ANOVA)

3.3 Topic Selection



3.4 Bookmark: LCDM Foundations



3.5 LCDM Background: Models

  **LCDM Background: Models**

Numerous DCMs have been developed of varying parametric complexity; some historically popular models include:

- Deterministic Inputs Noisy 'And' gate (**DINA**) model
- Deterministic Inputs Noisy 'Or' gate (**DINO**) model
- Noisy Inputs Deterministic 'And' gate (**NIDA**) model
- Noisy Inputs Deterministic 'Or' gate (**NIDO**) model
- Compensatory Reparameterized Unified Model (**CRUM**)
- Reduced Reparameterized Unified Model (**RRUM**)

3.6 LCDM Background: Attribute Interactions

  **LCDM Background: Attribute Interactions**

Each of these DCMs makes assumptions about how mastered attributes combine/interact to produce an item response:

- **compensatory/ disjunctive/ additive**
(e.g., for an item measuring two attributes, a respondent may not be a master of one attribute but still has a high chance of mastering the other attribute)

OR

- **non-compensatory/ conjunctive/ non-additive**
(e.g., for an item measuring two attributes, a respondent must have mastered both attributes in order to give a correct response to the item)

3.7 LCDM Background: Unified Frameworks

  **LCDM Background: Unified Frameworks**

Unified modeling frameworks have been developed which subsume these DCMs:

- Log-linear Cognitive Diagnosis Model (LCDM) 
- Generalized DINA Model (G-DINA)
- General Diagnostic Model (GDM)

Focus of this ITEMS Module

Despite some **statistical differences** in specification and estimation they all **share very similar analytic goals**

They can be estimated using specialized code for **commercial packages** like **Mplus** and **SAS** or **freeware packages** like 'CDM' or 'GDINA' in R

3.8 LCDM Background: Benefits

  **LCDM Background: Benefits**

The LCDM has a number of benefits:

- ✓ **it is flexible** as it allows for both additive and non-additive relationships between attributes / items
- ✓ **it overcomes confusion** for the analyst about choosing which model is most suitable for their needs
- ✓ **it aligns with other statistical and psychometric modeling frameworks** such as ANOVA and CFAs, thereby enabling greater understanding of the modeling process

3.9 LCDM Background (V)



The LCDM Framework: Features

- The LCDM models the **probability of a correct response** to an item as a function of the **latent attributes** of respondents in a given latent class
- The latent attributes are **categorical**, meaning there are a finite number of possible **statuses / attribute profiles**
- Each **status / attribute profile** in each latent class corresponds to a **predicted probability** of a correct response for each item

3.10 LCDM vs ANOVA (I)



LCDM and ANOVA: Similarities

The LCDM and ANOVA frameworks are structurally similar in some key ways:

- both predict a response using **dummy coded variables** (in the LCDM dummy coded variables [0,1] represent latent attributes)
- both frameworks can be used to test for the presence of **main effects and interactions**
- both frameworks offer a means of **reducing model complexity** (i.e., by removing non-significant interactions and/or main effects)

3.11 LCDM vs ANOVA (II)



LCDM and ANOVA: Differences

But there are some important differences between the two frameworks:

- the ANOVA framework models a **continuous outcome** whereas the LCDM models a function of the **probability of a correct response**
- the ANOVA framework uses **observed 'factors'** as predictors (no measurement error) whereas the LCDM uses **discrete latent variables** (the attributes being measured)
- the ANOVA framework uses a more simple **sums-of-squares-based estimation approach** whereas the LCDM requires **maximum likelihood-based or Bayesian estimation approaches**

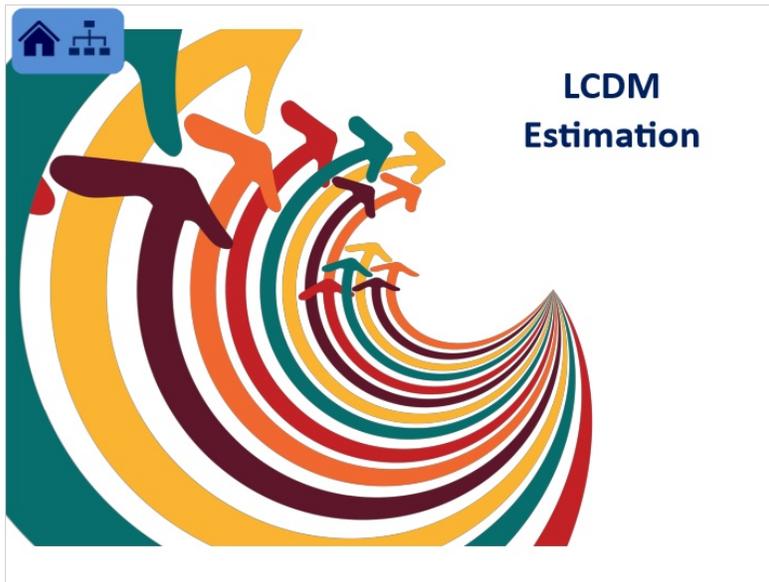
3.12 Bookend: LCDM Foundations



This is the end of this part.

[Topic Selection](#)

3.13 Bookmark: LCDM Estimation



3.14 LCDM Notation: Introduction

 **LCDM Notation: Introduction**

To understand the LCDM, let's consider a basic mathematics test with:

- **four attributes** (addition, subtraction, multiplication and division)
- **three items**

	Add	Sub	Mult	Div
2+3-1	1	1	0	0
4/2	0	0	0	1
(4 x 2)+3	1	0	1	0

1st item assesses two attributes:
- addition (attribute 1: α_{r1})
- subtraction (attribute 2: α_{r2})

1 = attribute it assessed by an item
0 = item is not assessed by an item

3.15 LCDM Notation: Structure

🏠
LCDM Notation: General Structure

The LCDM provides the **logit of a correct response** as a function of the **latent attributes** that are mastered by a respondent and **effect parameters**:

$$\text{Logit}(Y_{ri} = 1 | \alpha_r) = \lambda_{i,0} + \lambda_{i,1,(1)} \alpha_{r1} + \lambda_{i,1,(2)} \alpha_{r2} + \lambda_{i,2,(1,2)} \alpha_{r1} \alpha_{r2}$$

Logit = log-odds of the probability of a correct response

The logit is used because the items are either answered correctly (1) or incorrectly (0)

Latent attributes [mastered (1) or not mastered (0)]

Effect parameters (intercept, main effects, interaction effects)

3.16 LCDM Notation: Parameters

🏠
LCDM Notation: Parameters

$$\text{Logit}(Y_{ri} = 1 | \alpha_r) = \lambda_{i,0} + \lambda_{i,1,(1)} \alpha_{r1} + \lambda_{i,1,(2)} \alpha_{r2} + \lambda_{i,2,(1,2)} \alpha_{r1} \alpha_{r2}$$

$\text{Logit}(Y_{ri} = 1 | \alpha_r)$ logit of a correct response to item i by respondent r

$\lambda_{i,0}$	<p style="font-size: 12px; margin: 0;">intercept:</p> <ul style="list-style-type: none"> ▪ reference group is learners who have mastered neither attribute ($\alpha_{r1} = 0$ and $\alpha_{r2} = 0$)
$\lambda_{i,1,(1)}$	<p style="font-size: 12px; margin: 0;">conditional main effect for addition (attribute 1):</p> <ul style="list-style-type: none"> ▪ the increase in the logit for mastering addition (for someone who has not mastered subtraction)
$\lambda_{i,1,(2)}$	<p style="font-size: 12px; margin: 0;">conditional main effect for subtraction (attribute 2):</p> <ul style="list-style-type: none"> ▪ the increase in the logit for mastering subtraction (for someone who has not mastered addition)
$\lambda_{i,2,(1,2)}$	<p style="font-size: 12px; margin: 0;">2-way interaction between addition and subtraction (attributes 1 and 2):</p> <ul style="list-style-type: none"> ▪ change in the logit for mastering both addition & subtraction

3.17 LCDM Notation: Subscripts



LCDM Notation: Subscripts

$$\text{Logit}(Y_{ri} = 1 | \alpha_r) = \lambda_{i,0} + \lambda_{i,1,(1)}\alpha_{r1} + \lambda_{i,1,(2)}\alpha_{r2} + \lambda_{i,2,(1,2)}\alpha_{r1}\alpha_{r2}$$

(1) Subscript i relates to the item to which parameters belong

(2) Subscript following i represents the level of the effect:

- 0 = intercept
- 1 = main effect
- 2 = two-way interaction

(3) Subscript in parentheses and α_j relate to the attributes to which the effect applies:

- 1 = attribute 1
- 2 = attribute 2

3.18 LCDM Estimation: SAS Macro



LCDM Estimation: SAS Macro

The LCDM can be estimated in Mplus, however the syntax is quite lengthy



To make this process easier, Templin and Hoffman developed a **SAS macro** to facilitate model specification and output processing:

<https://jonathantemplin.com/sas-macro-estimation-lcdm-mplus/>

The user simply alters specific lines in the syntax to tailor the model to their data structure

3.19 LCDM Estimation: Video



“Using the SAS Macro to
Generate LCDM Mplus Code”

Video presentation by Daniel Jurich
[17 Minutes]

<https://jmu.njvid.net/show.php?pid=njcore:138598>



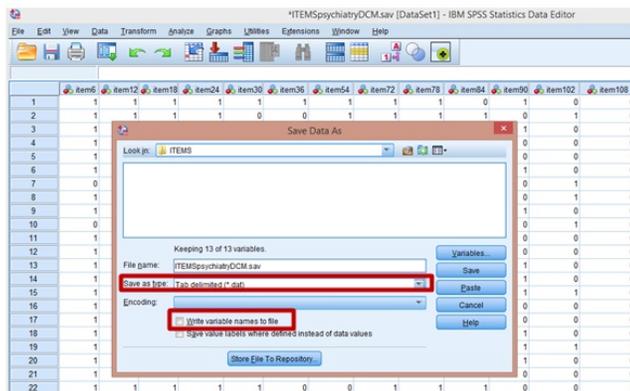
3.20 LCDM Estimation: Saving Data



Comment: Data File Structure

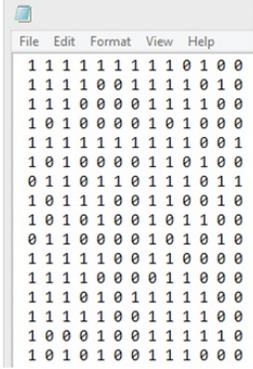
Save data in a Tab delimited data file in your data management program:

File – Save Data As . . .



3.21 LCDM Estimation: Data File

  **Comment: Data File Structure**



```
File Edit Format View Help
1 1 1 1 1 1 1 1 1 0 1 0 0
1 1 1 1 0 0 1 1 1 1 0 1 0
1 1 1 0 0 0 0 1 1 1 1 0 0
1 0 1 0 0 0 0 1 0 1 0 0 0
1 1 1 1 1 1 1 1 1 1 0 0 1
1 0 1 0 0 0 0 1 1 0 1 0 0
0 1 1 0 1 1 0 1 1 1 0 1 1
1 0 1 1 1 0 0 1 1 0 0 1 0
1 0 1 0 1 0 0 1 0 1 1 0 0
0 1 1 0 0 0 0 1 0 1 0 1 0
1 1 1 1 1 0 0 1 1 0 0 0 0
1 1 1 1 0 0 0 0 1 1 0 0 0
1 1 1 0 1 0 1 1 1 1 1 0 0
1 1 1 1 1 0 0 1 1 1 1 0 0
1 0 0 0 1 0 0 1 1 1 1 1 0
1 0 1 0 1 0 0 1 1 1 0 0 0
```

Here is an extract of what the data will look like in Notepad

3.22 LCDM Estimation: Syntax Adaptation

  **SAS Example File: Syntax Adaptation**

- Access and download the **Macro File.sas** and the **Example File.sas**
- In the **Example File.sas** change the following syntax lines to suit your data:

<pre>LIBNAME folder %LET macroloc= %LET filesave= %LET filename = %LET Qname= %LET dataname=</pre>	<pre>%LET itemlist= %LET numitem= %LET numatt= %LET numclass= INFILE "&filesave.\</pre>
--	---
- Do **NOT** make any changes to the **Macro File.sas**
- Import the **Q-matrix** from Excel below `DATALINES;`
- **Tip:** Save all files on the **Desktop** and work from the **Desktop**

3.23 Bookend: LCDM Estimation



This is the end of this part.

