

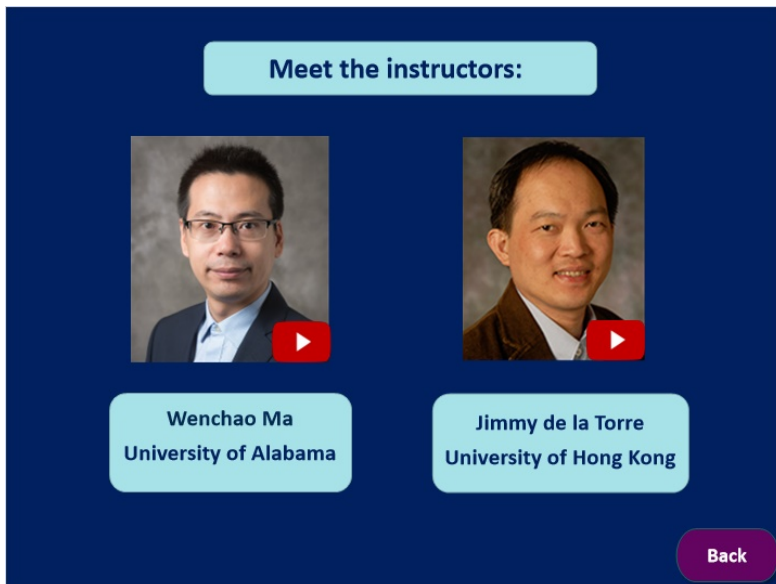
DM05 SLIDES (G-DINA Framework, Version 1.5)

1. Module Overview

1.1 Module Cover (START)




1.2 Instructors




1.3 Designers

Meet the designers:




André A. Rupp
Mindful Measurement



Xi Lu
Florida State University

Back

1.4 Welcome



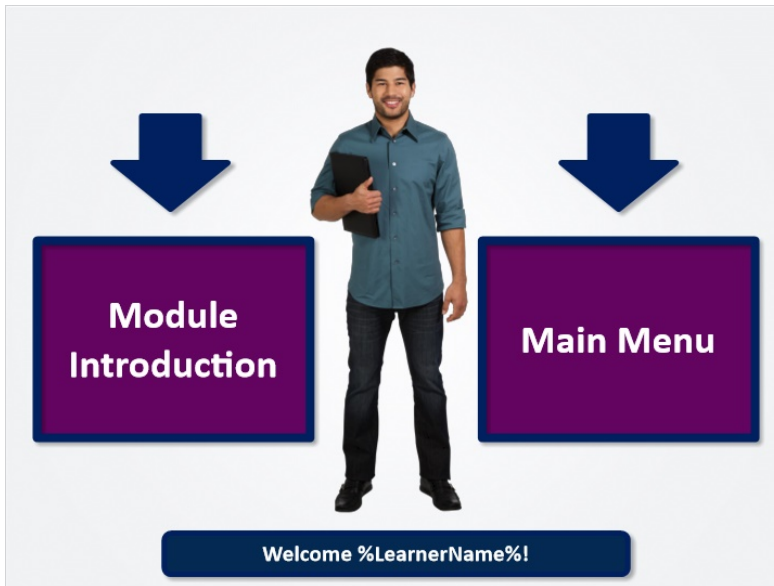
**Welcome to the
ITEMS Module!**

The man to the left is Jet!

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

Type your name here:

1.5 Path Choice



1.6 Overview



1.7 Target Audience

Target Audience

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

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



However, we hope that you find the information in this module useful no matter what your official title, role, or responsibility in an organization is!


1.8 Expectations (I)



Let's discuss expectations....


1.9 Expectations (II)

ITEMS Modules in Context



1.10 Learning Objectives

Learning Objectives




1. Understand the basic principles and ideas behind diagnostic measurement
2. Understand the similarities and differences between IRT and diagnostic models
3. Understand the structure and properties of the G-DINA model framework
4. Specify various diagnostic models as special cases of the G-DINA model
5. Investigate model diagnostics using various statistical procedures
6. Conduct diagnostic analyses using the GDINA R package

1.11 Prerequisites

Prerequisites

To get the most out of this module it is beneficial to have the following background knowledge and basic experiences:

- Basic knowledge of educational assessments
- Basic knowledge of item response theory
- Practical experience with analyzing item response data
- Practical experience with using *R* for data analysis




1.12 Resources


Resources

Ma, W., & de la Torre, J. (2019). Diagnostic measurement: The G-DINA framework (Digital ITEMS Module 05). *Educational Measurement: Issues and Practice*, 38(2), 114-115. Available online at <https://ncme.elevate.commpartners.com/>

Module Citation



Additional References



G-DINA Website

References (Slide Layer)

References (I)

The G-DINA Model Framework

de la Torre, J. (2011). The generalized DINA framework. *Psychometrika*, 76, 179-199.
<https://doi.org/10.1007/s11336-011-9207-7>

Q-matrix Validation Procedures

de la Torre, J., & Chlu, C. Y. (2016). A General Method of Empirical Q-matrix Validation. *Psychometrika*, 81, 253-273. <https://doi.org/10.1007/s11336-015-9467-8>

Ma, W., & de la Torre, J. (2019). An empirical Q-matrix validation method for the sequential generalized DINA model. *British Journal of Mathematical and Statistical Psychology*. <https://doi.org/10.1111/bmsp.12156>

Nájera, P., Sorrel, M. A., & Abad, F. J. (2019). Reconsidering Cutoff Points In the General Method of Empirical Q-Matrix Validation. *Educational and Psychological Measurement*. <https://doi.org/10.1177/0013164418822700>

More
References

Back

More References (Slide Layer)

References (II)

Item-level Model Comparison

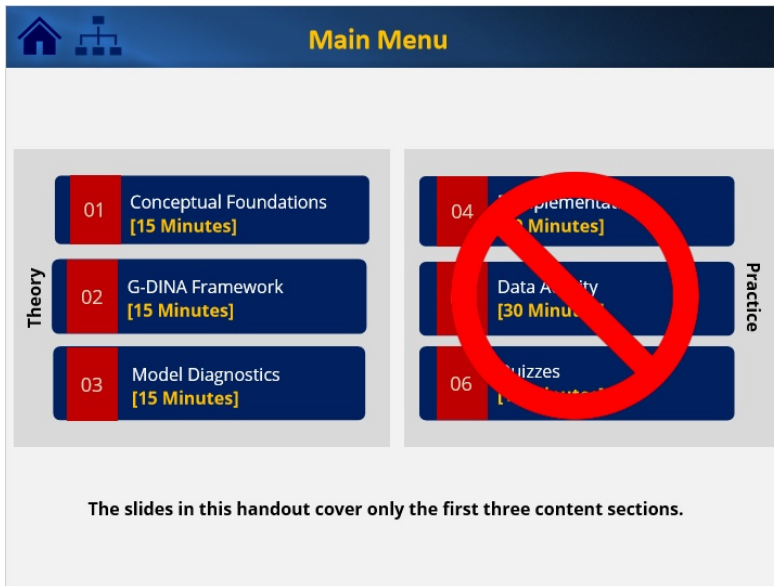
de la Torre, J., & Lee, Y. S. (2013). Evaluating the Wald test for Item-level comparison of saturated and reduced models in cognitive diagnosis. *Journal of Educational Measurement*, 50, 355-373. <https://doi.org/10.1111/jedm.12022>

Ma, W., Iaconangelo, C., & de la Torre, J. (2016). Model Similarity, Model Selection, and Attribute Classification. *Applied Psychological Measurement*, 40, 200-217. <https://doi.org/10.1177/0146621615621717>

Other
References

Back

1.13 Main Menu





2. Section 1: Conceptual Foundations


2.1 Cover: Section 1



2.2 Objectives: Section 1

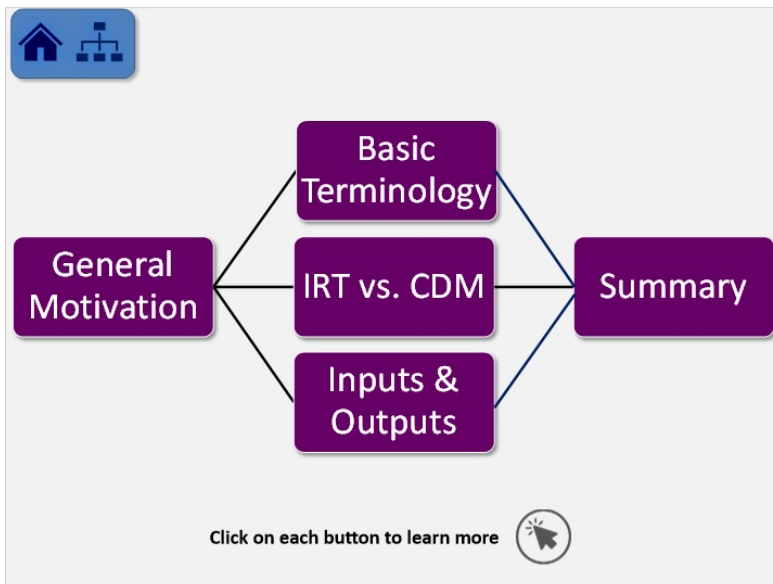


Learning Objectives



1. Understand the major motivations of cognitively diagnostic assessments
2. Understand the basic terminology of cognitive diagnosis models (CDMs)
3. Understand the similarities and differences between IRT and CDMs
4. Understand the inputs and primary outputs of CDM analyses


2.3 Topic Selection



2.4 Bookmark: General Motivation






2.5 Motivation (I)





General Motivation

Educational assessments typically used to support **school and system accountability** do not provide diagnostic information about individual students to support learning because these assessments typically:

- are based on **unidimensional measurement models** with test designs designed to maximize **group-level comparisons** at the institutional (or higher) level
- submerge any distinct set of skills into **a single reported value** for the institution (or level) associated with that dimension, which has a relatively **"coarse-grained"** meaning



2.6 Motivation (II)





Diagnostic Assessment



For assessments to **help inform classroom instruction and learning**, they must be **cognitively diagnostic**.


They must provide information that is **closely tied to classroom instruction**:

- standards- and skills-based
- conceptually multidimensional
- statistically reliable
- didactically actionable



2.7 Bookend: General Motivation





This is the end of this topic.


