

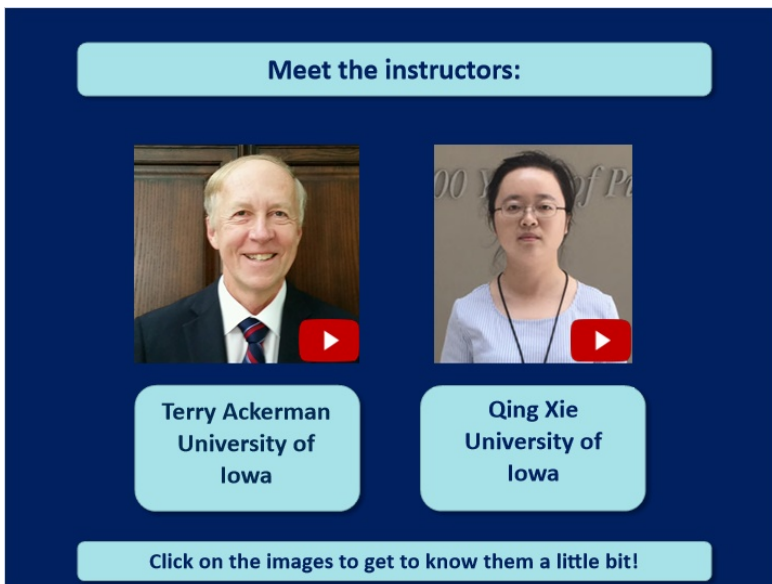
ITEMS: MIRT Graphics

1. Module Overview

1.1 Module Cover




1.2 Instructors



1.3 Designers

Meet the instructional designers:




Jonathan Lehrfeld
ETS

Xi Lu
Florida State
University

André A. Rupp
Mindful
Measurement

Click on the images to get to know them a little bit!

1.4 Welcome



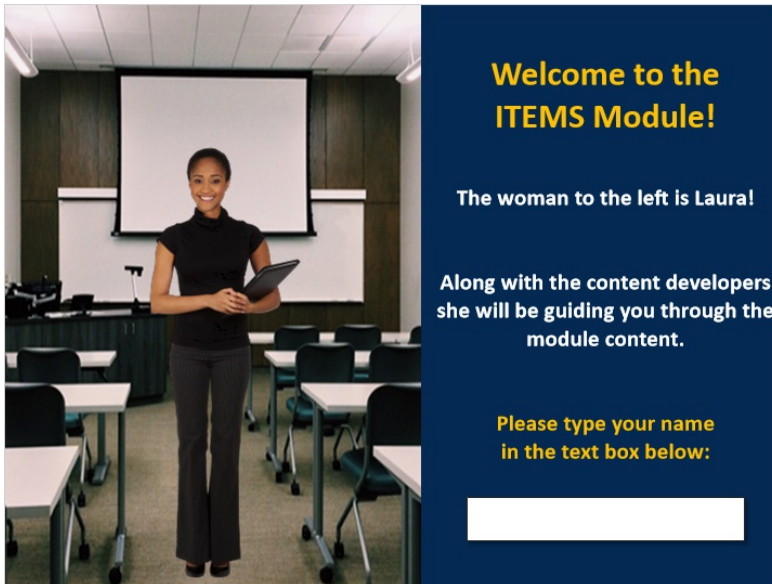
**Welcome to the
ITEMS Module!**

The woman to the left is Laura!

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

Please type your name
in the text box below:

Untitled Layer 1 (Slide Layer)



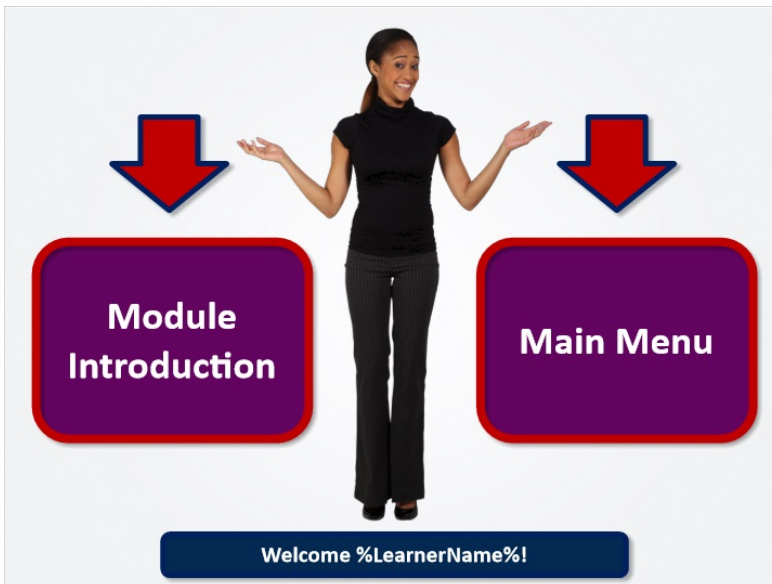
**Welcome to the
ITEMS Module!**

The woman to the left is Laura!