Topic Selection

3.24 Bookend: Section 2



If you are interested in taking a **quiz** on this section click here:

Quiz

If you are interested in **analyzing some sample data** using an applied exercise click here:

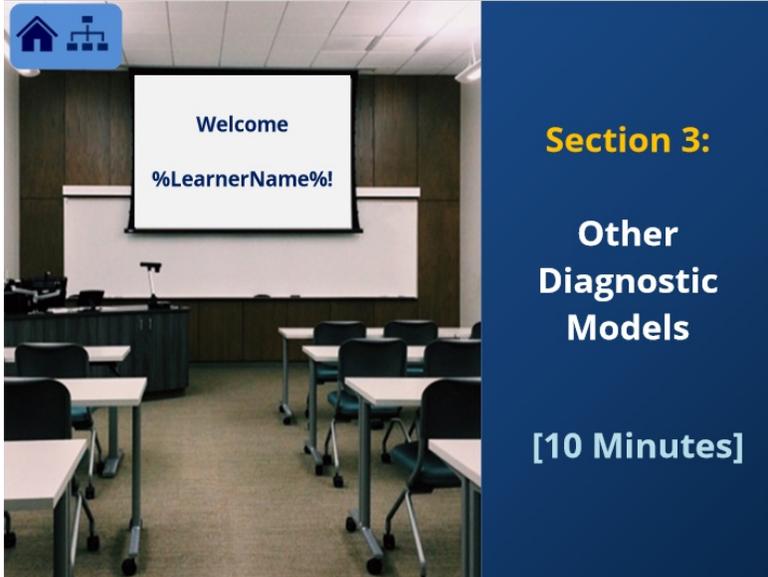
Applied Exercise

If you want to return to the **main menu**, click here:

Main Menu

4. Section 3: Other Diagnostic Models

4.1 Cover: Section 3



4.2 Objectives: Section 3

Learning Objectives

1. Understand the statistical differences between the LCDM, C-RUM, NIDO, DINO and DINA
2. Learn about the specific Mplus model estimation code of the LCDM, C-RUM, NIDO, DINO and DINA

4.3 Model Selection

Click on the four lower-level buttons above to learn more about the specific model estimation code in Mplus

Section End

4.4 C-RUM Specification (Mplus)

C-RUM Specification

- The main difference between the LCDM and the C-RUM is that the C-RUM **does NOT have any interactions**
- To specify a C-RUM, simply **remove the interaction terms** from the 'MODEL CONSTRAINT' command

LCDM	C-RUM
<pre>MODEL CONSTRAINT: ! Item 9: Define LCDM parameters present for item 9 NEW(L9_0 L9_11 L9_12 L9_212); T9_1=(L9_0); T9_2=(L9_0+L9_12); T9_3=(L9_0+L9_11); T9_4=(L9_0+L9_11+L9_12 L9_212); ! Main effect order constraints L9_11>0; L9_12>0; ! Two-way interaction order constraints L9_212>L9_11; L9_212>L9_12;</pre>	<pre>MODEL CONSTRAINT: ! Item 9: Define C-RUM parameters present for item 9 NEW(L9_0 L9_11 L9_12); T9_1=(L9_0); T9_2=(L9_0+L9_12); T9_3=(L9_0+L9_11); T9_4=(L9_0+L9_11+L9_12); ! Main effect order constraints L9_11>0; L9_12>0;</pre>
<pre>! Item 10: Define LCDM parameters present for item 10 NEW(L10_0 L10_11); T10_1=(L10_0); T10_2=(L10_0+L10_11); ! Main effect order constraints L10_11>0;</pre>	<pre>! Item 10: Define C-RUM parameters present for item 10 NEW(L10_0 L10_11); T10_1=(L10_0); T10_2=(L10_0+L10_11); ! Main effect order constraints L10_11>0;</pre>

4.5 NIDO Specification (Mplus)

  NIDO Specification

- The NIDO model is a **constrained version** of the C-RUM
- The NIDO has a **unique intercept parameter** for each item, but **constrains main effect parameters** to be equal across items

C-RUM	NIDO
<pre>MODEL CONSTRAINT: ! Item 9: Define C-RUM parameters present for item 9 NEW(L9_0 L9_11 L9_12); T9_1=-(L9_0); T9_2=-(L9_0+L9_11); T9_3=-(L9_0+L9_11); T9_4=-(L9_0+L9_11+L9_12); ! Main effect order constraints L9_11>0; L9_12>0; ! Item 10: Define C-RUM parameters present for item 10 NEW(L10_0 L10_11); T10_1=-(L10_0); T10_2=-(L10_0+L10_11); ! Main effect order constraints L10_11>0;</pre>	<pre>MODEL CONSTRAINT: ! Item 9: Define NIDO parameters present for item 9 NEW(L9_0); T9_1=-(L9_0); T9_2=-(L9_0+L_2); T9_3=-(L9_0+L_1); T9_4=-(L9_0+L_1+L_2); ! Item 10: Define NIDO parameters present for item 10 NEW(L10_0); T10_1=-(L10_0); T10_2=-(L10_0+L_1);</pre>

4.6 DINO Specification I (Mplus)

  DINO Specification (I)

The DINO model requires only **two probabilities** of a correct response per item for:

- respondents who have **not mastered any** of the required attributes
-> **guessing parameter (g)**
- respondents who have **mastered one or more** of the required attributes
-> **slipping parameter (1-s)**

Requires the specification of a **'NEW'** statement within the **'MODEL CONSTRAINT'** section to define:

- an **intercept parameter** (associated with "guessing")
- a **single main effect parameter** (associated with "slipping")

4.7 DINO Specification II (Mplus)

  DINO Specification (II)

```
! Item 9: Define DINO parameters present for item 9
NEW(L9_0 L9_e);
T9_1=- (L9_0);
T9_2=- (L9_0+L9_e);
T9_3=- (L9_0+L9_e);
T9_4=- (L9_0+L9_e);
L9_e>0;

! Item 10: Define DINO parameters present for item 10
NEW(L10_0 L10_e);
T10_1=- (L10_0);
T10_2=- (L10_0+L10_e);
L10_e>0;
```

4.8 DINA Specification I (Mplus)

  DINA Specification (I)

The DINA model requires **two probabilities** of a correct response per item i

- respondents who are **have not mastered one or more** of the attribut
-> **guessing parameter (g)**
- respondents who have **mastered all** of the attributes
-> **slipping parameter (1-s)**

Requires the specification of a **'NEW'** statement within the **'MODEL CONSTRAINT'** section to define for each item:

- **an intercept parameter** (associated with "guessing")
- **a single main effect parameter** (associated with "slipping")

4.9 DINA Specification II (Mplus)

 **DINA Specification (II)**

DINO	DINA
<pre>! Item 9: Define DINO parameters for item 9 NEW(L9_0 L9_e); T9_1=-(L9_0); T9_2=-(L9_0+L9_e); T9_3=-(L9_0+L9_e); T9_4=-(L9_0+L9_e); L9_e>0; ! Item 10: Define DINO parameters for item 10 NEW(L10_0 L10_e); T10_1=-(L10_0); T10_2=-(L10_0+L10_e); L10_e>0;</pre>	<pre>! Item 9: Define DINA parameters for item 9 NEW(L9_0 L9_e); T9_1=-(L9_0); T9_2=-(L9_0); T9_3=-(L9_0); T9_4=-(L9_0+L9_e); L9_e>0; ! Item 10: Define DINA parameters for item 10 NEW(L10_0 L10_e); T10_1=-(L10_0); T10_2=-(L10_0+L10_e); L10_e>0;</pre>

Thresholds / response probabilities for all (three) latent classes in which at least one attribute is mastered are **defined identically** (i.e., have an increase in the logit)

Thresholds / response probabilities for all (three) latent classes in which at least one attribute is lacking are **defined identically** (i.e., only have the base logit)

4.10 Bookend: Section 3





If you are interested in taking a **quiz** on this section click here: [Quiz](#)

If you are interested in **analyzing sample data** using an applied exercise click here: [Applied Exercise](#)

If you want to return to the **main menu** click here: [Main Menu](#)

5. Section 4: Checklist Development

5.1 Cover: Section 4



5.2 Learning Objectives: Section 4

Learning Objectives

1. Understand the motivation and purpose of checklists for DCMs
2. Understand the purpose and benefits of checklists in general
3. Understand how checklists are used in applied settings
4. Understand the structure of the four DCM checklists

5.3 Topic Selection



5.4 Bookmark: Checklist Motivations



5.5 Utility of Checklists



Utility of Checklists

Human error is inevitable, particularly when:

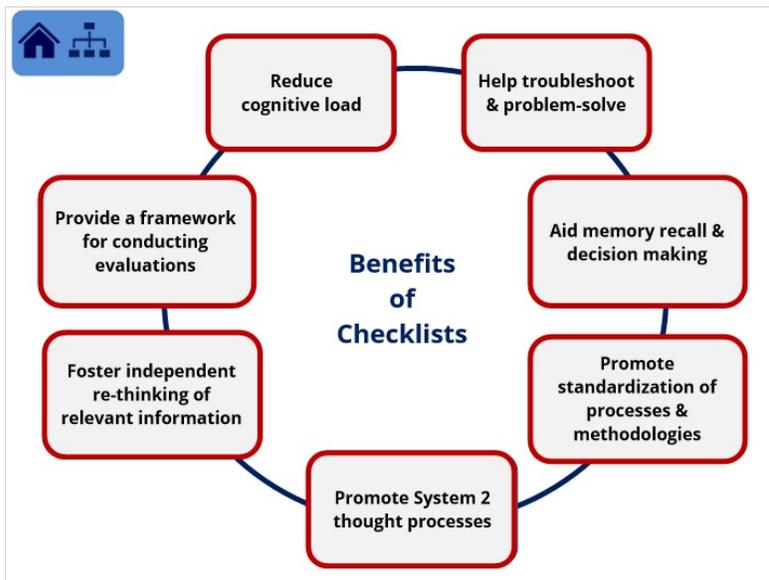
- performing under situations of duress, fatigue, task saturation
- performing complex tasks
- trying to master new skills or acquire new knowledge

Errors can have adverse consequences, compromising judgement, reducing compliance with standards and protocols, and decreasing proficiency



Simple checklists enable corrective steps to be taken before errors cause harm (known as “error trapping”)

5.6 Benefits of Checklists



5.7 Checklists in Daily Life



5.8 Gap in the literature

Gap in the Literature

- DCMs are a **relatively new psychometric framework** that holds considerable promise
- However, **their operational application to real-life data has been limited to date** due, in part, to a lack of **clear accessible guidelines** about how to apply these models in practice
- In the **structural equation modeling (SEM) field**, checklists have been developed for measurement invariance and factor analysis as well as for research publications (e.g., *Brown, 2015; Hancock et al., 2018; Merenda, 1997; Schumacker & Lomax, 2004; Shook, Ketchen, Hult, & Kacmar, 2004; van de Schoot, Lugtig, & Hox, 2012*)

However, no checklists or similar aids have been developed for DCMs

5.9 Checklists for DCMs

🏠
Checklists for DCMs

... an A-Z Guide to DCMs for Beginners

The purpose of the DCM checklists is to **increase the accessibility of DCMs**. They represent an effort to **translate the personal and published experience** of the authors, their colleagues, and students conducting psychometric analyses with DCMs

5.10 Development

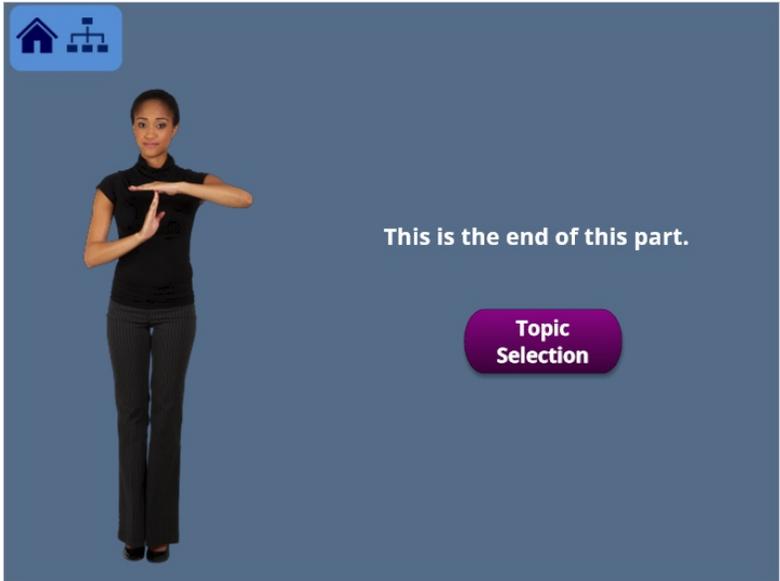
🏠
Development and Testing

Surgical Safety Checklist

 World Health Organization | Patient Safety
A World Alliance for Safer Health Care

Before induction of anaesthesia
(with at least nurse and anaesthetist)
→
Before skin incision
(with nurse, anaesthetist and surgeon)
→
Before patient leaves operating room
(with nurse, anaesthetist and surgeon)

5.11 Bookend: Checklist Motivations



5.12 Bookmark: Checklist Framework



5.13 Checklists Foreword (I)

**Checklists Foreword: Software**

- The DCM Checklists use **SAS, Mplus, R/RStudio** and **Microsoft Excel**
- DCMs are specified and estimated using **Mplus** and a freely available, non-commercial **SAS macro**:
 - ✓ **'SAS Macro for Estimation of the LCDM in Mplus'**
 - ✓ developed by Templin and Hoffman (2013) to facilitate automation of the **Mplus syntax** generation process

<http://jonathantemplin.com/?s=macro&submit=Go>
- Attribute reliability is estimated using an **R/RStudio** template: **'Estimation of Attribute Reliability'**
- Basic computations are performed in **Microsoft Excel**

5.14 Checklists Foreword (II)

**Checklists Foreword: DCMs**

The checklists were developed for use with the following five DCMs:

- Log-linear Cognitive Diagnosis Model (**LCDM**)
- Compensatory Reparameterized Unified Model (**C-RUM**)
- Noisy Inputs Deterministic 'Or' gate model (**NIDO**)
- Deterministic Inputs Noisy 'Or' gate model (**DINO**)
- Deterministic Inputs Noisy 'And' gate model (**DINA**)

The DCM Checklists can be applied to any one DCM or a combination of these models as part of a test of relative model-data fit

5.15 Checklists Foreword (III)

  **Checklists Foreword: Cautionary Notes**

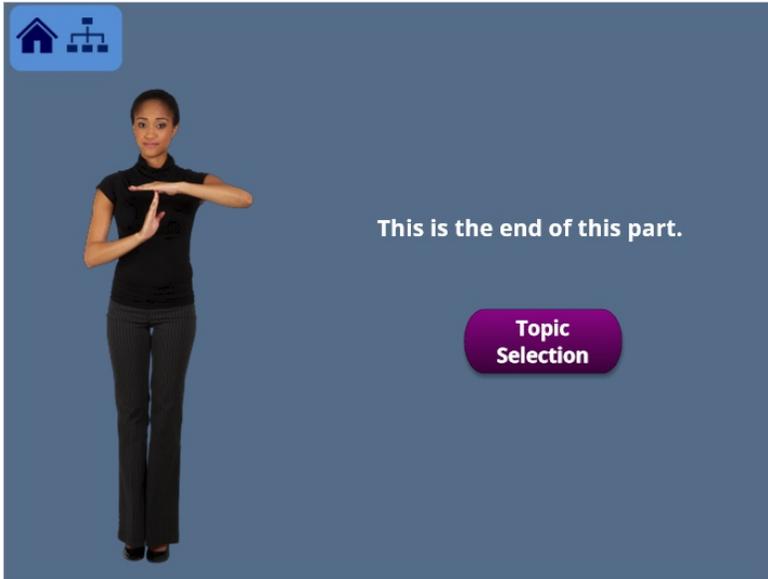
- DCMs are **complex** and analyses can be **idiosyncratic**; therefore, it is **NOT possible** to anticipate the nature of all **possible convergence problems**. The ‘**Estimation**’ checklist is therefore **NOT exhaustive**
- Given the **relative infancy of the field**, areas such as **model-data fit** are undergoing development
- Users are therefore encouraged to **modify** the ‘**Model Fit Evaluation**’, ‘**Interpretation**’, and ‘**Reporting**’ checklists in line with **advances in the literature and knowledge over time**

5.16 Checklists Foreword (IV)

  **Checklists Foreword: Terminology**

- The DCM Checklists relate to models for **binary attribute states** (i.e., present vs. absent, mastered vs. non-mastered) and are well suited for **application in education, psychiatry, and organizational psychology**
- The DCM Checklists use a **READ-DO approach** which involves reading a task, doing it, and checking off the task to confirm that it is completed before moving on to the next task (**similar to following a recipe while baking or cooking**)
- The DCM Checklists may be **used in sequence or separately** but we recommend using them **in sequence**

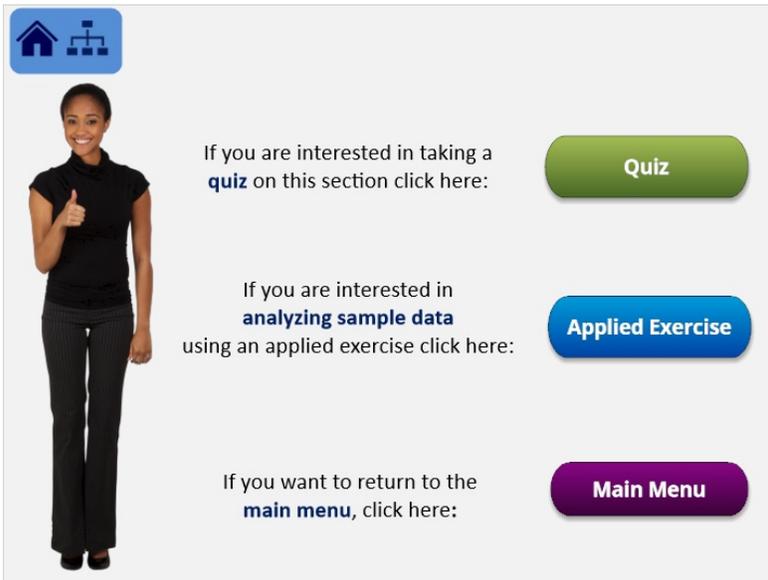
5.17 Bookend: Checklist Framework



This is the end of this part.

[Topic Selection](#)

5.18 Bookend: Section 4



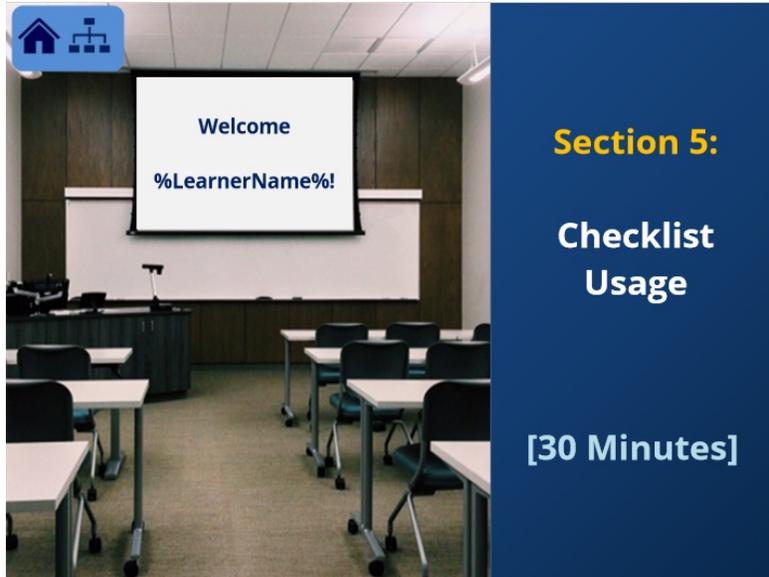
If you are interested in taking a **quiz** on this section click here: [Quiz](#)

If you are interested in **analyzing sample data** using an applied exercise click here: [Applied Exercise](#)

If you want to return to the **main menu**, click here: [Main Menu](#)

6. Section 5: Checklist Usage

6.1 Cover: Section 5



6.2 Learning Objectives: Section 5

Learning Objectives



Understand how to use the four DCM checklist:

1. Estimation
2. Model Fit
3. Interpretation
4. Reporting

6.3 Data Set



Data Set

- Walk through the '**Model Fit Evaluation**' and '**Interpretation**' checklists to **demonstrate concepts using real-life data** to aid understanding
- Data are derived from a **higher education medical assessment** with **13 psychiatry-related items** assessing **2 broadly defined attributes**:
 - Attribute 1 (A1)** = Ability to apply clinical knowledge in the assessment of clinical problems
 - Attribute 2 (A2)** = Ability to apply clinical knowledge in the management of clinical problems

6.4 Companion Files



Companion Files

- All **companion files** related to the data are provided in the '**Resources**' tab of this module and on Jon Templin's website (**click on URL for direct access**):



<https://jonathantemplin.com/ncme-items-module-2019-supplementary-material/>
- Lists of all companion files** can be accessed via the buttons below.

[Input / Output](#)

[Model Estimation](#)

[Model Evaluation](#)

Input / Output (Slide Layer)

Companion Files: Input / Output	
Raw data file: Psychiatry	ITEMSpsychiatry.dat
Q-matrix: Psychiatry	Q matrix ITEMS psychiatry.xlsx
LCDM SAS input file: Psychiatry	ITEMSpsychiatry.sas
LCDM SAS output file {1}: Psychiatry LCDM SAS output file {2}: Psychiatry LCDM SAS output file {3}: Psychiatry LCDM SAS output file {4}: Psychiatry LCDM SAS output file {5}: Psychiatry LCDM SAS output file {6}: Psychiatry LCDM SAS output file {7}: Psychiatry LCDM SAS output file {8}: Psychiatry LCDM SAS output file {9}: Psychiatry	psychiatry_attprobrel.sas7bdat psychiatry_classcounts.sas7bdat psychiatry_classmeans.sas7bdat psychiatry_itemparms.sas7bdat psychiatry_qmatrix.sas7bdat psychiatry_readmplus.sas7bdat psychiatry_respondents.sas7bdat psychiatry_structparms.sas7bdat psychiatry_thresholds.sas7bdat
LCDM Mplus respondents file: Psychiatry LCDM Mplus input file: Psychiatry LCDM Mplus output file: Psychiatry	respondents_ITEMSPsychiatry.dat ITEMSpsychiatry.inp itemspychiatry.out

Back

Model Estimation (Slide Layer)

Companion Files: Model Estimation	
Mplus C-RUM file {1}: Psychiatry Mplus C-RUM file {2}: Psychiatry Mplus C-RUM file {3}: Psychiatry Mplus C-RUM file {4}: Psychiatry	C-RUM, ITEMSpsychiatry.inp c-rum, itemspychiatry.out respondents_ITEMSPsychiatry_CRUM.dat ITEMSpsychiatry.dat
Mplus DINA file {1}: Psychiatry Mplus DINA file {2}: Psychiatry Mplus DINA file {3}: Psychiatry Mplus DINA file {4}: Psychiatry	DINA, ITEMSpsychiatry.inp dina, itemspychiatry.out respondents_ITEMSPsychiatry_DINA.dat ITEMSpsychiatry.dat
Mplus DINO file {1}: Psychiatry Mplus DINO file {2}: Psychiatry Mplus DINO file {3}: Psychiatry Mplus DINO file {4}: Psychiatry	DINO, ITEMSpsychiatry.inp dino, itemspychiatry.out respondents_ITEMSPsychiatry_DINO.dat ITEMSpsychiatry.dat
Mplus NIDO file {1}: Psychiatry Mplus NIDO file {2}: Psychiatry Mplus NIDO file {3}: Psychiatry Mplus NIDO file {4}: Psychiatry	NIDO, ITEMSpsychiatry.inp nido, itemspychiatry.out respondents_ITEMSPsychiatry_NIDO.dat ITEMSpsychiatry.dat

Back

Model Evaluation (Slide Layer)

Companion Files: Model Evaluation	
Attribute reliability file (1): Psychiatry Attribute reliability file (2): Psychiatry Attribute reliability file (3): Psychiatry	Attribute Reliability_Psychiatry.R MplusDCM_functions.R PsychiatryReliability.csv
Loglikelihood ratio calculation: Psychiatry	Loglikelihood ratio calculation psychiatry.xlsx
Attribute mastery profiles: Psychiatry	Attribute mastery profiles psychiatry.xlsx
Diagnostic Score Report: Psychiatry	Diagnostic Score Report psychiatry.pdf
Diagnostic quality of items: Psychiatry	Diagnostic quality of items psychiatry.xlsx
Method and Results write-up: Psychiatry	Method and Results write up psychiatry.pdf
DCM Checklists	DCM Checklists.pdf

Back

6.5 Pedagogical Note

  **Pedagogical Note**

- This example is based on **real data** from to illustrate the **potential use** of DCMs in an **authentic context** for **research purposes**
- The walk-through of different computations is done for **didactic purposes** with **lots of resources** provided to **explain computations**



While the **model-data fit** for these data is **acceptable**, the **statistical properties** for several items are **not optimal** under this model.

We would **NOT** be using this model with these data for **operational reporting** purposes.

6.6 Checklist Selection



6.7 Bookmark Estimation Checklist



6.8 Estimation Checklist (I)

Estimation Checklist (I)	
MPLUS CONVERGENCE PROBLEMS	SUGGESTED SOLUTION
Do you receive a warning message that some input may be truncated?	Lines > 90 characters are truncated and the specifications are ignored. In the <i>Mplus</i> input file, you can amend this.
Have you used an alternative estimator to MLR to achieve convergence?	MLF often eliminates error messages. The same fit indices are available with MLR. Change the estimator in the <i>Mplus</i> input file, however, remember to use the same estimator for all models.
Have you added class constraints?	In the <i>Mplus</i> input file (under MODEL command), constraints can be added to a subset (see below) or all classes. MODEL CONSTRAINT: m1>-10; ! Constrain latent class 1 estimates to > -10 m2>-10; ! Constrain latent class 2 estimates to > -10 etc.
Have you investigated the size of the non-converged solution's parameters?	Items with parameters > 10 or < -10 often cause convergence issues. In the <i>Mplus</i> input file (under MODEL command), simplify the model for these items by removing the highest interaction terms incrementally. Set constraints for the parameter ranges (constraints highlighted in bold below): MODEL CONSTRAINT: NEW(L1_0 L1_11 L1_12 L1_212); T1_1=-(L1_0); T1_2=-(L1_0+L1_12); T1_3=-(L1_0+L1_11); T1_4=-(L1_0+L1_11+L1_12+L1_212); L1_11=0; L1_12=0; L1_212=L1_11; L1_212=L1_12; L1_0 > -10; L1_0 < -10; L1_11 > -10; L1_11 < -10; L1_12 > -10; L1_12 < -10; L1_212 > -10; L1_212 < -10;
If problems persist consider removing the item(s). Note: Any items removed from the model should be justifiable on statistical and conceptual grounds.	