Topic Selection

2.8 Bookmark: Terminology





2.9 Term Selection

Key Terms		
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 for a more detailed definition 

Topic End

2.10 Attribute



Attribute





Conceptually:
A skill, disposition, or any other construct needed for problem solving


Psychometrically:
A latent variable in a statistical model measured by assessment items

Back

2.11 Diagnosis



Diagnosis





Conceptually:
An act of identifying a disease from its signs and symptoms / identifying skill mastery states for learners

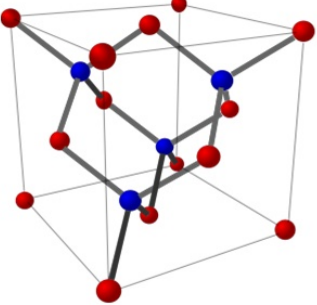
Psychometrically:
A classification of a learner into one of several latent classes

Back

2.12 Dimension



Dimension





Conceptually:
An aspect or facet of a cognitive response process


Psychometrically:
A continuous or categorical variable underlying a statistical model

Back

2.13 Item



Item





Conceptually:
A physical or digital stimulus presented to a learner

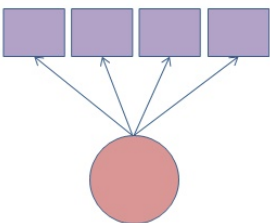
Psychometrically:
A way of recording learners' problem solving performance

Back

2.14 Latent Variable



Latent Variable





Conceptually:
A hypothetical construct of interest


Psychometrically:
An unobserved statistical quantity to be measured (circle on the left)

Back

2.15 Latent Class



Latent Class





Conceptually:
An unobserved grouping of learners that share similar characteristics

Psychometrically:
An unobserved classification state representing a unique mastery profile in CDMs

Back

2.16 Q-matrix



Q-matrix

	+	-	X
5+3	✓		
6x2-5		✓	✓
9+4x2	✓		✓

Conceptually:
An association between items and attributes for an assessment

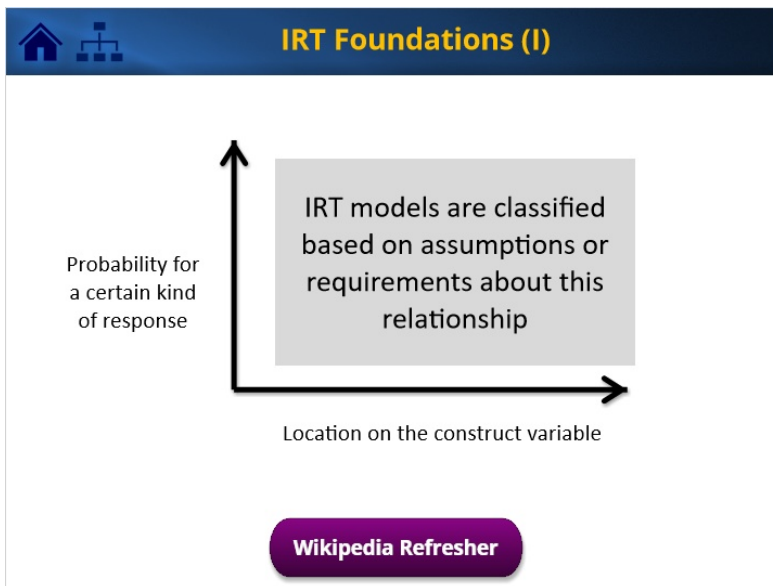
Psychometrically:
A two-dimensional table with numeric entries

Back

2.17 Bookmark: IRT vs CDM



2.18 Item Response Function (IRF)





2.19 IRT Principles

The diagram illustrates the Item Response Theory (IRT) principles with three key points, each in a separate box:

- performance is based on a **single continuous latent trait θ** (*unidimensional IRT*) or **multiple continuous latent traits θ** (**trait vector**) (*multidimensional IRT*)
- learners with **higher latent trait values** have **higher probabilities** of getting an **item correct** (*dichotomous items*) or obtaining a **higher score** (*polytomous items*)
- item parameters** can be used to characterize the **operating characteristics** such as **difficulty**, **discrimination**, or **guessing**

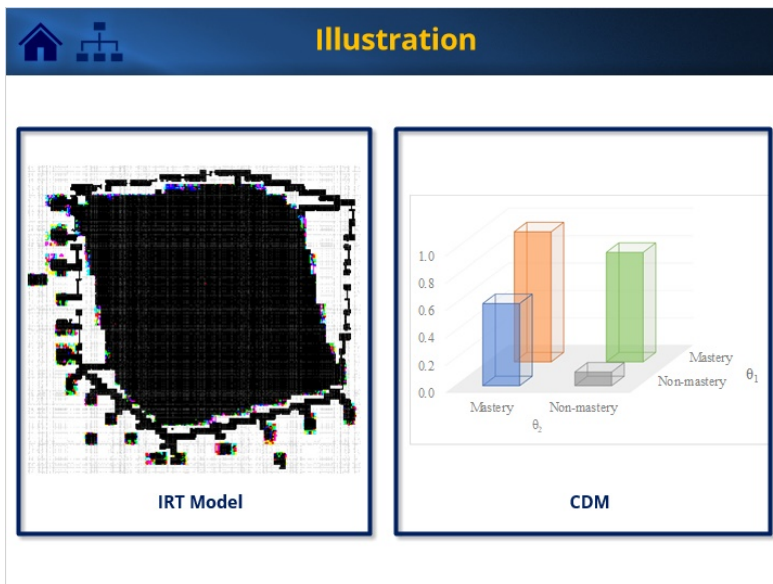
2.20 CDM Principles



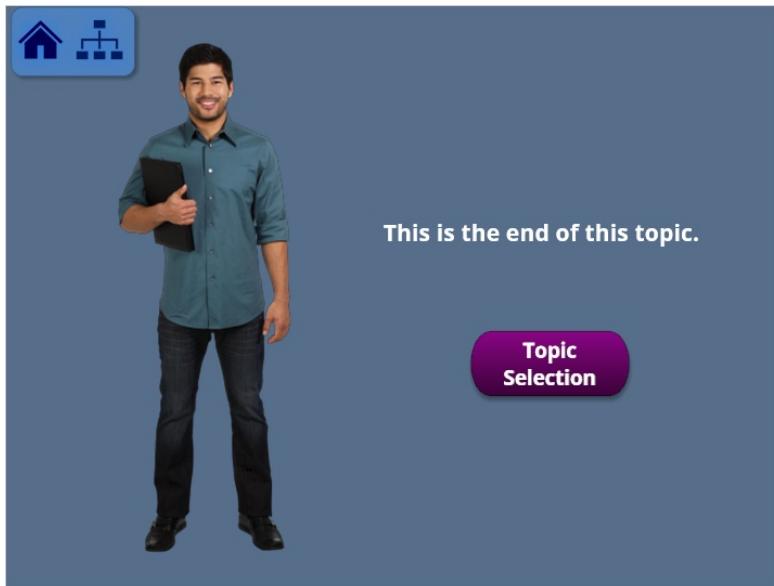
CDM Foundations

- performance is based on **multiple latent variables**, which are also called **attributes** $\alpha = (\alpha_1, \dots, \alpha_K)$
- attributes are typically **binary (0-1)** to indicate **mastery status**, which results in 2^K **latent classes**, each with a **unique attribute profile**
- goal is to **estimate learners' attribute profiles** by classifying them into different **latent classes** based on the **discrete attribute variables**
- each item is designed to **measure one or more** of the latent attributes with the exact design captured in a **Q-matrix** with 0s and 1s
- response probabilities are class-specific** and depends on which attributes are **mastered in the particular latent class**

2.21 Illustration





2.22 Bookend: Framework Comparison

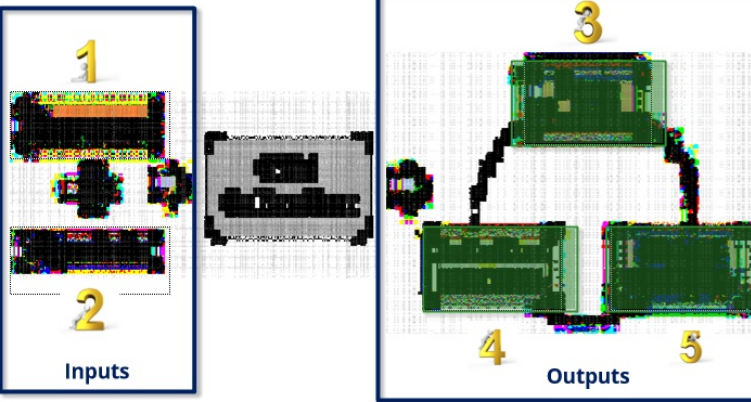



2.23 Bookmark: CDM Outputs



2.24 Component Selection



  **CDM Inputs and Outputs**



Click on any of the numbered boxes to learn more 

Topic End

2.25 Data File Structure

  **Data File Structure**



Columns = Items

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

Rows = Learners

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

2.26 Q-matrix Structure



Q-matrix Structure



Columns = Attributes

	α_1	α_2	α_3	α_4
Item 1	1	0	0	0
Item 2	0	1	0	1
Item 3	1	0	1	0
Item 4	1	1	0	1
Item 5	1	0	0	1
Item 6	0	1	1	0

Rows = Items

Cells = Measurement Structure
(0 – attribute not measured, 1 – attribute measured)

2.27 Item Parameters



Item Parameters

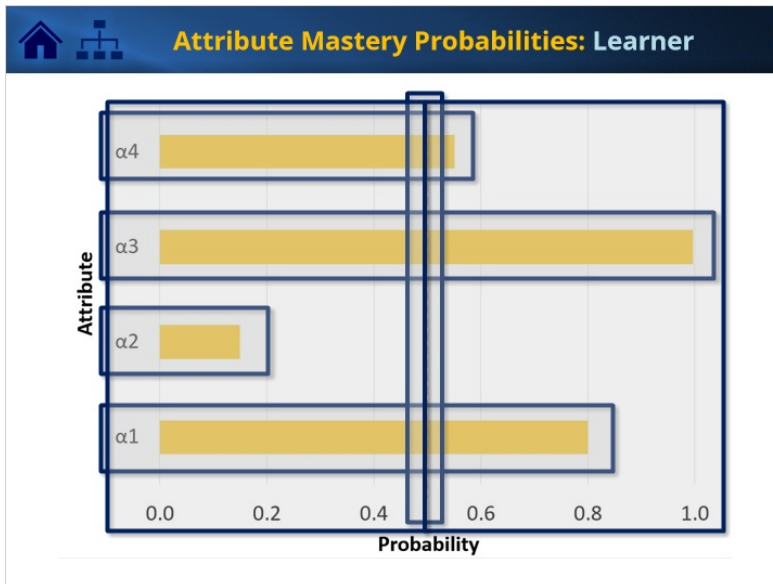
- Different CDMs have **different parameterizations**

Deterministic input noisy “and” gate (DINA) model:

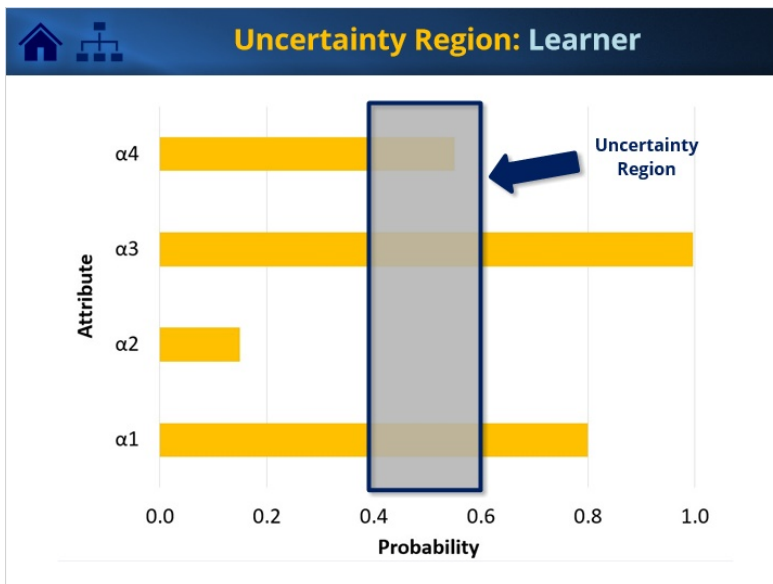
- ✓ guessing parameter
- ✓ slipping parameter

- Guessing** = probability of correct response when at least one required attribute is not mastered
- Slipping** = probability of incorrect response when all required attributes are mastered

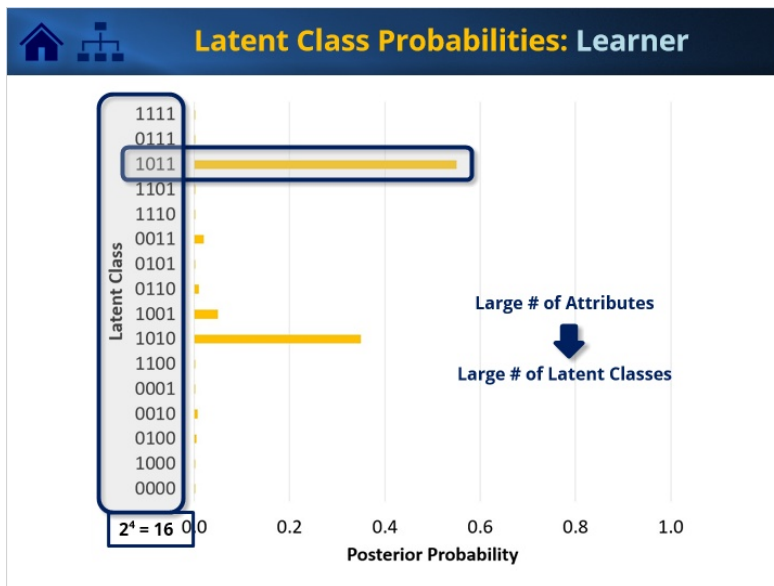
2.28 Attribute Probabilities (Person)



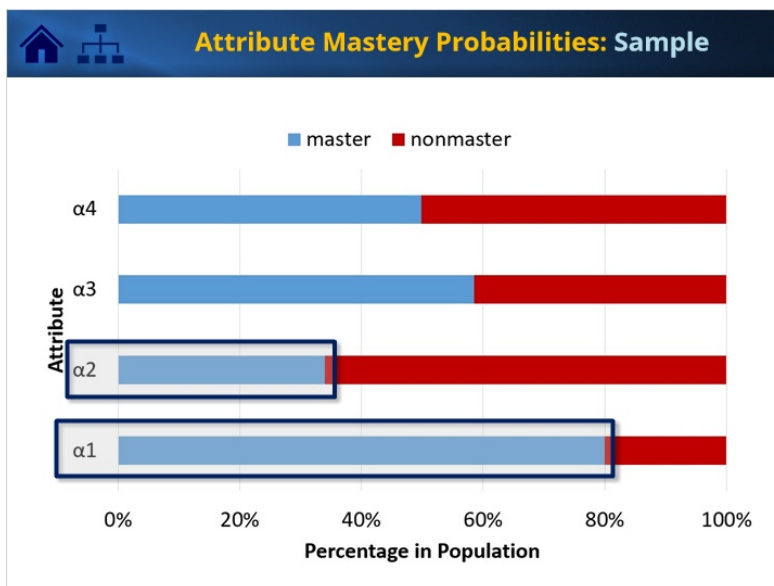
2.29 Attribute Uncertainties (Person)



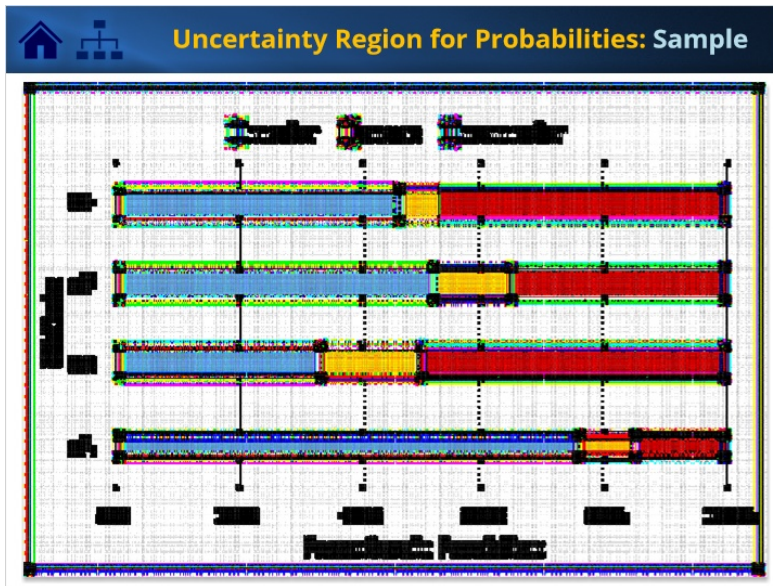
2.30 Latent Class Probabilities (Person)



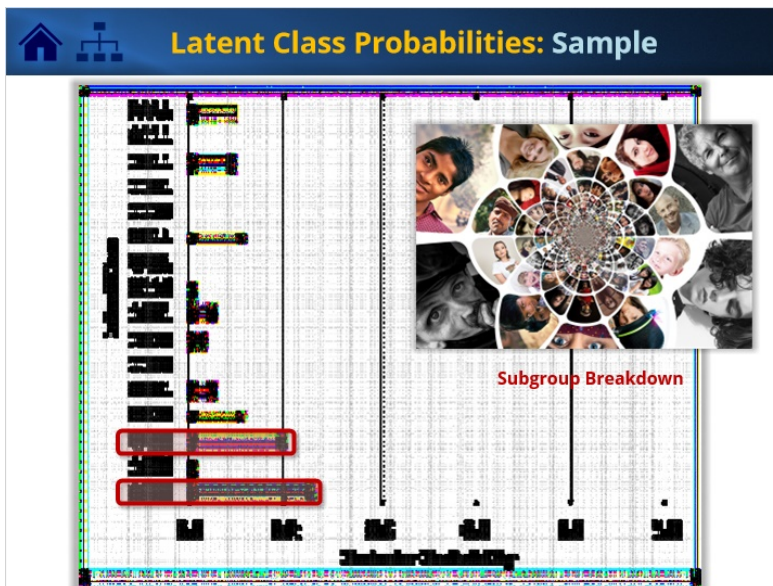
2.31 Attribute Probabilities (Sample)





2.32 Attribute Uncertainties (Sample)



2.33 Latent Class Probabilities (Sample)





2.34 Summary




Summary (I)

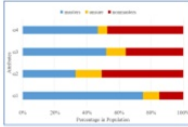
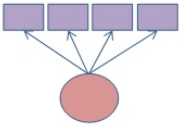
- CDMs aim to **classify learners** into different **latent classes** and provide **diagnostic feedback** to learners and teachers
- CDMs and IRT models are similar in some respects (**multidimensional structure**) but different in others (**discrete vs continuous latent variables**)
- CDMs are often used for **dichotomous responses and attributes** but can also handle **polytomous items and attributes**

2.35 Summary






Summary (II)


- CDMs require that analysts provide **item responses** and **Q-matrix** and choose a **modeling framework** or **specific model**


	+	-	X
5+3	✓		
6x2-5		✓	✓
9+4x2	✓		✓
- CDM analyses will produce **various outputs** including **item parameter estimates** to gauge item quality and **attribute profile estimates** to evaluate **learner characteristics**



2.36 Bookend: Section 1



If you are interested in taking a **self-assessment** on this section click here:



If you are interested in seeing worked data examples of analyses in an **R package** click here:



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

3. Section 2: GDINA Framework

3.1 Cover: Section 2





Section 2:


The G-DINA Framework

[15 Minutes]

3.2 Objectives: Section 2





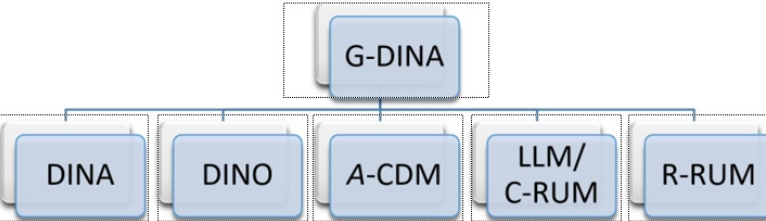
Learning Objectives




1. Provide definitions of key modeling components necessary for working with the G-DINA framework
2. Identify key parameters in the basic structure of the G-DINA model framework and describe how they relate to response probabilities
3. Specify several reduced CDMs defined in the G-DINA framework

3.3 Model Selection

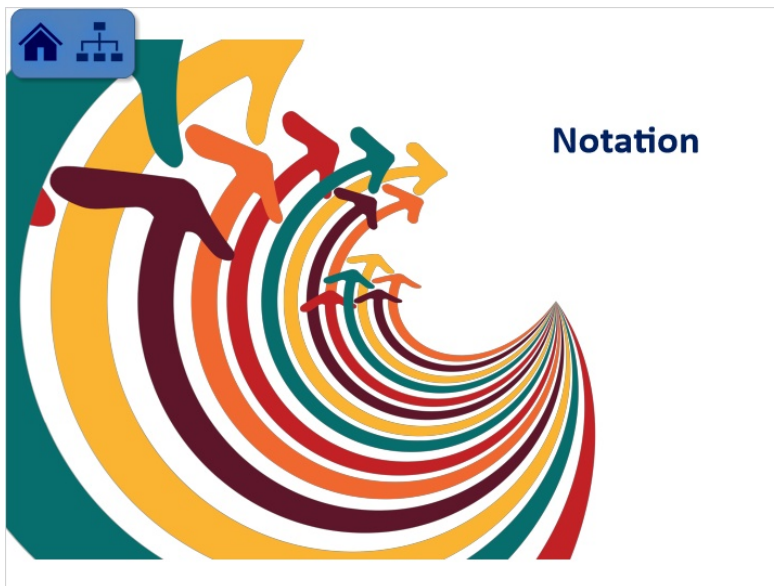






Click on the model buttons to learn more 

[Notation](#) [Section End](#)

3.4 Bookmark: Notation



3.5 Latent Classes



Notation: Latent Classes

K = # of attributes measured by an assessment

If all attributes are binary (0/1) then there are 2^K latent classes



For example, if $K = 4$ there are

$2^4 = 2 \times 2 \times 2 \times 2 = 16$

latent classes

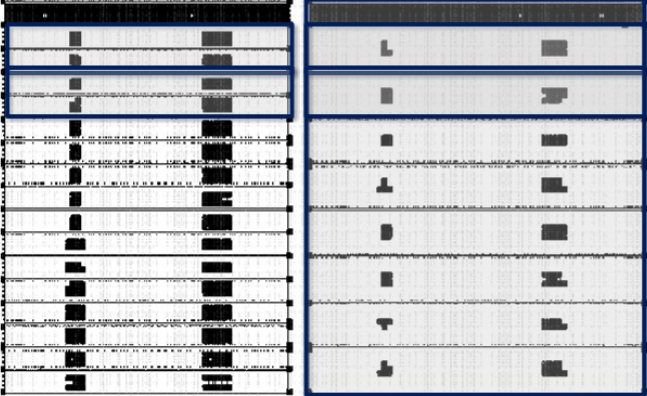
Latent class	Attribute profile α_k
1	0000
2	0001
3	0010
4	0011
5	0100
6	0101
7	0110
8	0111
9	1000
10	1001
11	1010
12	1011
13	1100
14	1101
15	1110
16	1111