6.9 Estimation Checklist: ID Variable

Estimation Checklist: ID Variable

If your dataset does NOT contain an ID variable:

when you use the **SAS** macro to run the LCDM your **.dat** file will change to include an **additional column** indicating the ID of each person

File	Edit	Format	View	Help									
1	1	1	1	1	1	1	1	0	1	0	0		
2	1	1	1	0	0	1	1	1	1	0	1	0	
3	1	1	0	0	0	0	1	1	1	1	0	0	
4	1	0	1	0	0	0	0	1	0	1	0	0	
5	1	1	1	1	1	1	1	1	1	0	0	1	
6	0	1	0	0	0	0	1	1	0	1	0	0	
7	0	1	1	0	1	1	0	1	1	1	0	1	
8	1	0	1	1	1	0	0	1	1	0	0	1	0
9	1	0	1	0	1	0	0	1	0	1	1	0	0
10	0	1	1	0	0	0	0	1	0	1	0	1	0

6.10 Estimation Checklist: Convergence



Estimation Checklist: Convergence

If the LCDM does NOT successfully converge:

- **alter** the **Mplus input file** that is generated by the **SAS macro** using suggested solutions from the '**Estimation**' checklist
- **re-run** the model from within **Mplus**
- **remember** to use the **original .dat file WITHOUT the ID variable**

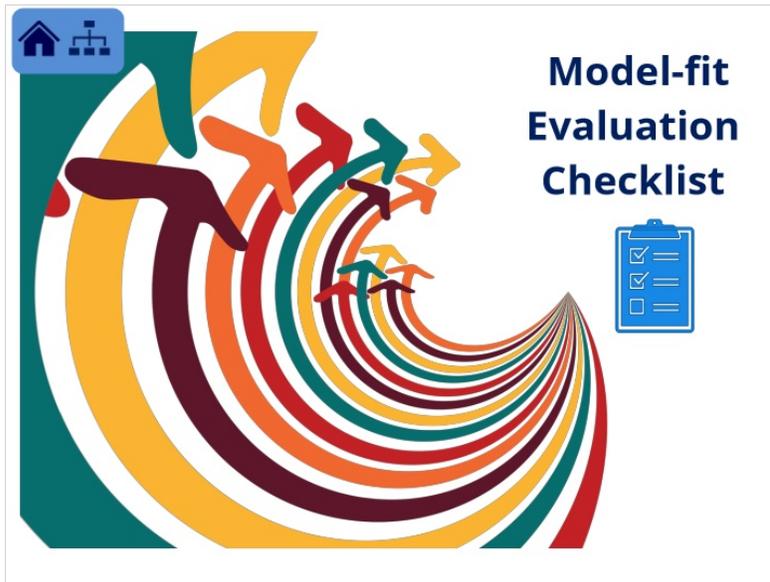
6.11 Bookend: Estimation Checklist



This is the end of this part.

[Checklist Selection](#)

6.12 Bookmark: Evaluation Checklist



6.13 Model Fit Selection



6.14 Bookmark: Absolute Fit



6.15 Absolute Global Fit: General Principle

Absolute Global Fit: General Principle

Whether you are evaluating one or multiple DCMs, **assess each model independently** to determine whether the model provides a good fit to the data based on **bivariate model-data fit information**

6.16 Absolute Global Fit: Computation (I)

  **Absolute Global Fit: Computation (I)**

- **Step 1:** Calculate the total number of item pairs:

$$\frac{\text{Number of items} \times (\text{number of items} - 1)}{2}$$


6.17 Absolute Global Fit: Computation (II)

  **Absolute Global Fit: Computation (II)**

- **Step 2:** Compare the number of item pairs to the 'Overall Bivariate Pearson Chi-Square' and 'Overall Bivariate Log-Likelihood Chi-Square' values:

 If the bivariate values are **lower** than the total number of item pairs, the model provides a **good** fit to the data

 If the bivariate values are **higher** than the total number of item pairs, the model provides a **poor** fit to the data



6.18 Absolute Global Fit: Example



Absolute Global Fit: Example

13 items: $N \text{ of items} \times (N \text{ of items} - 1) / 2$
 $13 \times (13 - 1) / 2 = \mathbf{78 \text{ item pairs}}$

BIVARIATE MODEL FIT INFORMATION				
		Estimated Probabilities		
Variable	Variable	H1	H0	Standardized Residual (z-score)
MITEM1	MITEM2			
Category 1	Category 1	0.046	0.038	0.719
Category 1	Category 2	0.154	0.162	-0.369
Category 2	Category 1	0.121	0.129	-0.377
Category 2	Category 2	0.679	0.671	0.265
Bivariate Pearson Chi-Square				0.758
Bivariate Log-Likelihood Chi-Square				0.731

Overall Bivariate Pearson Chi-Square	70.247
Overall Bivariate Log-Likelihood Chi-Square	70.657

Interpretation:
Since $70 < 78$, the model provides a **good fit** to the data

6.19 Absolute Local Fit: General Principle



Absolute Local Fit: General Principle



Whether you are evaluating one or multiple DCMs, **assess each model independently** to determine whether the model provides a good fit to the data based on **the number of observed and expected pairs of misfitting items**

6.20 Absolute Local Fit: Computation (I)

🏠 **Absolute Local Fit: Computation (I)**

Step 1: Identify the Number of Observed Pairs of Misfitting Items

In the 'Bivariate Model Fit Information' section of the Mplus output, inspect the 'Bivariate Pearson Chi-Square' and 'Bivariate Log-Likelihood Chi-Square' values for each item:

Identify the number of item pairs with values ≥ 3.84



6.21 Absolute Local Fit: Computation (II)

🏠 **Absolute Local Fit: Computation (II)**

Step 2: Identify the Number of Expected Pairs of Misfitting Items

$0.05 \times$ total number of items pairs

The number of observed misfitting pairs should be no larger than the number of misfitting item pairs expected by chance.



6.22 Absolute Local Fit: Example (I)

BIVARIATE MODEL FIT INFORMATION				
Variable	Variable	Estimated Probabilities		Standardized Residual (z-score)
		H1	H0	
MITEM1	MITEM2			
Category 1	Category 1	0.046	0.038	0.719
Category 1	Category 2	0.154	0.162	-0.369
Category 2	Category 1	0.121	0.129	-0.377
Category 2	Category 2	0.679	0.671	0.265
Bivariate Pearson Chi-Square				0.758
Bivariate Log-Likelihood Chi-Square				0.731
Well-fitting item pair (value <3.84)				
MITEM4	MITEM9			
Category 1	Category 1	0.050	0.037	1.168
Category 1	Category 2	0.532	0.545	-0.434
Category 2	Category 1	0.007	0.020	-1.519
Category 2	Category 2	0.411	0.398	0.424
Bivariate Pearson Chi-Square				3.768
Bivariate Log-Likelihood Chi-Square				4.376
Misfitting item pair (value > 3.84)				

Overall, we observed 4 misfitting item pairs for our data

6.23 Absolute Local Fit: Example (II)

Calculate the number of misfitting item pairs expected by chance:

$$0.05 \times \text{total number of item pairs}$$

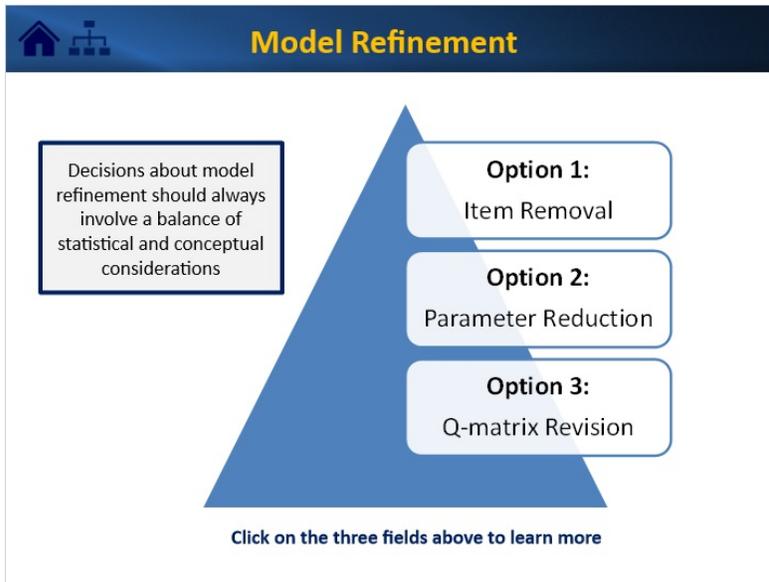
$$0.05 \times 78 = 4 \text{ misfitting item pairs expected by chance}$$

If the number of observed misfitting item pairs is **less than** the number of expected misfitting item pairs, the model provides a **good fit** to the data

If the number of observed misfitting item pairs is **greater than** the number of expected misfitting item pairs, the model provides a **poor fit** to the data

In this example, there were 4 misfitting item pairs expected by chance and 4 misfitting item pairs observed – this means the model provides a good fit

6.24 Model Refinement



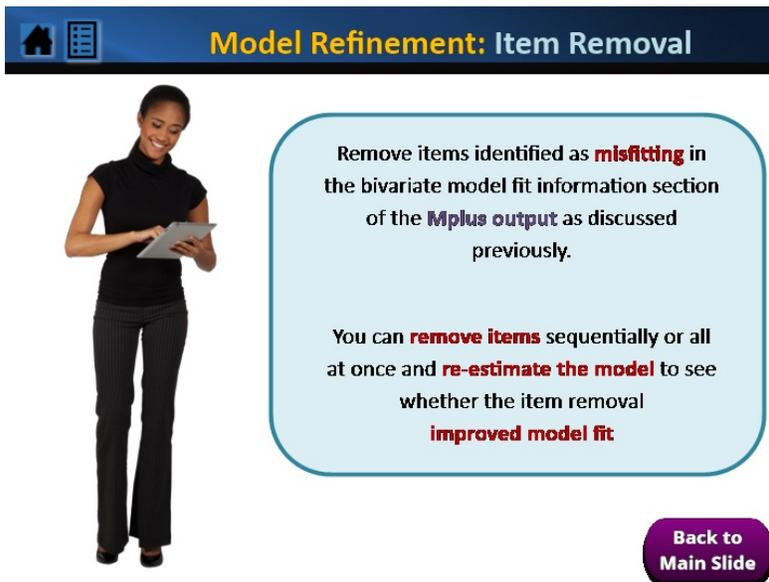
Model Refinement

Decisions about model refinement should always involve a balance of statistical and conceptual considerations

- Option 1: Item Removal
- Option 2: Parameter Reduction
- Option 3: Q-matrix Revision

Click on the three fields above to learn more

Item Removal (Slide Layer)



Model Refinement: Item Removal



Remove items identified as **misfitting** in the bivariate model fit information section of the **Mplus output** as discussed previously.

You can **remove items** sequentially or all at once and **re-estimate the model** to see whether the item removal **improved model fit**

[Back to Main Slide](#)

Parameter Reduction (Slide Layer)

Model Refinement: Parameter Reduction



In the fully specified (saturated) LCDM:

non-significant interaction terms may be removed from the model, which may help to **improve model fit**

[Back to Main Slide](#)

Q-matrix Revision (Slide Layer)

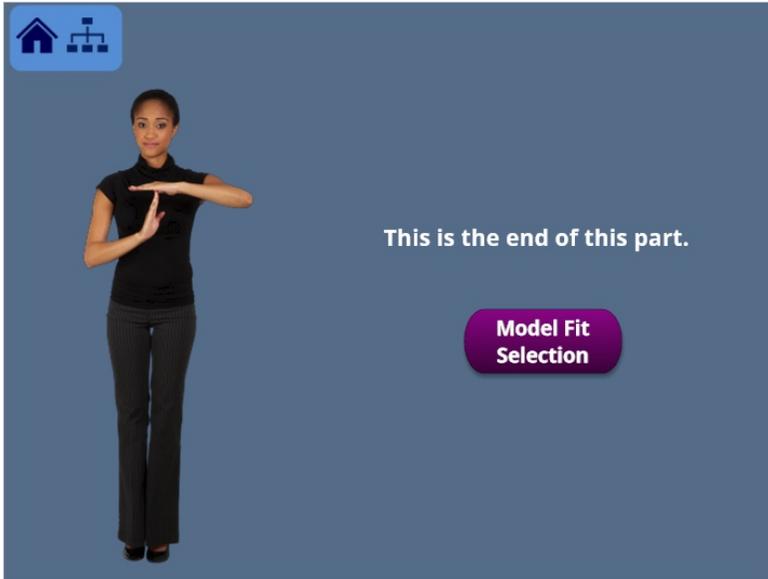
Model Refinement: Q-matrix Modifications



You may refine the Q-matrix by **removing an attribute, combining attributes, and/or changing item-to-attribute alignment** (i.e., switching 0s and 1s).

[Back to Main Slide](#)

6.25 Bookend: Absolute Fit



6.26 Bookmark: Relative Fit



6.27 Relative Model Fit: General Principles (I)

🏠 **Relative Model Fit: General Principles (I)**



Whenever you are **evaluating multiple DCMs**, only those models displaying **acceptable absolute model fit** should be subsequently **evaluated for relative model fit**

Item Removal (Slide Layer)



Option 1: Removal of misfitting items

- In some cases, the bivariate model fit information may indicate that the model does not provide a good fit to the data. Multiple misfitting items may be observed, with some loading on a single attribute and others loading on multiple attributes.
- Relatedly, if the number of observed misfitting item pairs exceeds the number expected by chance, the user may wish to consider removing problematic items and re-estimating the model to improve model fit.
 - All items repeatedly identified as misfitting (i.e., Bivariate Pearson Chi-Square and Bivariate Log-Likelihood Chi-Square values ≥ 3.84) could be flagged as candidates for removal in the first instance. These items could be removed sequentially or all at once, and the model re-estimated.

Note: Decisions about model refinement should involve a balance of statistical and conceptual considerations

Only models displaying acceptable absolute model fit should be subsequently evaluated for relative model fit.
Note: Obtaining absolute model fit may be challenging in large samples.

Parameter Reduction (Slide Layer)



Option 2: Removal of non-significant interaction terms (LCDM only)

- As discussed in detail in the DCM Interpretation Checklist (see 'Parameter estimates'), in the LCDM, non-significant interaction terms may be removed from the model. This may help to improve model fit.

Note: Decisions about model refinement should involve a balance of statistical and conceptual considerations

Only models displaying acceptable absolute model fit should be subsequently evaluated for relative model fit.
Note: Obtaining absolute model fit may be challenging in large samples.

Q-matrix Revision (Slide Layer)



Option 3: Changes to Q-matrix

- The user may wish to consider if there are any empirical misspecifications in the Q-matrix and whether it is necessary to refine the Q-matrix by removing an attribute and/or changing item-to-attribute alignment.

Note: Decisions about model refinement should involve a balance of statistical and conceptual considerations

Only models displaying acceptable absolute model fit should be subsequently evaluated for relative model fit.
Note: Obtaining absolute model fit may be challenging in large samples.

6.28 Relative Model Fit: General Principles



Relative Model Fit: General Principles



- To compare models, all candidate models must:
 - have the same number of items
 - be based on the same sample
- Currently, **fit indices in DCMs remain in their infancy** and practical fit indices used in other modeling approaches such as SEM are **not yet available for DCMs**

Item Removal (Slide Layer)



Option 1: Removal of misfitting items

- In some cases, the bivariate model fit information may indicate that the model does not provide a good fit to the data. Multiple misfitting items may be observed, with some loading on a single attribute and others loading on multiple attributes.
- Relatedly, if the number of observed misfitting item pairs exceeds the number expected by chance, the user may wish to consider removing problematic items and re-estimating the model to improve model fit.
 - All items repeatedly identified as misfitting (i.e., Bivariate Pearson Chi-Square and Bivariate Log-Likelihood Chi-Square values ≥ 3.84) could be flagged as candidates for removal in the first instance. These items could be removed sequentially or all at once, and the model re-estimated.

Note: Decisions about model refinement should involve a balance of statistical and conceptual considerations

Only models displaying acceptable absolute model fit should be subsequently evaluated for relative model fit.
Note: Obtaining absolute model fit may be challenging in large samples.

Parameter Reduction (Slide Layer)



Option 2: Removal of non-significant interaction terms (LCDM only)

- As discussed in detail in the DCM Interpretation Checklist (see 'Parameter estimates'), in the LCDM, non-significant interaction terms may be removed from the model. This may help to improve model fit.

Note: Decisions about model refinement should involve a balance of statistical and conceptual considerations

Only models displaying acceptable absolute model fit should be subsequently evaluated for relative model fit.
Note: Obtaining absolute model fit may be challenging in large samples.

Q-matrix Revision (Slide Layer)



Option 3: Changes to Q-matrix

- The user may wish to consider if there are any empirical misspecifications in the Q-matrix and whether it is necessary to refine the Q-matrix by removing an attribute and/or changing item-to-attribute alignment.

Note: Decisions about model refinement should involve a balance of statistical and conceptual considerations

Only models displaying acceptable absolute model fit should be subsequently evaluated for relative model fit.
Note: Obtaining absolute model fit may be challenging in large samples.

6.29 Relative Model Fit: Entropy



Relative Model Fit: Entropy

- Entropy is based on the **posterior class membership probabilities**
- Entropy is used in latent class analysis to evaluate how well each of the classes is **separated and represented** by the data
- Entropy values **range from 0 to 1** with high values preferred

 **In the context of DCMs, however, entropy is misleading as poorly fitting models often have a high entropy value** 

Item Removal (Slide Layer)



Option 1: Removal of misfitting items

- In some cases, the bivariate model fit information may indicate that the model does not provide a good fit to the data. Multiple misfitting items may be observed, with some loading on a single attribute and others loading on multiple attributes.
- Relatedly, if the number of observed misfitting item pairs exceeds the number expected by chance, the user may wish to consider removing problematic items and re-estimating the model to improve model fit.
 - All items repeatedly identified as misfitting (i.e., Bivariate Pearson Chi-Square and Bivariate Log-Likelihood Chi-Square values ≥ 3.84) could be flagged as candidates for removal in the first instance. These items could be removed sequentially or all at once, and the model re-estimated.

Note: Decisions about model refinement should involve a balance of statistical and conceptual considerations

Only models displaying acceptable absolute model fit should be subsequently evaluated for relative model fit.
Note: Obtaining absolute model fit may be challenging in large samples.

Parameter Reduction (Slide Layer)



Option 2: Removal of non-significant interaction terms (LCDM only)

- As discussed in detail in the DCM Interpretation Checklist (see 'Parameter estimates'), in the LCDM, non-significant interaction terms may be removed from the model. This may help to improve model fit.

Note: Decisions about model refinement should involve a balance of statistical and conceptual considerations

Only models displaying acceptable absolute model fit should be subsequently evaluated for relative model fit.
Note: Obtaining absolute model fit may be challenging in large samples.

Q-matrix Revision (Slide Layer)



Option 3: Changes to Q-matrix

- The user may wish to consider if there are any empirical misspecifications in the Q-matrix and whether it is necessary to refine the Q-matrix by removing an attribute and/or changing item-to-attribute alignment.

Note: Decisions about model refinement should involve a balance of statistical and conceptual considerations

Only models displaying acceptable absolute model fit should be subsequently evaluated for relative model fit.
Note: Obtaining absolute model fit may be challenging in large samples.

6.30 Model Nesting Selection



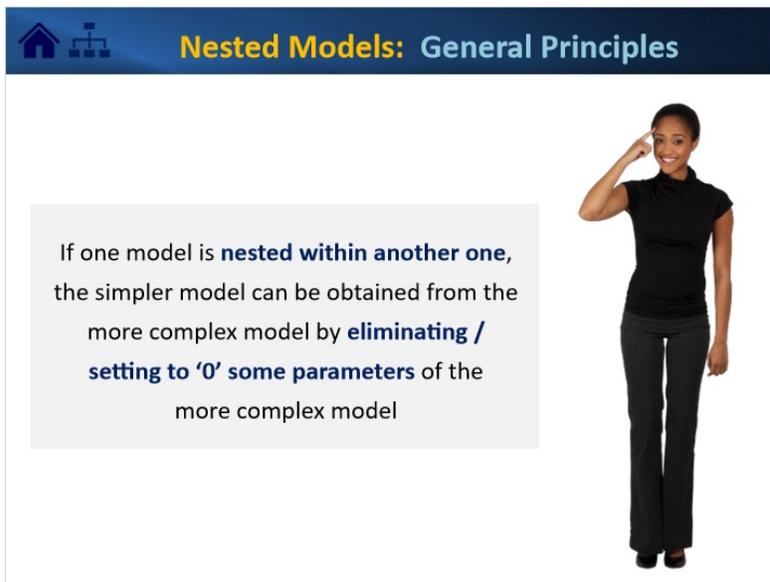
Home icon and a tree diagram icon.

Nested Models

Non-nested Models

Section End

6.31 Nested Models: General Principles



Home icon and a tree diagram icon.

Nested Models: General Principles

If one model is **nested within another one**, the simpler model can be obtained from the more complex model by **eliminating / setting to '0'** some **parameters** of the more complex model

A woman in a black top and dark pants, standing with her hand on her forehead in a thoughtful pose.

6.32 Nested Models: Illustrations



Nested Models: Illustrations

- A **Rasch model** is nested within a two-parameter model
- A **model with one latent dimension (unidimensional model)** is nested within a model with two latent dimensions (multidimensional model)
- A **model with main effect parameters only** is nested within a model that also has interaction effect parameters



6.33 Nested Models: Evaluation



Nested Models: Evaluation

- Locate the '**Model Fit Information**' section in the **Mplus output**
- Retrieve the '**Loglikelihood**' and '**Number of Free Parameters**' values for the null model (**H0: simpler model**) and the alternative model (**H1: more complex model**)

Using Excel, calculate the likelihood ratio test p -value:

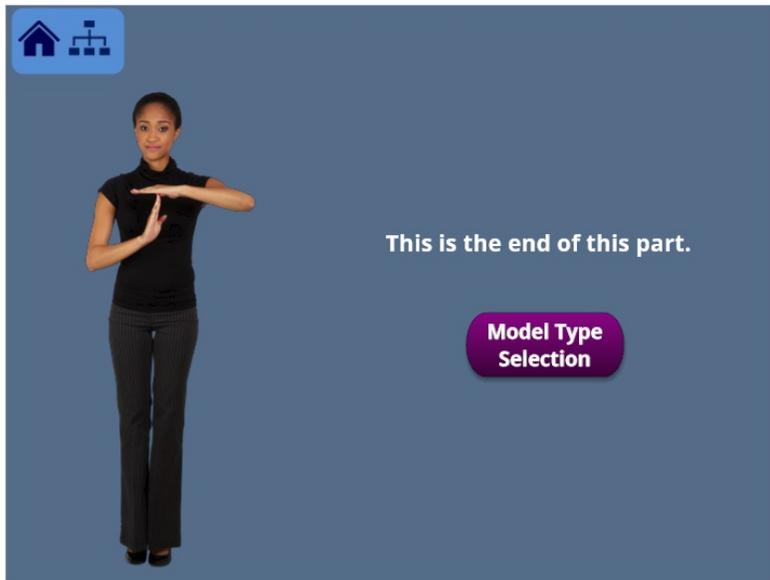
- $p\text{-value} \geq 0.05$ = simpler model provides acceptable fit
- $p\text{-value is} < 0.05$ = more complex model provides better fit

6.34 Nested Models: Example

Nested Models: Example	
LCDM (alternative / more complex / H1 model)	
Loglikelihood value	1825.370
Number of free parameters	39
C-RUM (null / simpler / H0 model)	
Loglikelihood value	-1825.161
Number of free parameters	34
LR: $-2 \times (\text{loglikelihood value H0 model} - \text{loglikelihood value H1 model})$	7301.06
df_{LR}: df of H1 model – df of H0 model	5
LR p-value: $\text{chidist}(\text{LR}, df_{LR})$	0.000

The LCDM provides a **better fit** than the C-RUM

6.35 Bookend: Nested Models



6.36 Non-nested Models: General Principles

🏠 📊 Non-nested Models: General Principles

Two models are **non-nested** if neither model can be represented as a **special case** of the other.

In other words, one model is **NOT** a **restricted version** of the other model



6.37 Non-nested Models: Evaluation

🏠 📊 Non-nested Models: Evaluation

- Locate the 'Model Fit Information' section in the **Mplus output**
- Compare the '**AIC**', '**BIC**', and '**SSABIC**' values across key DCMs:

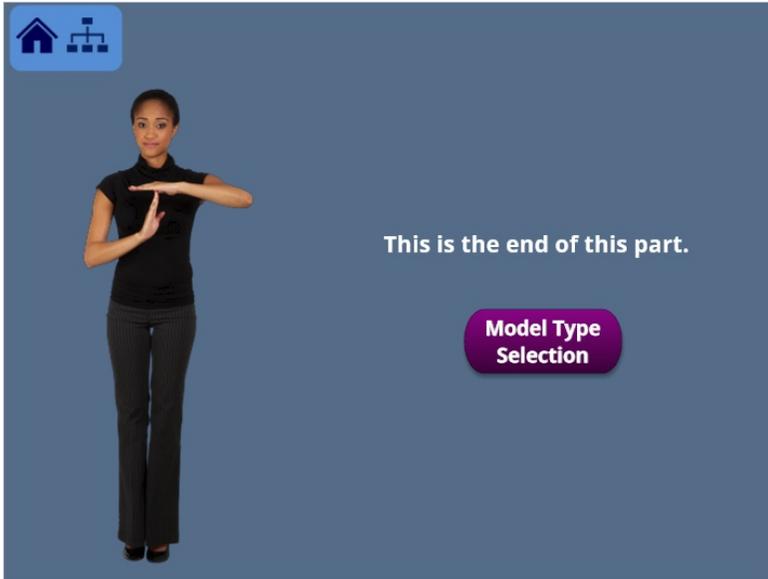
lower values indicate better model fit for a given model

```
MODEL FIT INFORMATION
Number of Free Parameters          39

Loglikelihood
  H0 Value                        -1825.370
  H0 Scaling Correction Factor    0.8505
  for MLR

Information Criteria
  Akaike (AIC)                    3728.740
  Bayesian (BIC)                  3870.497
  Sample-Size Adjusted BIC       3746.831
  (n* = (n + 2) / 24)
```