3.6 Attribute Profiles





Notation: Reduced Attribute Profiles


Suppose the first K_j^* attributes are measured by item j and denote the profile of these required attributes by α_{lj}^*



3.7 Response Probability



Notation: Response Probability

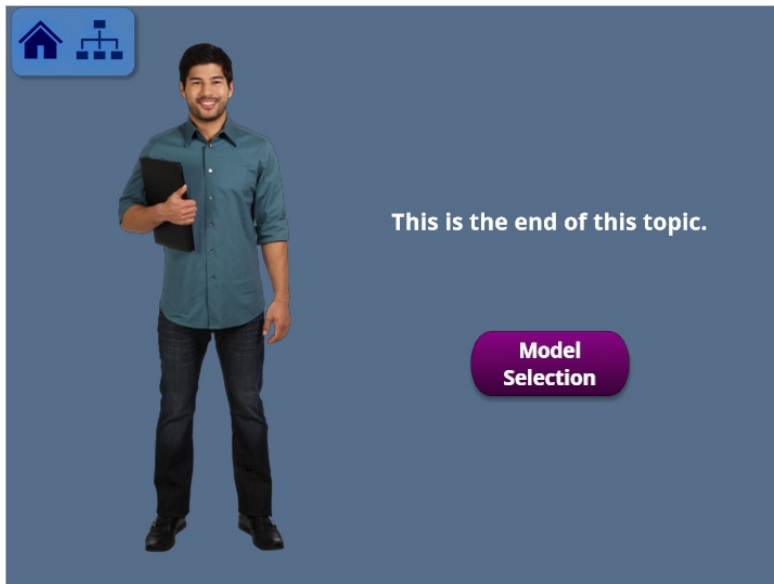


Let Y_j be a **binary response** variable for item j

The **conditional probability** of answering item j correctly given a **reduced attribute profile** is denoted by:

$$P(\alpha_{lj}^*) = P(Y_j = 1 | \alpha_{lj}^*)$$



3.8 Bookend: Notation



3.9 Bookmark: G-DINA





3.10 Framework Properties



The Generalized DINA (G-DINA) Model


- The **generalized deterministic input noisy “and” gate (G-DINA)** model is **one of the most general frameworks** in the literature
- The model subsumes **many well-known CDMs as special cases** via **parameter restrictions** or changes in the **link function**
- The modeling framework is **very flexible** and tends to **fit real data** sets better than many restricted models
- **Item quality** can be measured via item parameters and **learner characteristics** can be measured via **person parameters** as in other measurement models

3.11 Link Function Selection



G-DINA Model: Link Functions

Link Function	Outcome	Special Cases
Identity link	Response Probability	DINA, DINO, A-CDM
Logit link	Log-odds of Response Probability	LLM/C-RUM
Log link	Log of Response Probability	R-RUM

Click on each row to see the model variant equation 

Topic End

3.12 Identity Link (I)

G-DINA Model: Identity Link

For the **identity link**, the **item response probability** is:

$$P(\alpha_{ij}^*) = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk} \alpha_{1k} + \sum_{k'=k+1}^{K_j^*} \sum_{k=1}^{K_j^*-1} \delta_{jkk'} \alpha_{1k} \alpha_{1k'} + \dots + \delta_{j12\dots K_j^*} \prod_{k=1}^{K_j^*} \alpha_{1k}$$

Identity Link
Intercept
Main Effects
Interaction Effects

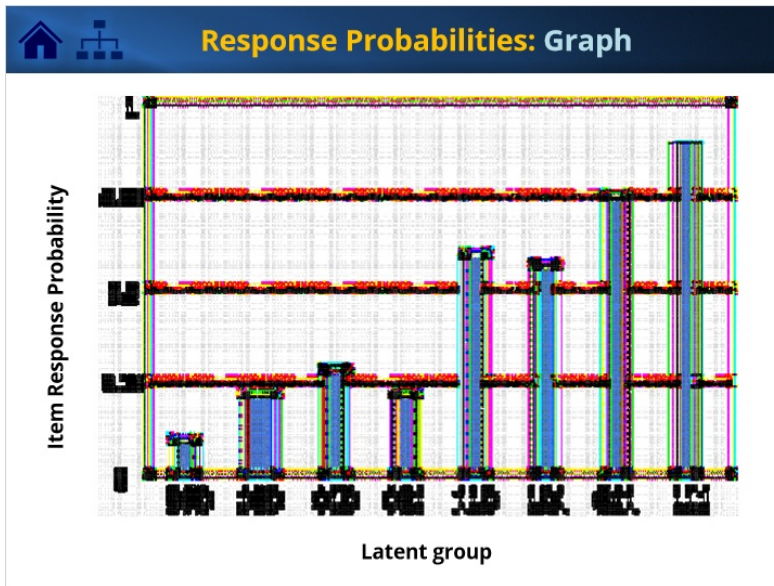
3.13 Identity Link (II)

Response Probabilities: Equations

Item	Equation	Equation
1	δ_{10}	δ_{10}
2	δ_{20}	δ_{20}
3	δ_{30}	δ_{30}
4	δ_{40}	δ_{40}
5	δ_{50}	$\delta_{50} + \delta_{51}\alpha_{11} + \delta_{52}\alpha_{12} + \delta_{53}\alpha_{13} + \delta_{54}\alpha_{14}$
6	δ_{60}	$\delta_{60} + \delta_{61}\alpha_{11} + \delta_{62}\alpha_{12} + \delta_{63}\alpha_{13} + \delta_{64}\alpha_{14} + \delta_{65}\alpha_{11}\alpha_{12} + \delta_{66}\alpha_{11}\alpha_{13} + \delta_{67}\alpha_{11}\alpha_{14} + \delta_{68}\alpha_{12}\alpha_{13} + \delta_{69}\alpha_{12}\alpha_{14} + \delta_{70}\alpha_{13}\alpha_{14}$
7	δ_{70}	$\delta_{70} + \delta_{71}\alpha_{11} + \delta_{72}\alpha_{12} + \delta_{73}\alpha_{13} + \delta_{74}\alpha_{14} + \delta_{75}\alpha_{11}\alpha_{12} + \delta_{76}\alpha_{11}\alpha_{13} + \delta_{77}\alpha_{11}\alpha_{14} + \delta_{78}\alpha_{12}\alpha_{13} + \delta_{79}\alpha_{12}\alpha_{14} + \delta_{80}\alpha_{13}\alpha_{14}$
8	δ_{80}	$\delta_{80} + \delta_{81}\alpha_{11} + \delta_{82}\alpha_{12} + \delta_{83}\alpha_{13} + \delta_{84}\alpha_{14} + \delta_{85}\alpha_{11}\alpha_{12} + \delta_{86}\alpha_{11}\alpha_{13} + \delta_{87}\alpha_{11}\alpha_{14} + \delta_{88}\alpha_{12}\alpha_{13} + \delta_{89}\alpha_{12}\alpha_{14} + \delta_{90}\alpha_{13}\alpha_{14}$

parameters = # latent groups

3.14 Identity Link (III)



3.15 Logit Link

G-DINA Model: Logit Link

For the **logit link**, the **log-odds of the item response probability** is:

$$\text{logit} [P(\alpha_{ij}^*)] = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk} \alpha_{ik} + \sum_{k'=k+1}^{K_j^*} \sum_{k=1}^{K_j^*} \delta_{jkk'} \alpha_{ik} \alpha_{ik'} + \dots + \delta_{j12\dots K_j^*} \prod_{k=1}^{K_j^*} \alpha_{ik}$$

Logit Link **Intercept** **Main Effects** **Interaction Effects**

The following table shows the log-odds of the item response probability for different latent groups (j) and items (i). The table is divided into columns for the intercept, main effects, and interaction effects.

Latent group (j)	Intercept	Main Effects	Interaction Effects
1	0.00	0.00	0.00
2	0.00	0.00	0.00
3	0.00	0.00	0.00
4	0.00	0.00	0.00
5	0.00	0.00	0.00
6	0.00	0.00	0.00
7	0.00	0.00	0.00
8	0.00	0.00	0.00
9	0.00	0.00	0.00
10	0.00	0.00	0.00
11	0.00	0.00	0.00
12	0.00	0.00	0.00
13	0.00	0.00	0.00
14	0.00	0.00	0.00
15	0.00	0.00	0.00
16	0.00	0.00	0.00
17	0.00	0.00	0.00
18	0.00	0.00	0.00
19	0.00	0.00	0.00
20	0.00	0.00	0.00
21	0.00	0.00	0.00
22	0.00	0.00	0.00
23	0.00	0.00	0.00
24	0.00	0.00	0.00
25	0.00	0.00	0.00
26	0.00	0.00	0.00
27	0.00	0.00	0.00
28	0.00	0.00	0.00
29	0.00	0.00	0.00
30	0.00	0.00	0.00
31	0.00	0.00	0.00
32	0.00	0.00	0.00
33	0.00	0.00	0.00
34	0.00	0.00	0.00
35	0.00	0.00	0.00
36	0.00	0.00	0.00
37	0.00	0.00	0.00
38	0.00	0.00	0.00
39	0.00	0.00	0.00
40	0.00	0.00	0.00
41	0.00	0.00	0.00
42	0.00	0.00	0.00
43	0.00	0.00	0.00
44	0.00	0.00	0.00
45	0.00	0.00	0.00
46	0.00	0.00	0.00
47	0.00	0.00	0.00
48	0.00	0.00	0.00
49	0.00	0.00	0.00
50	0.00	0.00	0.00
51	0.00	0.00	0.00
52	0.00	0.00	0.00
53	0.00	0.00	0.00
54	0.00	0.00	0.00
55	0.00	0.00	0.00
56	0.00	0.00	0.00
57	0.00	0.00	0.00
58	0.00	0.00	0.00
59	0.00	0.00	0.00
60	0.00	0.00	0.00
61	0.00	0.00	0.00
62	0.00	0.00	0.00
63	0.00	0.00	0.00
64	0.00	0.00	0.00
65	0.00	0.00	0.00
66	0.00	0.00	0.00
67	0.00	0.00	0.00
68	0.00	0.00	0.00
69	0.00	0.00	0.00
70	0.00	0.00	0.00
71	0.00	0.00	0.00
72	0.00	0.00	0.00
73	0.00	0.00	0.00
74	0.00	0.00	0.00
75	0.00	0.00	0.00
76	0.00	0.00	0.00
77	0.00	0.00	0.00
78	0.00	0.00	0.00
79	0.00	0.00	0.00
80	0.00	0.00	0.00
81	0.00	0.00	0.00
82	0.00	0.00	0.00
83	0.00	0.00	0.00
84	0.00	0.00	0.00
85	0.00	0.00	0.00
86	0.00	0.00	0.00
87	0.00	0.00	0.00
88	0.00	0.00	0.00
89	0.00	0.00	0.00
90	0.00	0.00	0.00
91	0.00	0.00	0.00
92	0.00	0.00	0.00
93	0.00	0.00	0.00
94	0.00	0.00	0.00
95	0.00	0.00	0.00
96	0.00	0.00	0.00
97	0.00	0.00	0.00
98	0.00	0.00	0.00
99	0.00	0.00	0.00
100	0.00	0.00	0.00