6.38 Bookend: Non-nested Models



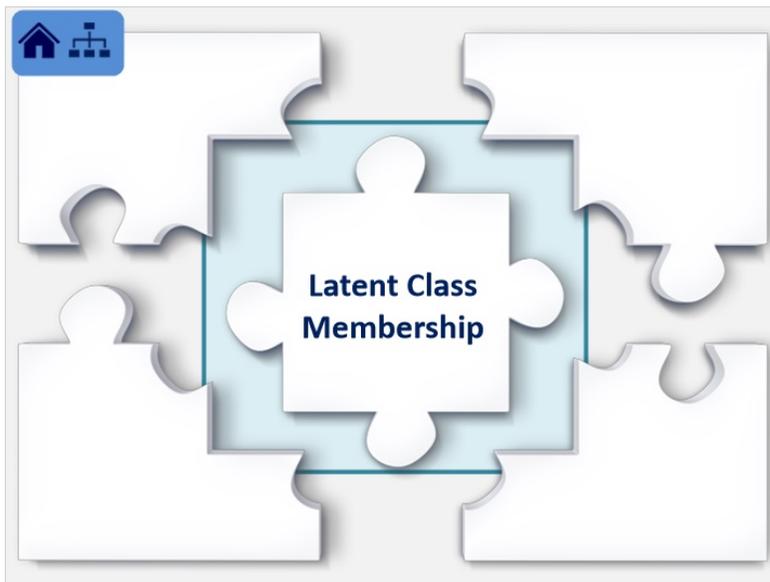
6.39 Bookmark: Interpretation Checklist



6.40 Interpretation Topic Selection



6.41 Bookmark: Membership



6.42 Latent Class Membership: Counts

  **Latent Class Membership: Counts**

Locate the 'Final class counts and proportions for the latent classes based on the estimated model' section in the Mplus output

FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASSES
BASED ON THE ESTIMATED MODEL

Latent Classes	Number of examinees in each latent class	Proportion of examinees in each latent class
1	78.81545	0.28148
2	81.54375	0.29123
3	40.58269	0.14494
4	79.05811	0.28235

Class 1 = [0,0]
Class 2 = [0,1]
Class 3 = [1,0]
Class 4 = [1,1]

6.43 Latent Class Proportions: Mastery Profiles

  **Latent Class Proportions: Mastery Profiles**

Identify which attributes are **mastered and not mastered** in each latent class by retrieving the SAS output file called '_classcounts'

	class	estcount	estprop	classatt1	classatt2
1	1	78.81545	0.28148	0	0
2	2	81.54375	0.29123	0	1
3	3	40.58269	0.14494	1	0
4	4	79.05811	0.28235	1	1

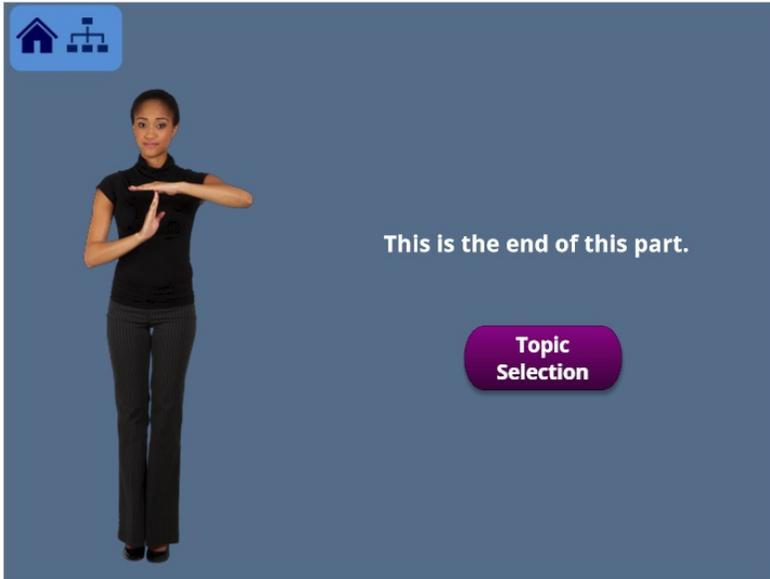
Class 1 [0,0]: examinees who have not mastered either of the attributes

Class 2 [0,1]: examinees who have mastered only attribute 2

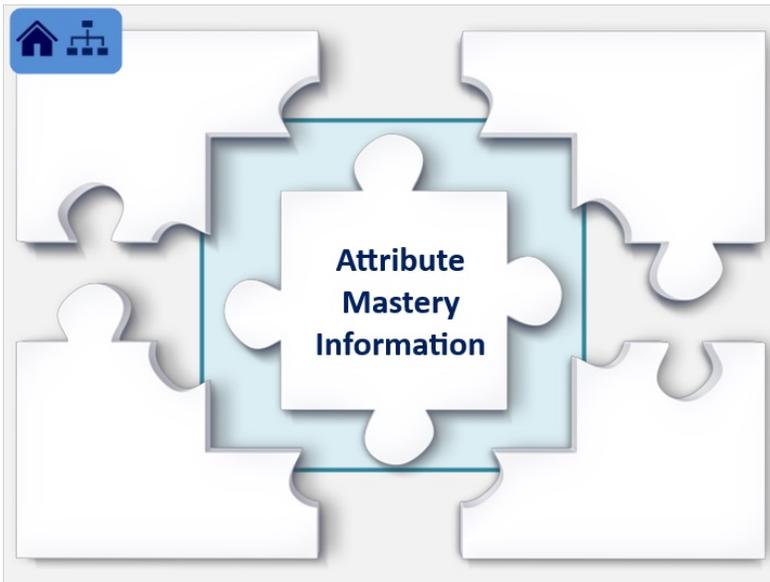
Class 3 [1,0]: examinees who have mastered only attribute 1

Class 4 [1,1]: examinees who have mastered attribute 1 and attribute 2

6.44 Bookend: Proportions



6.45 Bookmark: Mastery



6.46 Marginal Attribute Mastery: Probabilities



Marginal Attribute Mastery: Probabilities

For the **full / saturated LCDM**, mastery probabilities are available for each attribute in the **SAS output file** called '**_respondents**'

	ID	cprob1	cprob2	cprob3	cprob4	class	prob_attribute1	prob_attribute2
1	1	0.02709	0.03823	0.312	0.622	4	0.93469	0.66063
2	2	0.0622	0.10936	0.243	0.585	4	0.82845	0.69455
3	3	0.32058	0.4524	0.076	0.151	2	0.22702	0.60357
4	4	0.97556	0.0216	0.003	15E-5	1	0.00285	0.02175
5	5	0.00534	0.0144	0.317	0.664	4	0.98026	0.67792

For all other DCMs, it is necessary to **calculate these probabilities manually** from the probabilities of latent class membership.

Manual Computation

Part 1 (Slide Layer)

Part 1: Transfer Posterior Probabilities

To facilitate calculation of the **marginal probabilities**, transfer the **Mplus .dat file** containing the **posterior probabilities of class membership into Excel**

- Open the file '**respondents_ITEMSpsychiatry**' in the supporting documents
- The initial series of columns represent the **item response data**
- The next column represents the **ID variable**
- The next series of columns contain the **posterior latent class membership probabilities for each class**
- The next column indicates the **latent class with highest posterior probability**

Part 1

Part 2

Part 3

Back

Part 2 (Slide Layer)

Part 2: Calculate Marginal Probabilities

- For each attribute, use the **SUM function in Excel** to calculate the marginal probability of mastery for each attribute as the **sum of the latent class membership probabilities** across all of the latent classes for which the attribute is actually mastered
- Since **latent class membership is mutually exclusive**, the sum of all probabilities across all latent classes is always 1 and computing sums is a legitimate statistical operation for this purpose
- Perform this equation for examinee 1. **Drag the formula down the column to fill the remaining cells.** Repeat for each attribute.

Part 1

Part 2

Part 3

Back

Part 3 (Slide Layer)

Part 3: Example

	Q	R	S	T	U	V	X	Y
15	Posterior latent class membership probability				Latent class with highest posterior probability	Marginal probabilities		
	Latent class 1	Latent class 2	Latent class 3	Latent class 4		Attribute 1: [0,0,1,1]	Attribute 2: [0,1,0,1]	
16	[0,0]	[0,1]	[1,0]	[1,1]				
17	0.02709	0.03823	0.31229	0.6224	4	0.93469	0.66063	
18	0.0622	0.10936	0.24326	0.58519	4	0.82845	0.69455	
19	0.32058	0.4524	0.07585	0.15117	2	0.22702	0.60357	
20	0.97556	0.0216	0.0027	0.00015	1	0.00285	0.02175	
21	0.00534	0.0144	0.31674	0.66352	4	0.98026	0.67792	
22	0.58578	0.40042	0.00762	0.00618	1	0.0138	0.4066	

$$=SUM(Q17*{0},R17*{0},S17*{1},T17*{1})$$

$$=SUM(Q17*{0},R17*{1},S17*{0},T17*{1})$$

Part 1

Part 2

Part 3

Back

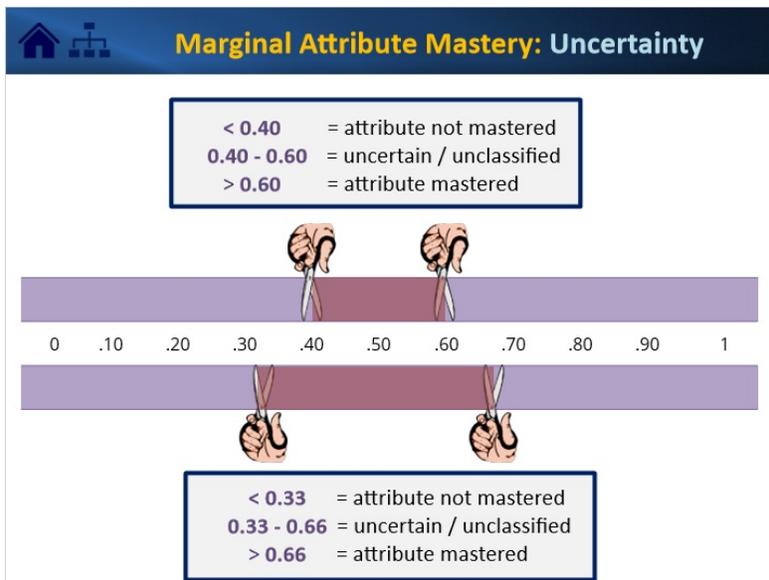
6.47 Marginal Attribute Mastery: Classifications

🏠 📊 **Marginal Attribute Mastery: Classifications**

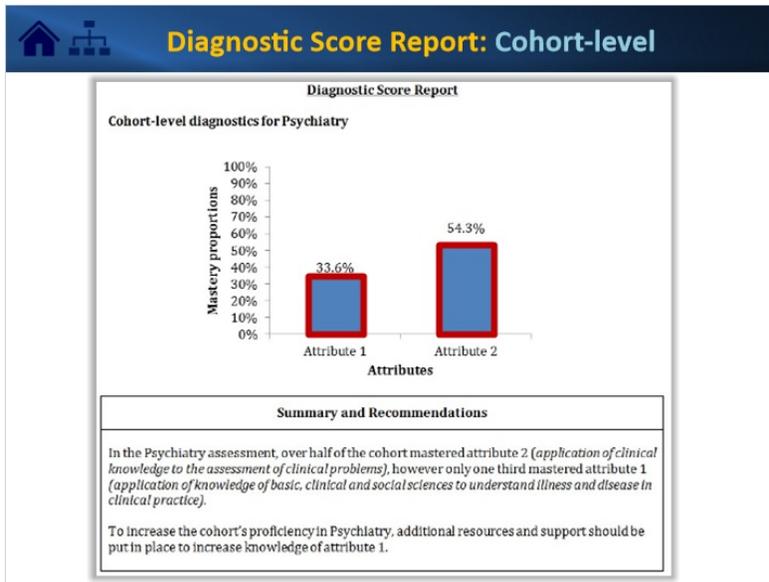


- For summary and descriptive purposes, it is useful to **convert the probabilities into binary mastery states**.
- However, there is **no consensus** on where the conversion thresholds should be as it depends on how **“conservative”** an analyst wants to be
- Larger uncertainty region → more conservative mastery state assignment / **more evidence needed** for state assignment

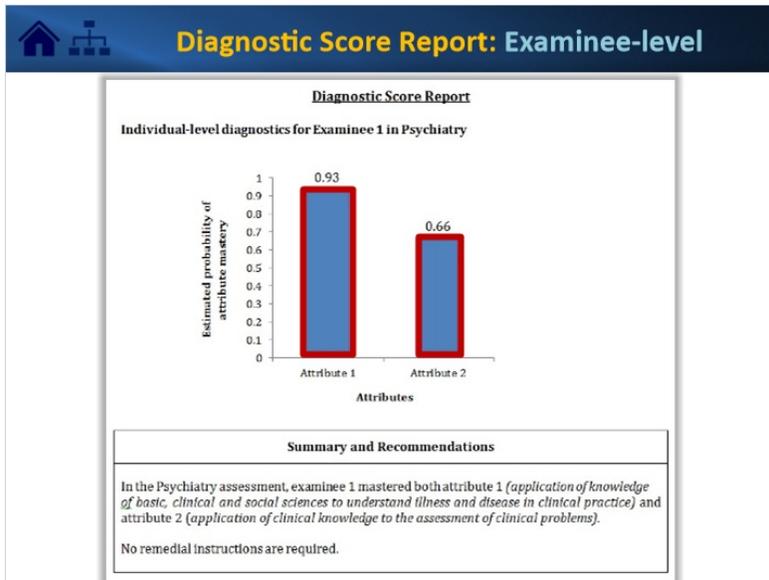
6.48 Marginal Attribute Mastery: Uncertainty



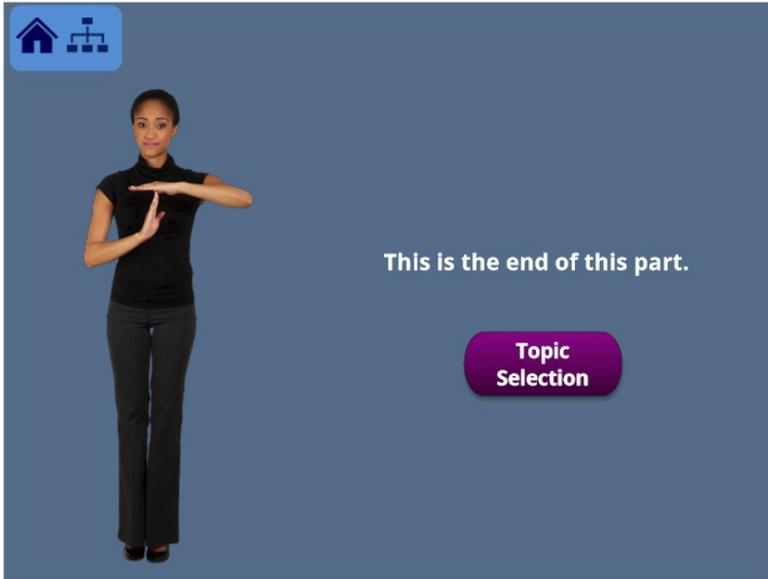
6.49 Diagnostic Score Report: Cohort-level



6.50 Diagnostic Score Report: Examinee-level



6.51 Bookend: Mastery



6.52 Bookmark: Quality



6.53 Diagnostic Quality: General Principles

 **Diagnostic Quality: General Principles**

- **Parameter estimates** provide useful information on the diagnostic quality of each item. This information may be of particular interest to course co-coordinators
- **Parameter estimates** are located in the 'Estimate' column of the 'New / Additional Parameters' section in the **Mplus output**
- Next, we will demonstrate how to use these parameter estimates – specifically **intercept and main effect parameters** – to calculate and interpret associated **response probabilities**

6.54 Interaction Selection



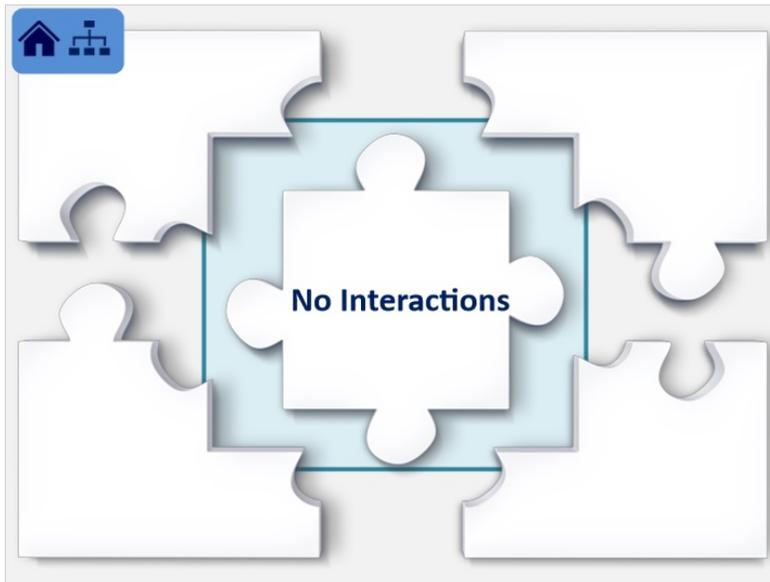


No Interactions

Interactions

Part End

6.55 Bookmark: No Interactions



6.56 Intercept: Computational Steps (I)

**Intercept: Computational Steps (I)**

Step 1: Identify Relevant Parameters

Retrieve the **intercept parameter estimate** for each item from the **'New / Additional Parameters'** section of the **Mplus output**:

L6_0: L6 indicates item 6, 0 indicates intercept parameter



6.57 Intercept: Computational Steps (II)



Intercept: Computational Steps (II)

Step 2: Calculate Response Probability

Enter the **intercept parameter estimate** into the following formula using **Excel**, which is the **conversion of the logit into a probability**:

$$\frac{\exp(\text{'intercept parameter estimate'})}{[1 + \exp(\text{'intercept parameter estimate'})]}$$



Formula
Derivation

Derivation (Slide Layer)

Derivation of Formula

$$\text{Logit}(Y_{ri}) = \lambda_{i,0}$$

↓

$$\text{Log} \left(\frac{P(Y_{ri})}{1 - P(Y_{ri})} \right) = \lambda_{i,0}$$

↓

$$P(Y_{ri}) = \frac{\exp(\lambda_{i,0})}{1 + \exp(\lambda_{i,0})}$$

Back

6.58 Intercept: Example

🏠 📊 Intercept: Example

Parameter label	Parameter estimate	Calculation	
L6_0	-2.403	exp(L6_0)=	0.09
		1+exp(L6_0)=	1.09
		exp(L6_0)/1+exp(L6_0)=	0.08

Perform this calculation for **all intercept parameters** regardless of whether or not the estimate is **statistically significant**



6.59 Intercept: Interpretation

🏠 📊 Intercept: Interpretation

- Probability values **range from 0 to 1**

- For the intercept, the resulting probabilities indicate the chances of a **correct response for non-masters of all required attributes for the item** (i.e., “guessing”)

- **High probabilities** thus indicate that the item may **not be well measured by the attribute** and vice versa.

- In the example on the previous slide, non-masters of all required attributes have a **very low probability (8%)** of a correct response to item 6 (i.e., **this is a good item**)

6.60 Main Effect: Computational Steps (I)



Main Effect: Computational Steps (I)

Step 1: Identify Relevant Parameters

Retrieve the **main effect parameter estimate** for each item from the **'New / Additional Parameters'** section of the **Mplus output**:

L6_11:

- L6 indicates **item 6**
- the first '1' after the underscore indicates a **main effect**
- the second '1' after the underscore indicates that this item measures **attribute 1**



6.61 Main Effect: Computational Steps (II)



Main Effect: Computational Steps (II)

Step 2: Calculate Response Probability

Enter the main effect parameter estimate into the following formula using **Excel**, which is the **conversion of the logit into a probability**:

$$\frac{\exp(\text{'intercept' + 'main effect'})}{[1 + \exp(\text{'intercept' + 'main effect'})]}$$

Perform this calculation for **all main effect parameters**, regardless of whether or not the estimate is **statistically significant**



Derivation (Slide Layer)

Derivation of Formula

$$\text{Logit}(Y_{ri}) = \lambda_{i,0} + \lambda_{i,1,(a)}$$



$$\text{Log} \left(\frac{P(Y_{ri})}{1-P(Y_{ri})} \right) = \lambda_{i,0} + \lambda_{i,1,(a)}$$



$$P(Y_{ri}) = \frac{\exp(\lambda_{i,0} + \lambda_{i,1,(a)})}{1 + \exp(\lambda_{i,0} + \lambda_{i,1,(a)})}$$

Back

6.62 Main Effect: Example

 **Main Effect: Example**

L6_11	1.205		
		L6_0 + L6_11=	-1.20
		exp(L6_0 + L6_11)=	0.30
		1+exp(L6_0 + L6_11)=	1.30
		exp(L6_0 + L6_11)/1+exp(L6_0 + L6_11)=	0.23



6.63 Main Effect: Interpretation



Main Effect: Interpretation

- The **parameter estimate** indicates the **increase in logit** that masters of the attribute in question have relative to non-masters of all attributes

In the example, examinees who have mastered attribute 1 have an **increase in logit of 1.205** over examinees who are non-masters

- Probability values **range from 0 to 1** and express the relative advantage of examinees who are masters more naturally

In the example, examinees who are masters of attribute 1 **have a relatively low probability (23%)** of answering item 6 correctly

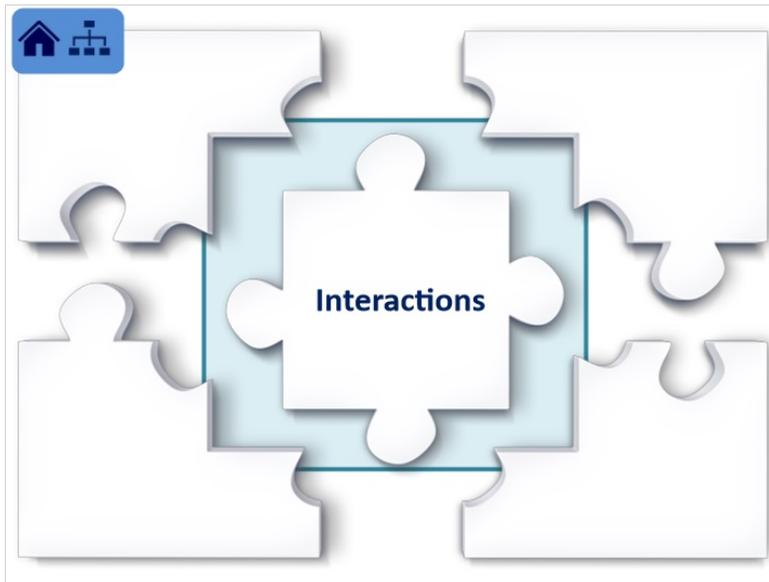
6.64 Bookend: Models without Interaction



This is the end of this part.

Topic Selection

6.65 Bookmark: Interactions



6.66 Interactions: General Principles (I)

 **Interactions: General Principles (I)**

Calculate the response probabilities when interaction terms are present (i.e., for the full / saturated LCDM)

- Regardless of **statistical significance**, if an interaction term is included in the model it **must be included** in all predicted probability calculations
- When **interaction terms are present**, the main effects are called **conditional (or simple) main effects**. This means that their interpretation depends on **other variables in the model being 0**
- If the interaction parameter is **NOT statistically significant** it can be omitted from the model. **Removing the interaction term** may be useful if the model provides a poor fit to the data and **requires modification to improve fit**

6.67 Interactions: General Principles (II)



Interactions: General Principles (II)

- An interaction term is an **over-additive** or an **under-additive** effect
- If the interaction parameter estimate is **negative**, there is an **under-additive effect** of mastering both attributes
 -  mastery of one attribute is sufficient to have a high chance of answering the item correctly
- If the raw interaction parameter estimate is **positive**, there is an **over-additive effect** of mastering both attributes
 -  there is a “bonus” for mastery of both attributes

6.68 All Parameters: Computational Steps (I)



All Parameters: Computational Steps (I)

Step 1: Identify Relevant Parameters

Retrieve the raw intercept, main effect and interaction parameter estimates for each item from the ‘New / Additional Parameters’ section of the **Mplus** output:

		Estimate
New/Additional Parameters		
Intercept	L2_0	0.632
Main Effects {	L2_11	2.611
	L2_12	0.725
Interaction Effect	L2_212	0.175

6.69 All Parameters: Computational Steps (II)

**All Parameters: Computational Steps (II)**

Step 2: Calculate Predicted Logits

Calculate the **predicted logit for each latent class** using this formula in **Excel**:

$$\text{logit} = \text{'intercept'} + (\text{'main effect for attribute 1'} * A1) + (\text{'main effect for attribute 2'} * A2) + (\text{'interaction effect'} * A1 * A2)$$



Formula Derivation

Derivation (Slide Layer)

Derivation

$$\text{Logit}(Y_{r2} = 1 | \alpha_r) = \lambda_{2,0} + \lambda_{2,1,(1)} + \lambda_{2,1,(2)} + \lambda_{2,2,(12)}$$

↓

$$\text{Log} \left(\frac{P(Y_{r2})}{1 - P(Y_{r2})} \right) = \lambda_{2,0} + \lambda_{2,1,(1)} + \lambda_{2,1,(2)} + \lambda_{2,2,(12)}$$

↓

$$P(Y_{r2}) = \frac{\exp(\lambda_{2,0} + \lambda_{2,1,(1)} + \lambda_{2,1,(2)} + \lambda_{2,2,(12)})}{1 + \exp(\lambda_{2,0} + \lambda_{2,1,(1)} + \lambda_{2,1,(2)} + \lambda_{2,2,(12)})}$$