3.16 Log Link

G-DINA Model: Log Link

For the **log link**, the **log of the item response probability** is:

$$\log [P(\alpha_{ij}^*)] = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk} \alpha_{lk} + \sum_{k'=k+1}^{K_j^*} \sum_{k=1}^{K_j^* - k'} \delta_{jkk'} \alpha_{lk} \alpha_{lk'} + \dots + \delta_{j12\dots K_j^*} \prod_{k=1}^{K_j^*} \alpha_{lk}$$

Log Link
Intercept
Main Effects
Interaction Effects

3.17 Bookmark: DINA



3.18 DINA Equation

🏠
📊
Diagnostic Inputs Noisy "and" Gate (DINA)

$$P(\alpha_{lj}^*) = \delta_{j0} + \delta_{j12\dots K_j^*} \prod_{k=1}^{K_j^*} \alpha_{lk}$$

$P(\alpha_{lj}^*)$
 Identity Link

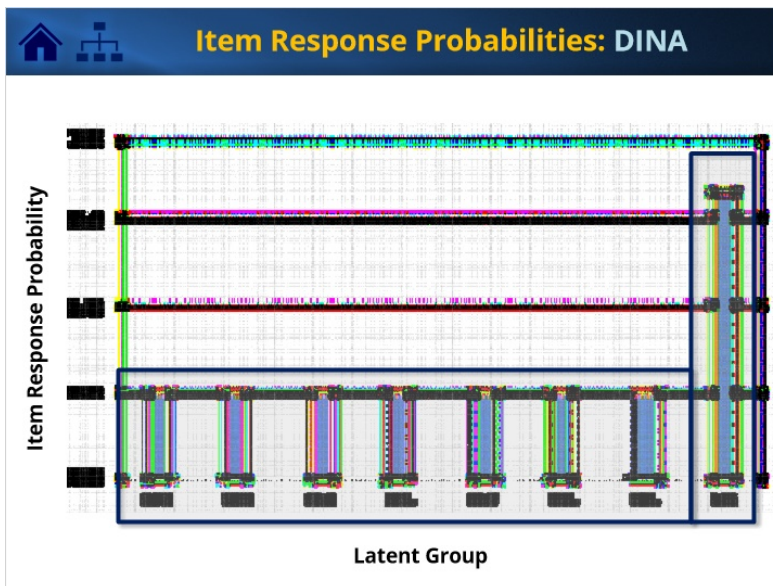
δ_{j0}
 Intercept

$\delta_{j12\dots K_j^*}$
 Interaction Effect

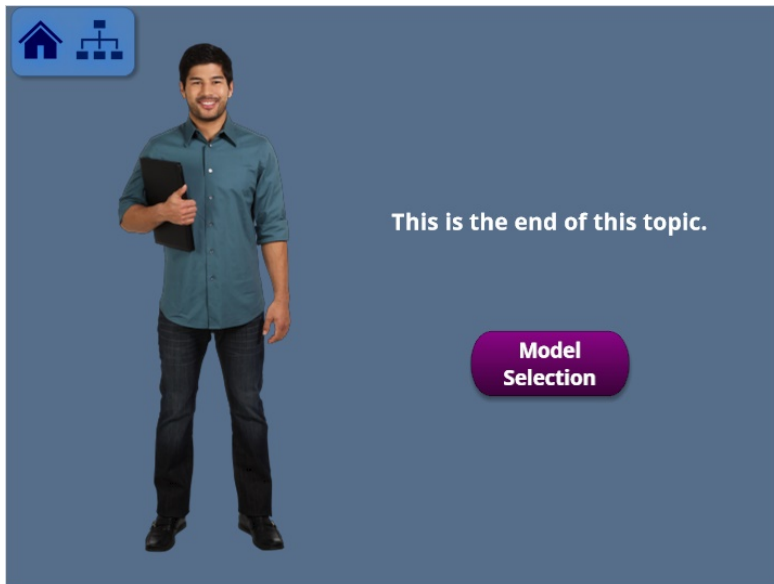
$\prod_{k=1}^{K_j^*} \alpha_{lk}$
 Attribute Product

This is the **identity link** G-DINA model with the **constraint** that **all** but the **intercept** and the **highest-order interaction** are **equal to 0**.

3.19 DINA Probabilities



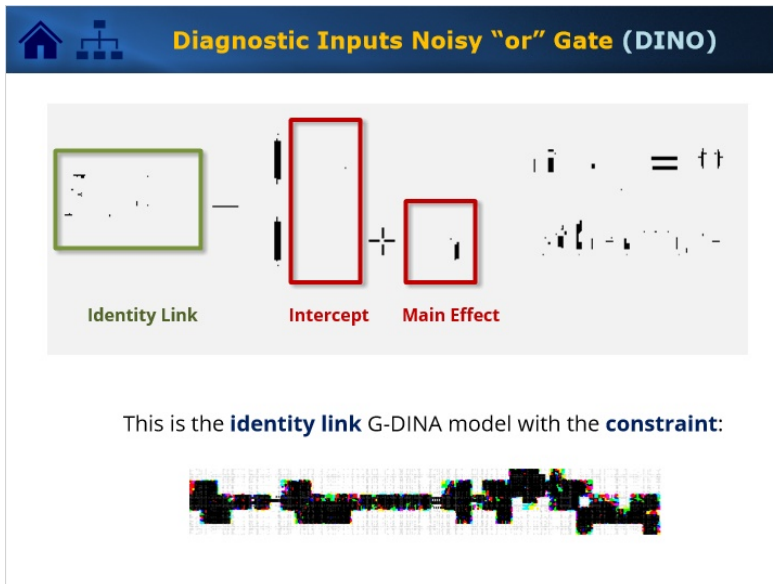
3.20 Bookend: DINA



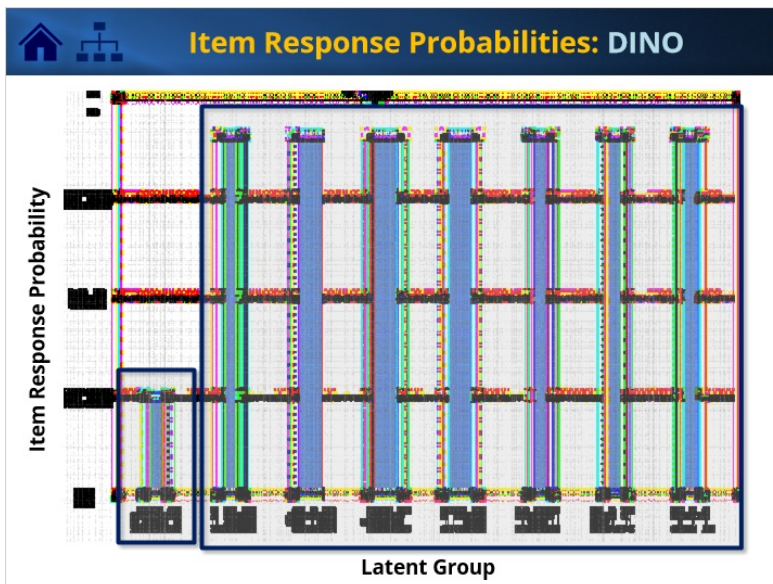
3.21 Bookmark: DINO



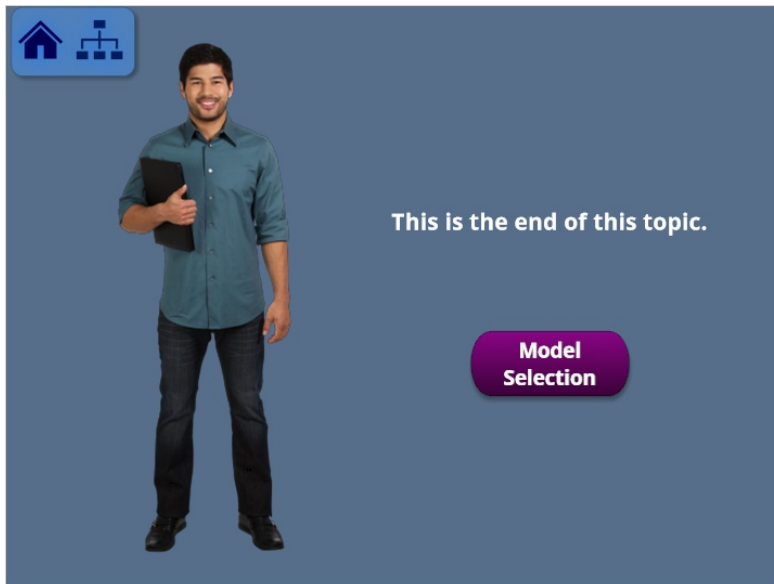
3.22 DINO Equation



3.23 DINO Probabilities





3.24 Bookend: DINO



3.25 Bookmark: A-CDM



3.26 A-CDM Equation

 **Additive Cognitive Diagnosis Model (A-CDM)**

$$P(\alpha_{lj}^*) = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk} \alpha_{lk}$$

Identity Link

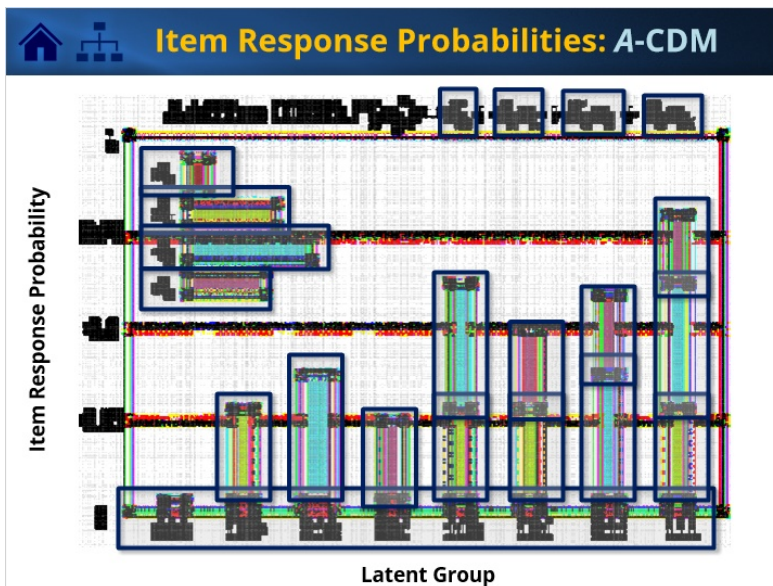
Intercept

Main Effects

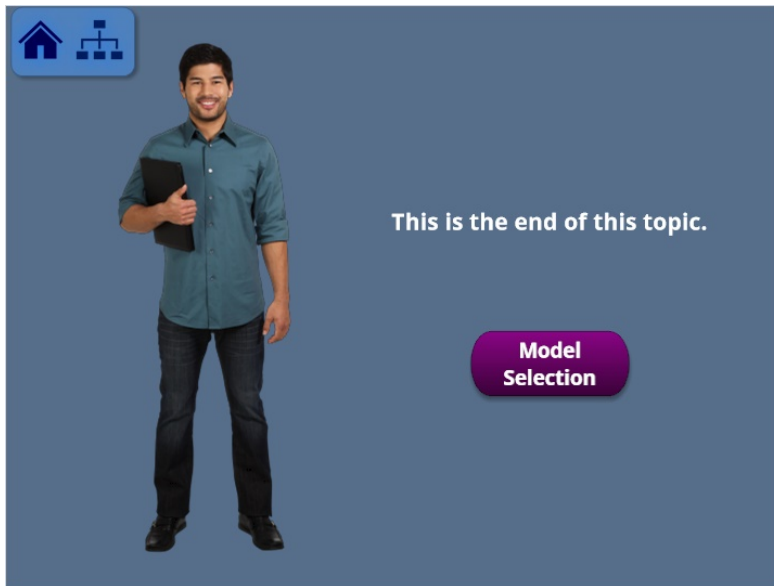
This is the **identity link** G-DINA model **without interaction effects**.

It indicates that **mastering one attribute increases the probability of success** on item j by δ_{jk} **independent of the contributions of the other attributes**.

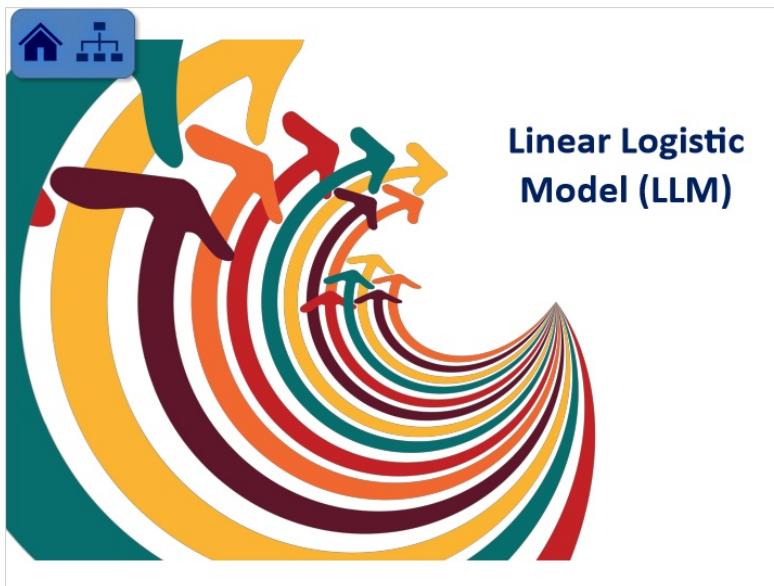
3.27 A-CDM Probabilities





3.28 Bookend: ACDM



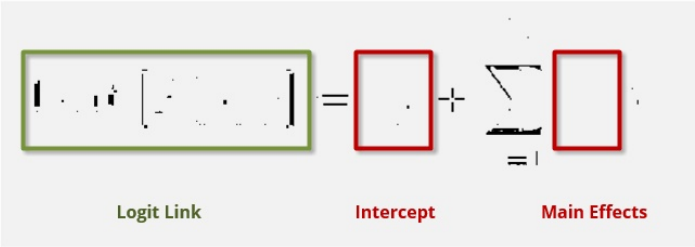
3.29 Bookmark: LLM



3.30 LLM Equation



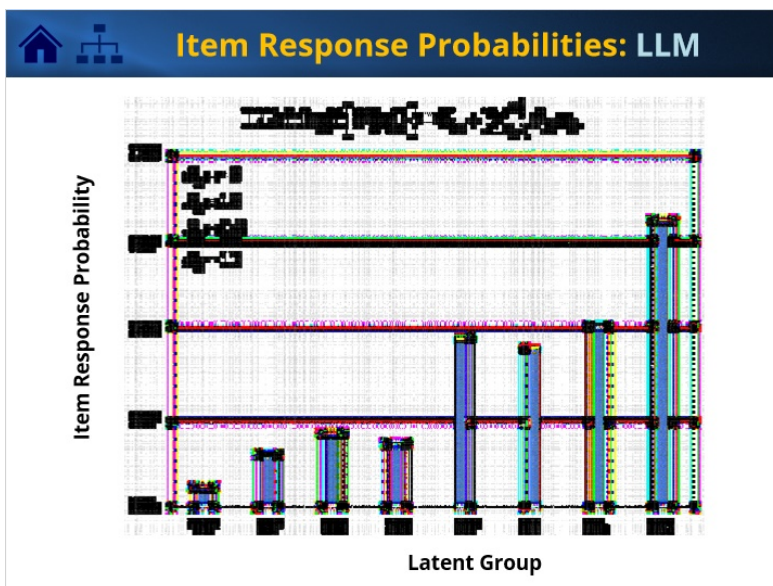
Linear Logistic Model (LLM)



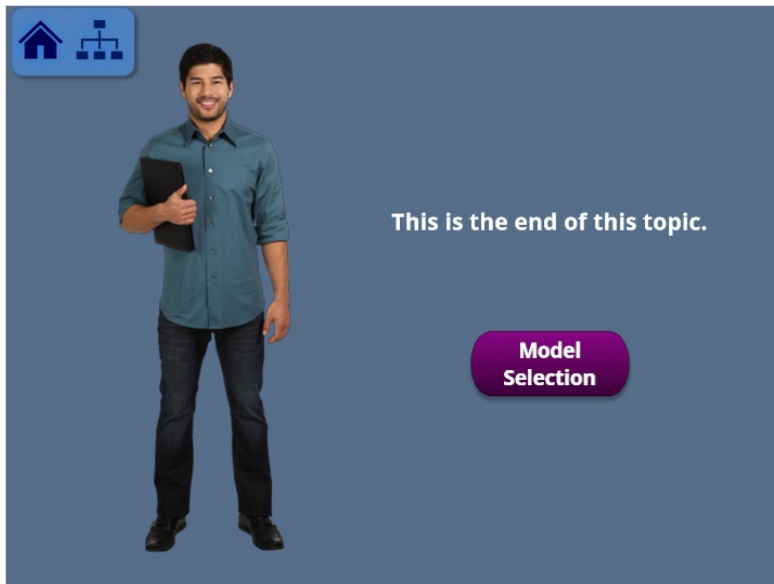
Logit Link Intercept Main Effects

This is the **logit link** G-DINA model with **intercept** and **main effects** only. The LLM is also called **compensatory reparameterized unified model (C-RUM)**.

3.31 LLM Probabilities



3.32 Bookend: LLM



3.33 Bookmark: R-RUM



3.34 R-RUM Equation

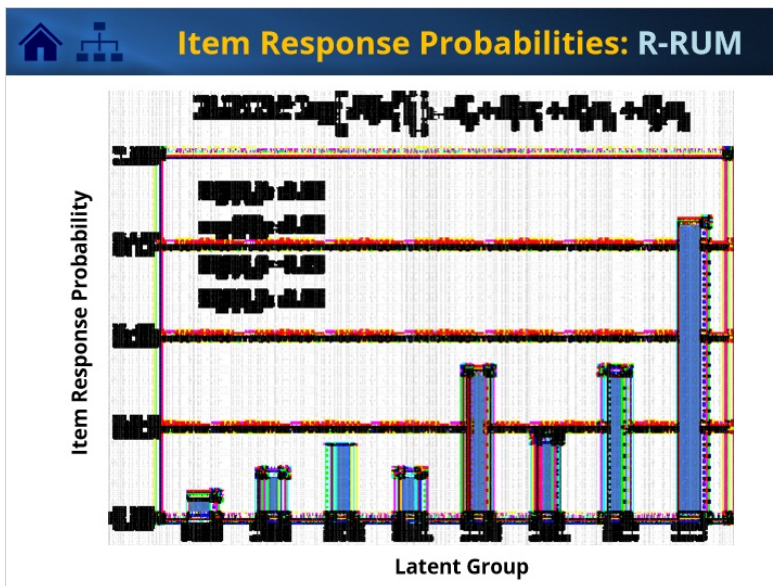
🏠
📊
Reduced Reparam. Unified Model (R-RUM)

$$\ln(\cdot) = \mu_j + \sum_{k=1}^K \beta_k x_k$$

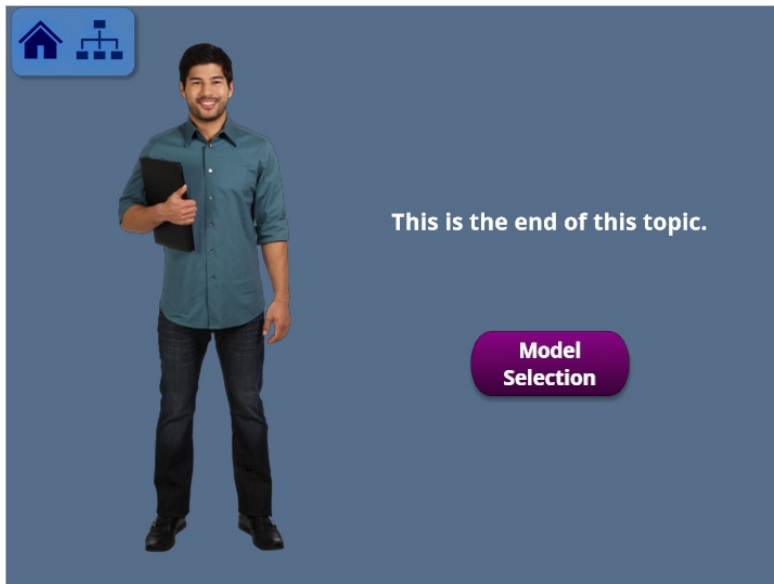
Log Link
Intercept
Main Effects

This is a **log link** G-DINA model with **intercept** and **main effects** only.



3.35 R-RUM Probabilities



3.36 Bookend: R-RUM



3.37 Summary






Summary


The G-DINA model...

- defines the **conditional probability of item responses**
- is very **general** because it considers main effects and all interactions
- can serve as a **framework** where reduced models can be obtained by setting appropriate constraints
- item parameter estimates can be used to assess the **quality of item** and **validate the Q-matrix**
- **attribute profiles** can be estimated after item parameters are estimated


3.38 Bookend: Section 2



If you are interested in taking a **self-assessment** on this section click here:



If you are interested in seeing worked data examples of analyses in an **R package** click here:



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

4. Section 3: Model Diagnostics

4.1 Cover: Section 3





Section 3:


Model Diagnostics

[15 Minutes]

4.2 Objectives: Section 3





Learning Objectives



1. Articulate the differences between procedures for assessing model-data fit
2. Describe how the Wald test can be used to compare models at the item level
3. Describe how a general procedure for validating Q-matrix works

4.3 Introduction (I)




Introduction


The G-DINA modeling framework allows analysts perform the usual psychometric evaluations of assessment quality focusing on:

- ✓ **operating characteristics for items** (discrimination, guessing)
- ✓ **classification accuracy for learners** (attribute reliability)
- ✓ **model-data fit** (absolute and relative, item-level and test-level)
- ✓ **item-by-attribute alignment** (Q-matrix modification)
- ✓ **item-level model simplification** (through parameter reduction)
- ✓ **differential item functioning** (based on multiple-group models)