Back

6.70 All Parameters: Example (I)

All Parameters: Example (I)					
	Intercept	Main Effect for A1	Main Effect for A2	A1*A2 interaction	
Class [A1, A2]	Calculation				Logit
Class 1 [0,0]:	0.632	+ (2.611*0)	+ (0.725*0)	+ (0.175*0*0)	0.632
Class 2 [1,0]:	0.632	+ (2.611*1)	+ (0.725*0)	+ (0.175*1*0)	3.243
Class 3 [0,1]:	0.632	+ (2.611*0)	+ (0.725*1)	+ (0.175*0*1)	1.357
Class 4 [1,1]:	0.632	+ (2.611*1)	+ (0.725*1)	+ (0.175*1*1)	4.143

6.71 All Parameters: Computational Steps (III)

Step 3: Calculate Predicted Response Probabilities

Convert each of the predicted logits into predicted response probabilities using the following formula:

$$\frac{\exp(\text{logit})}{1 + \exp(\text{logit})}$$

logit = 'intercept' + ('main effect for attribute 1' * A1) + ('main effect for attribute 2' * A2) + ('interaction effect' * A1 * A2)

Formula Derivation

Derivation (Slide Layer)

Derivation

$$\text{Logit}(Y_{r2} = 1 | \alpha_r) = \lambda_{2,0} + \lambda_{2,1,(1)} + \lambda_{2,1,(2)} + \lambda_{2,2,(12)}$$



$$\text{Log} \left(\frac{P(Y_{r2})}{1 - P(Y_{r2})} \right) = \lambda_{2,0} + \lambda_{2,1,(1)} + \lambda_{2,1,(2)} + \lambda_{2,2,(12)}$$



$$P(Y_{r2}) = \frac{\exp(\lambda_{2,0} + \lambda_{2,1,(1)} + \lambda_{2,1,(2)} + \lambda_{2,2,(12)})}{1 + \exp(\lambda_{2,0} + \lambda_{2,1,(1)} + \lambda_{2,1,(2)} + \lambda_{2,2,(12)})}$$

Back

6.72 All Parameters: Example (II)

All Parameters: Example (II)		
Class [A1, A2]	Calculation	Probability
Class 1 [0,0]:	$\frac{\exp(0.632)}{(1 + \exp(0.632))}$	= 0.65
Class 2 [1,0]:	$\frac{\exp(3.243)}{(1 + \exp(3.243))}$	= 0.96
Class 3 [0,1]:	$\frac{\exp(1.357)}{(1 + \exp(1.357))}$	= 0.80
Class 4 [1,1]:	$\frac{\exp(4.143)}{(1 + \exp(4.143))}$	= 0.98

6.73 All Parameters: Interpretation

 **All Parameters: Interpretation**

- Respondents who mastered neither attribute 1 nor 2 (**latent class 1**) had a **moderate chance** of answering item 2 correctly (**0.65**)
- Respondents who mastered attribute 1 but not attribute 2 (**latent class 2**) had a **very high chance** of answering item 2 correctly (**0.96**)
- Respondents who did not master attribute 1 but mastered attribute 2 (**latent class 3**) had a **high chance** of answering item 2 correctly (**0.80**)
- Respondents who mastered both attributes 1 and 2 (**latent class 4**) had a **high chance** of answering item 2 correctly (**0.98**)

6.74 Bookend: Models with Interaction





This is the end of this part.

Topic Selection

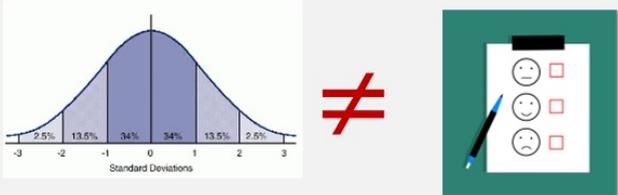
6.75 Bookmark: Reliability



6.76 General Principles (I)

 **General Principles (I)**

- It is important to report the **reliability of the attributes / attribute scores**
- Reliability in a model with **discrete latent variables differs mathematically** from reliability in more traditional **continuous latent variable models**



The figure shows a normal distribution curve on the left, divided into sections representing standard deviations from -3 to 3. The percentages for each section are: 2.5% (between -3 and -2), 13.5% (between -2 and -1), 34% (between -1 and 0), 34% (between 0 and 1), 13.5% (between 1 and 2), and 2.5% (between 2 and 3). A red not-equal sign (\neq) is placed between the curve and a clipboard on the right. The clipboard has a checklist with three items, each with a smiley face and a checkbox: a sad face, a neutral face, and a happy face.

6.77 General Principles (II)



General Principles (II)

- Templin and Bradshaw (2013) define reliability via **posterior probabilities of attribute mastery** that indicate which mastery status is more likely and the amount of **uncertainty (error)** in that choice

 **If error is low, reliability should be high and vice-versa**

- This index **aggregates the posterior probabilities** by considering **how consistently** respondents would be **classified if they took the test twice**

 **If they are always classified in the same way, reliability is high; if not, reliability is low**

6.78 Ad-hoc Approach (I)



Ad-hoc Approach (I)

- generate** two “pseudo” classifications from a **Bernoulli distribution** using a randomly sampled person’s **posterior probability of attribute mastery**
- replicate** this process **many times** (e.g., 10,000)
- estimate** the **tetrachoric correlation** of the resulting two-way table as an approximation of the **reliability coefficient**

As it is based on **posterior probabilities** of an attribute, it **assumes**:

 - **good model-data fit**  - **strictly parallel tests**

Research on **optimal estimation** of reliability is an **ongoing area of research**

6.79 Ad-hoc Approach (II)



Ad-hoc Approach (II)



Measuring the Reliability of Diagnostic Classification Model Examiner Estimates

Isaiah Templin
The University of Georgia
Lara Bradshaw
The University of Georgia

Abstract: Over the past decade, diagnostic classification models (DCMs) have become an active area of psychometric research. Despite their use, the reliability of examiner estimates in DCM applications has seldom been reported. In this paper, we study reliability measures for the categorical latent variables of DCMs, a related issue that has not been addressed in the literature. We use two different methods to estimate the reliability of examiner estimates. First, we use the test-retest method to estimate the reliability of examiner estimates. Second, we use the test-retest method to estimate the reliability of examiner estimates. We use two different methods to estimate the reliability of examiner estimates. We use two different methods to estimate the reliability of examiner estimates.

Keywords: Diagnostic classification models, Cognitive diagnosis, Reliability, Classification



THE POWER TO KNOW.



Templin & Bradshaw (2013)

Estimation Macros for Mplus

6.80 Bookend: Reliability



Bookend: Reliability



This is the end of this part.

Topic Selection

6.81 Bookmark: Reporting Checklist



6.82 Reporting

DCM REPORTING REQUIREMENTS	
Method	(*)
• Did the author(s) provide the Q-matrix and describe its development?	<input type="checkbox"/>
• Did the author(s) specify the number of items, latent classes and attributes assessed?	<input type="checkbox"/>
• Did the author(s) describe the sample size and justify its appropriateness for their study?	<input type="checkbox"/>
• Did the author(s) estimate alternative DCMs and provide descriptions?	<input type="checkbox"/>
• Was the software package name and version specified?	<input type="checkbox"/>
• Was the estimator specified?	<input type="checkbox"/>
• Did the author(s) provide an overview of the fit indices used to evaluate model fit?	<input type="checkbox"/>
• If the data was derived from a complex survey design, did the author(s) include weights in the model estimation?	<input type="checkbox"/>
• If the author(s) made any changes to the Q-matrix, were these changes explained?	<input type="checkbox"/>
Results	
• For each DCM tested, did the author(s) evaluate the Overall Bivariate Pearson Chi-Square and Overall Log-likelihood Chi-Square (absolute model fit)? (Model Fit Evaluation Checklist, point a)	<input type="checkbox"/>
• For each model tested, did the author(s) compare the number of observed misfitting item pairs and the number of misfitting pairs expected by chance (absolute model fit)? (Model Fit Evaluation Checklist, point b)	<input type="checkbox"/>
• If the author(s) removed any non-significant interactions or misfitting items to improve/refine model fit, were these changes described and justified? (Model Fit Evaluation Checklist, point c)	<input type="checkbox"/>
• If the author(s) compared alternative DCMs, did they identify which DCMs were excluded due to poor absolute fit? (Model Fit Evaluation Checklist, points a-d)	<input type="checkbox"/>
• Did the author(s) report relative model fit statistics (for non-nested & nested models) and loglikelihoodratio tests (for nested models) for those DCMs demonstrating good absolute fit? (Model Fit Evaluation Checklist, point e & f)	<input type="checkbox"/>
• Did the author(s) exclude entropy, which is a poor model fit index for DCMs? (DCM Model Fit Evaluation Checklist, point e)	<input type="checkbox"/>
• Did the author(s) provide sufficient justification for their choice of optimal model? (Model Fit Evaluation Checklist, points a-f)	<input type="checkbox"/>
• Did the author(s) present the proportion of examinees assigned to each class in the best-fitting model? (Interpretation Checklist, point a)	<input type="checkbox"/>
• Did the author(s) provide an overview of the diagnostic quality of the items (i. e., intercepts, main effects, interactions) in the best-fitting model? (Interpretation Checklist, points e-g)	<input type="checkbox"/>
• Did the author(s) report the reliability of the best-fitting model? (Interpretation Checklist, point h)	<input type="checkbox"/>

6.83 Reporting Checklist Benefits



Reporting Checklist: Benefits

- **provides a framework** for conducting and reporting DCM evaluations
- **encourages accurate and thorough reporting** to facilitate replication, transparency and reliability judgements
- **helps to prepare** comprehensive and usable articles
- **helps to assess** the quality of empirical studies
- **facilitates the development** of knowledge among non-expert readers

6.84 Reporting Checklists in other fields

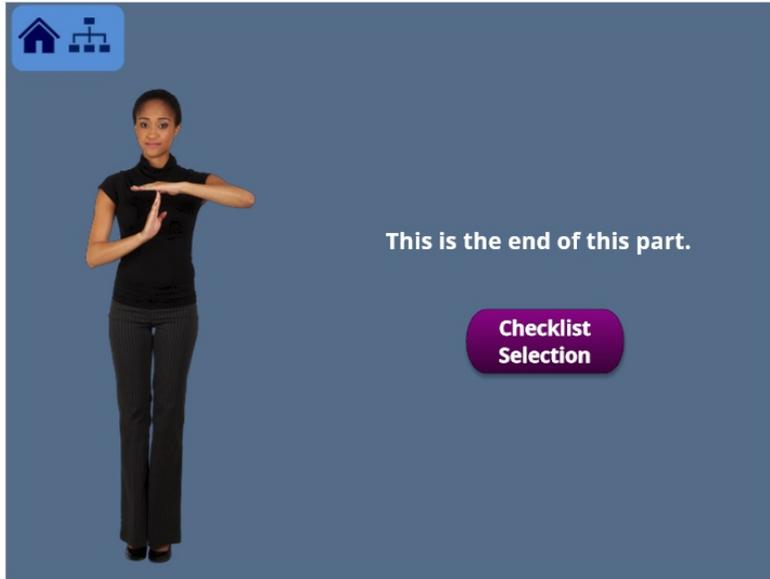


Reporting Checklists: Other Fields

Similar guidelines have been developed elsewhere and have been useful for reporting the results of:

- **Randomized Control Trials**
CONSORT: CONSolidated Standards Of Reporting Trials
- **Case-control, Cohort, and Cross-sectional Studies**
STROBE: STrengthening the Reporting of OBservational Studies in Epidemiology
- **Diagnostic Studies**
QUORUM: QUality Of Reporting of Meta-Analyses
MOOSE: Meta-analysis Of Observational Studies in Epidemiology
STARD: STAndards for Reporting of Diagnostic Accuracy

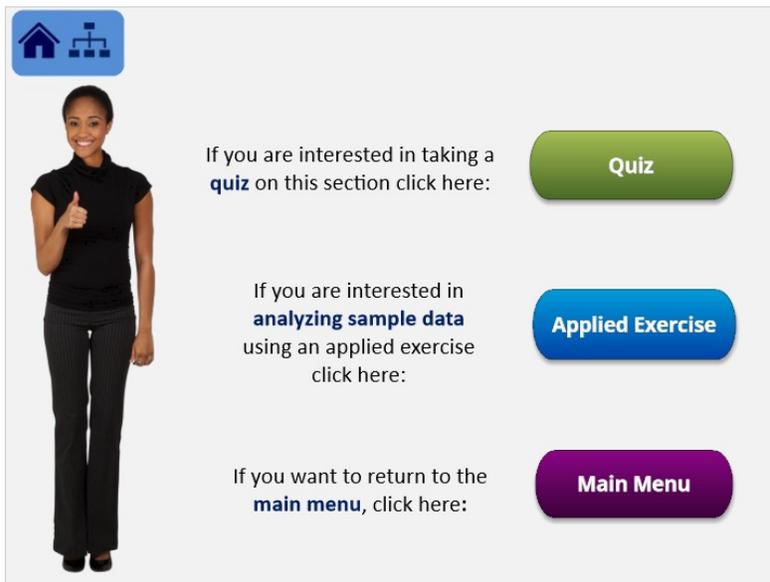
6.85 Bookend: Reporting Checklist



This is the end of this part.

Checklist Selection

6.86 Bookend: Section 5



If you are interested in taking a **quiz** on this section click here:

Quiz

If you are interested in **analyzing sample data** using an applied exercise click here:

Applied Exercise

If you want to return to the **main menu**, click here:

Main Menu

6.87 Module Cover (END)