4.4 Model Fit Selection



Relative Fit Item-level Fit Q-matrix Fit

Click on each button to learn more 

Section End

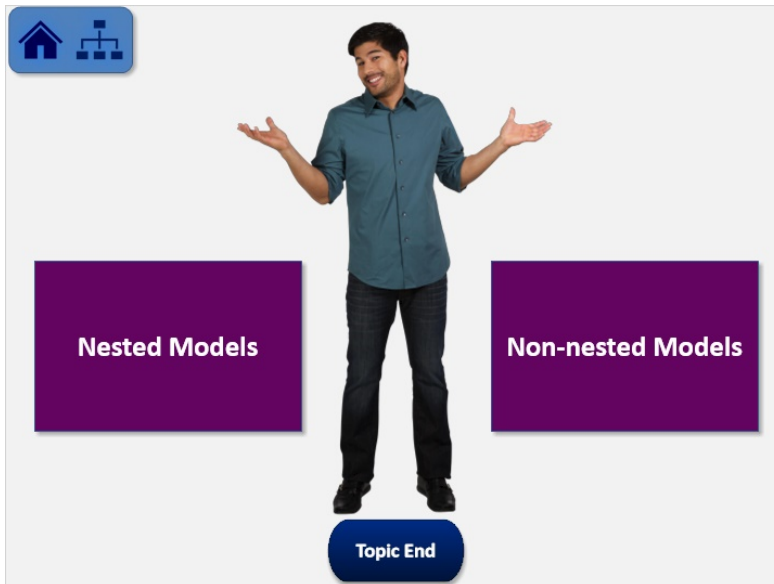
4.5 Bookmark: Relative Fit




Relative Fit



4.6 Model Selection





4.7 LR Test (I)



The Likelihood-ratio Test (I)

- Can be used for **nested models** where one model can be obtained as a **special case** of the other through **eliminating model parameters**
- The **null hypothesis** of the test is that the reduced model fits data as well as the saturated model while the **alternative hypothesis** says that the two models are not equivalent (i.e., that the simplification is not justified)
- The test compares the **log-likelihoods of the two models**, which separately capture how likely it is that each of the two models, as specified, **could have given rise to the observed data**

4.8 LR Test (II)



The Likelihood-ratio Test (II)



- Let S be a **more complex model** and R be a **reduced model**. Then the **likelihood ratio (LR) statistic** for comparing R and S is computed via the **log-likelihood (LL)**:


$$LR = [-2LL^{(R)}] - [-2LL^{(S)}]$$

- The LR statistic is **χ^2 -distributed** with **degrees of freedom (df)** equal to

$$df = p^{(S)} - p^{(R)}$$

4.9 Bookend: Nested Models







This is the end of this topic.

Model Type Selection

4.10 Information Indices



Information Indices

Uses the observed marginalized **maximum likelihood (ML)** value and adjusts it via a **penalty term** that includes the **number of model parameters** and **sample size**:



Deviance (-2LL) = -2 x log-likelihood

Akaike information criterion (AIC) = $-2LL + 2P$

Bayesian information criterion (BIC) = $-2LL + P \times \ln(N)$

where **P** is the total number of **model parameters** and **N** is the **total sample size**

4.11 Information Indices Video



Number of Parameters

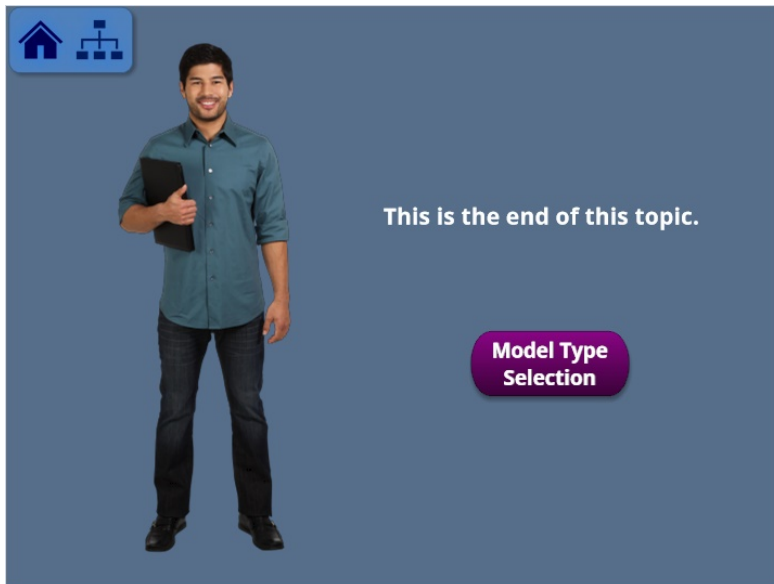
Q-matrix

	α_1	α_2	α_3
Item 1	1	0	0
Item 2	0	1	0
Item 3	1	0	1
Item 4	1	1	1

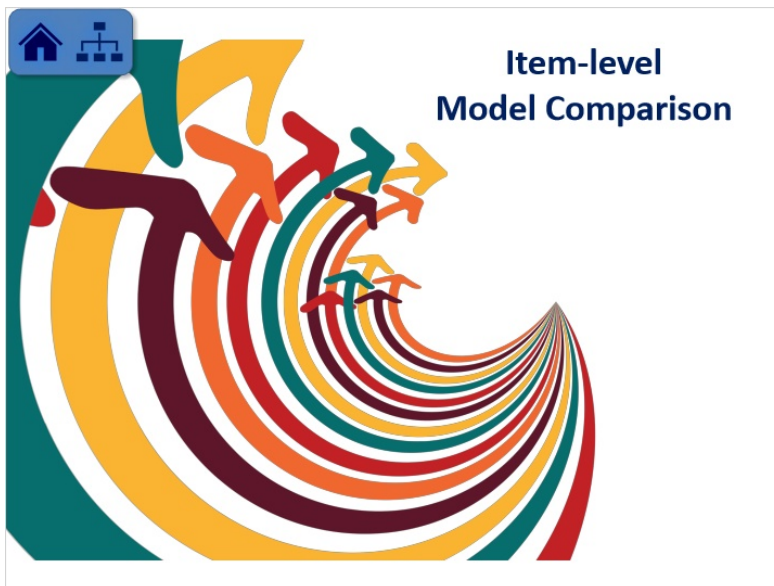
Parameter Table

Model for all items	# of item parameters	# of joint attribute distribution parameters	Total # of parameters
DINA	$2 + 2 + 2 + 2 = 8$	$2^3 - 1 = 7$	15
ACDM	$2 + 2 + 3 + 4 = 11$	$2^3 - 1 = 7$	18
G-DINA	$2 + 2 + 4 + 8 = 16$	$2^3 - 1 = 7$	23



4.12 Bookend: Non-nested Models



4.13 Bookmark: Item-level Fit




4.14 General Principles





General Principles

The **Wald test** can be used to compare the **saturated G-DINA model** and **reduced models** when items measure **more than one attribute**



H_0 : The reduced model fits data as well as the saturated G-DINA model
 H_1 : The reduced model fits data worse than the saturated G-DINA model

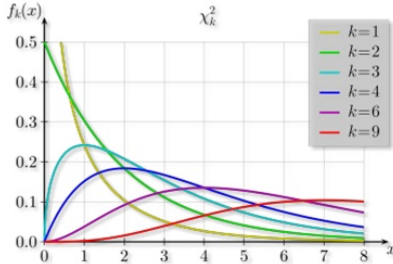
4.15 General Principles



General Principles



Wald test statistic is distributed as χ^2 with degrees of freedom:

parameters of G-DINA model – # parameters of reduced model



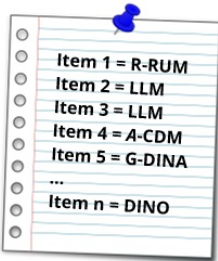

By Geek3 - Own work, CC BY 3.0, <https://commons.wikimedia.org/w/index.php?curid=9884213>

4.16 General Principles



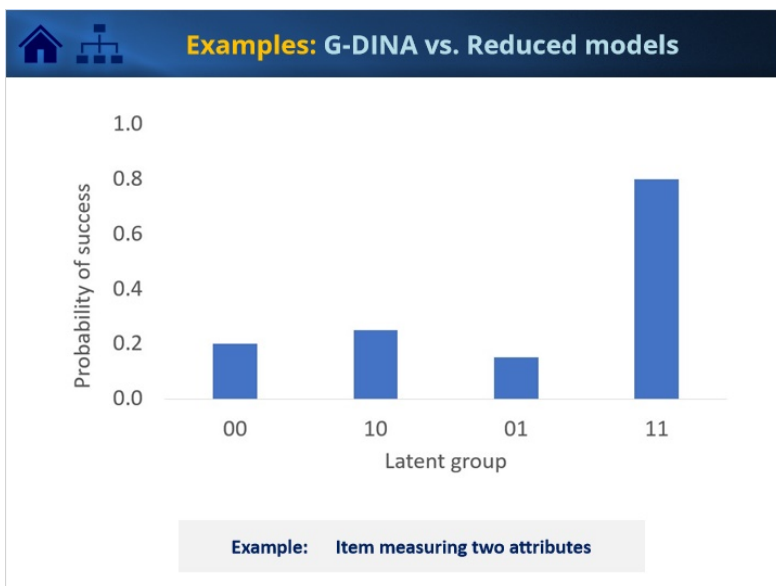
General Principles

Multiple CDMs can be used **simultaneously across items** without prescribing a **one-size-fits-all solution**

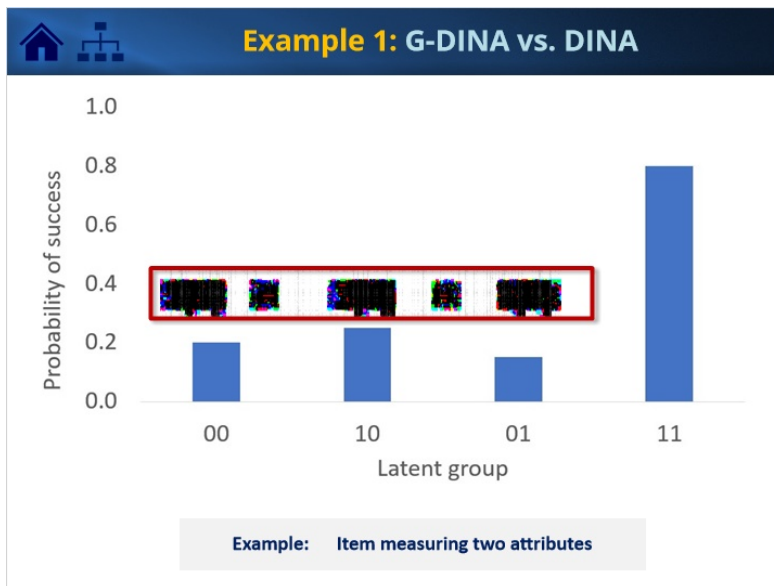


The models selected by the Wald test tend to produce **better attribute profile estimation** than the saturated G-DINA model

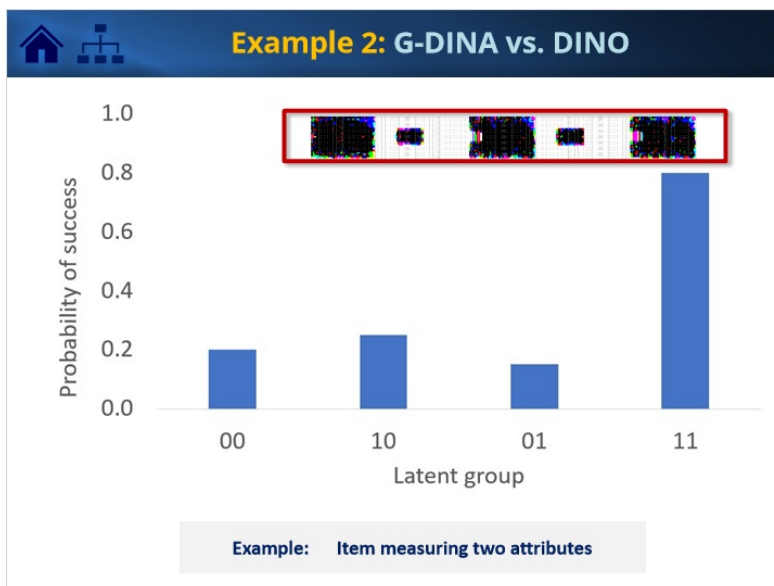
4.17 GDINA vs. DINA



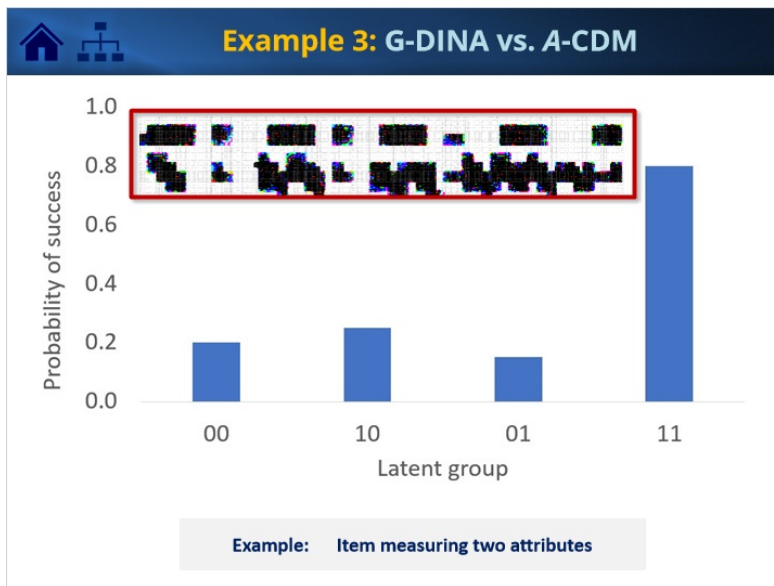
4.18 GDINA vs. DINA



4.19 GDINA vs. DINO



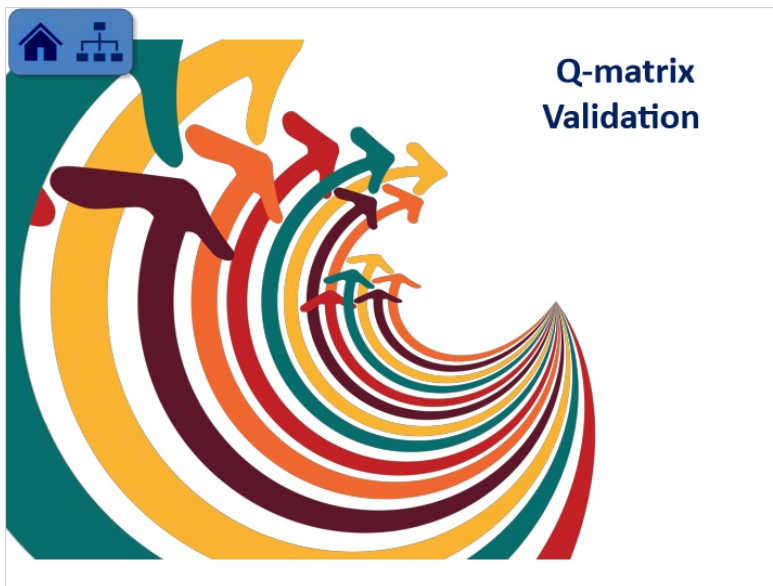
4.20 GDINA vs ACDM





4.21 Bookend: Item-level Fit



4.22 Bookmark: Q-matrix Fit




4.23 General Principles



General Principles

The **saturated G-DINA model** may overfit the data because it involves **many parameters** when items measure **multiple attributes**



Possibilities



The diagram shows a central vertical line with an upward-pointing arrow at the top and a downward-pointing arrow at the bottom. From this central line, several horizontal arrows branch out to the left and right. Each of these horizontal arrows is labeled with the word "Possibility".

Determining the **form of reduced models** may not be easy as different models make **different assumptions** about how **learners use attributes** to respond to items

4.24 General Principles





General Principles

- A **Q-matrix specification** tends to be **partially subjective** in nature
- Misspecifications** in the Q-matrix affect the **attribute classifications**
- The **following method** was designed specifically for the **G-DINA model**

Q-matrix		α_1	α_2	α_3	α_4
	Item 1	1	0	0	0
	Item 2	0	1	0	1
	Item 3	1	0	1	0
	Item 4	1	1	0	1

4.25 Overview of Procedure (I)



Overview of Procedure (I)

For each item, calculate the G-DINA discrimination index (GDI) for each possible q -vector (based on G-DINA estimates)

Properties of the GDI:

- the **true q -vector** and **over-specified q -vectors** have **identical and largest GDI** in principle
- over-specified q -vectors** have **slightly larger GDI** than the **true q -vector**
- the **GDI increases** when there are **more '1's** in the q -vectors; the q -vector with **all '1's** has the **largest GDI**

Reference



Reference (Slide Layer)

Reference



Back

4.26 Overview of Procedure (II)



Overview of Procedure (II)

Use the **Proportion of Variance-accounted-for (PVAF)** by each q -vector for decision-making:

$$\hat{\mathbf{P}} = \frac{\mathbf{Y} \mathbf{Y}^T}{\mathbf{Y} \mathbf{Y}^T + \mathbf{I}_n} = \frac{\mathbf{Y} \mathbf{Y}^T}{\mathbf{Y} \mathbf{Y}^T + \mathbf{I}_n}$$

Decision Rule:

Out of **all q -vectors** with **PVAF > 0.95**, the one having **fewest '1's** is the **suggested q -vector**

4.27 Illustration (I)

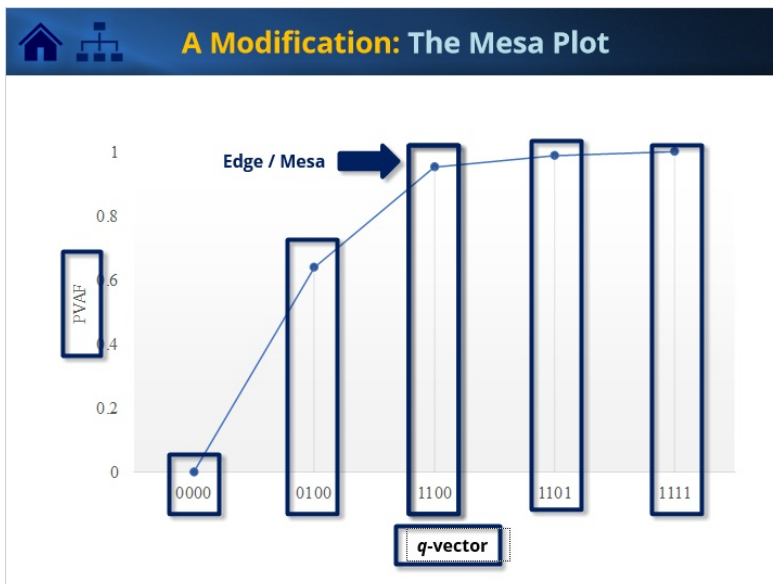
An Illustration: Vector Selection

All possible q-vectors	GDI	PVAF
1000	0.032 / 0.033 = 0.386	
0100	0.053	0.639
0010	0.018	0.217
0001	0.021	0.253
1100	0.079	0.952
1010	0.035	0.422
1001	0.034	0.410
0110	0.061	0.735
0101	0.056	0.675
0011	0.02	0.241
1110	0.081	0.976
1101	0.082	0.988
1011	0.046	0.554
0111	0.063	0.759
1111	0.083	1.000



suggested q-vector

Appropriate q-vectors
If .95 is used as cutoff

4.28 Illustration (II)



4.29 Other Options: Wald Test and Predicted Cutoffs



Other Options: Wald Test and Predicted Cutoffs

British Journal of Mathematical and Statistical Psychology (2019)
© 2019 The British Psychological Society
www.bjps.oxfordjournals.org

An empirical Q-matrix validation method for the sequential generalized DINA model

Wenchao Ma¹ and Jimmy de la Torre²

¹Department of Educational Studies in Psychology, Research Methodology and Counseling, University of Alabama, Tuscaloosa, Alabama, USA
²Faculty of Education, University of Hong Kong, Hong Kong

As a core component of most cognitive diagnosis models, the Q-matrix, or item and attribute association matrix, is typically developed by domain experts, and tends to be subjective. It is critical to validate the Q-matrix empirically because a misspecified Q-matrix could result in erroneous attribute estimation. Most existing Q-matrix validation procedures are developed for dichotomous responses. However, in this paper, we propose a method to empirically detect and correct the misspecifications in the Q-matrix for graded response data based on the sequential generalized deterministic inputs, noisy 'and' gate (G-DINA) model. The proposed Q-matrix validation procedure is implemented in a stepwise manner based on the Wald test and an effect size measure. The feasibility of the proposed method is examined using simulation studies. Also, a set of data from the Trends in International Mathematics and Science Study (TIMSS) 2011 mathematics assessment is analyzed for illustration.


Educational and Psychological Measurement
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journalsPermissions
DOI: 10.1177/0013164418822700
jpe.sagepub.com/home/jpe

Reconsidering Cutoff Points in the General Method of Empirical Q-Matrix Validation


Pablo Nájera¹, Miguel A. Sorrel¹, and Francisco José Abad¹


Abstract
Cognitive diagnosis models (CDMs) are latent class multidimensional statistical models that help classify people accurately by using a set of discrete latent variables, commonly referred to as attributes. These models require a Q-matrix that indicates the attributes involved in each item. A potential problem is that the Q-matrix construction process, typically performed by domain experts, is subjective in nature. This might lead to the existence of Q-matrix misspecifications that can lead to inaccurate classifications. For this reason, several empirical Q-matrix validation methods have been developed in the recent years. de la Torre and Chiu proposed one of the most popular methods, based on a discrimination index. However, some questions related to the usefulness of the method with empirical data remained open due to the restricted number of conditions examined, and the use of a unique cutoff point (EPS) regardless of the data conditions. This article includes two simulation studies to test this validation method under a wider range of conditions, with the purpose of providing it with a higher generalization, and to empirically determine the most suitable EPS considering the data conditions. Results show a good overall performance of the method, the relevance of the different studied factors, and that using a single indiscriminate EPS is not acceptable. Specific guidelines for selecting an appropriate EPS are provided in the discussion.

Click on the images to go to the publishers' web portals



4.30 Bookend: Q-matrix Fit







This is the end of this topic.

Model Fit Selection

4.31 Summary






Summary

There are several analytic methods currently available within the G-DINA model framework for model diagnostics:

- **PVAF and Mesa Plots** can be used to validate the Q-matrix without assuming a specific form of the item response function
- **Wald tests** can be used to assess whether a reduced model can be used in place of the G-DINA model for each item
- **LR tests** can be used to compare two nested models at the test level such as the saturated G-DINA model and the reduced models
- **Deviance, AIC and BIC** can be used to compare various non-nested models at the test level such as different reduced models

4.32 Bookend: Section 3





If you are interested in taking a **self-assessment** on this section click here:

If you are interested in seeing worked data examples of analyses in an **R package** click here:

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

Quiz

Data Examples

Main Menu

4.33 Module Cover (END)