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

Please type your name
in the text box below:

1.5 Path Choice

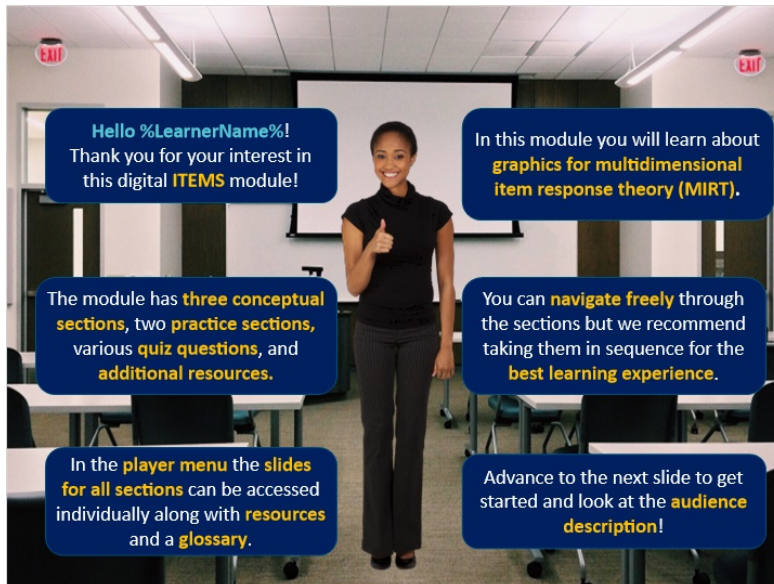


**Module
Introduction**

Main Menu

Welcome %LearnerName%!

1.6 Overview




1.7 Target Audience

Target Audience

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

- graduate students and faculty 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

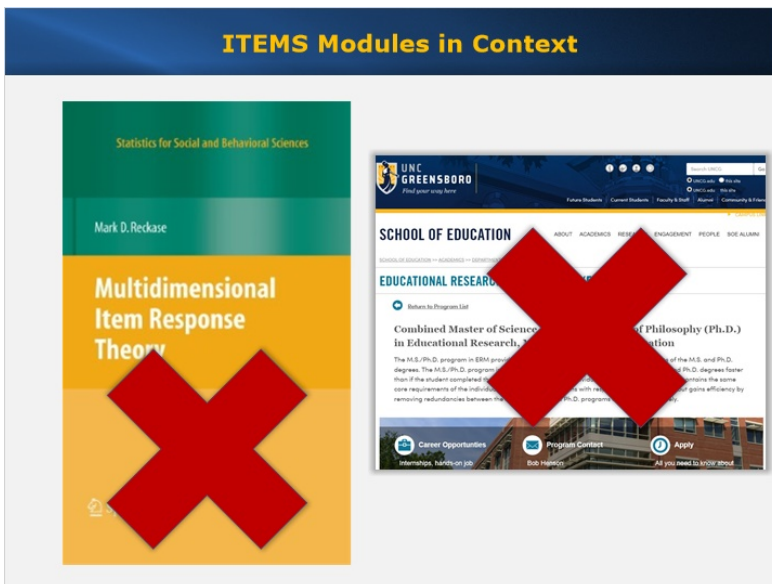


However, we hope that you find the information in this module useful no matter what your official title or role in an organization is!

1.8 Expectations (I)




1.9 Expectations (II)



1.10 Learning Objectives

Learning Objectives



1. Understand how to interpret evidence of dimensionality
2. Understand how the concepts and models change from UIRT to MIRT
3. Understand different ways to represent 2-D IRT item and test characteristics
4. Understand different ways to represent true score information and score scale consistency using centroids
5. Create MIRT graphics using an Rshiny app

1.11 Prerequisites


Prerequisites

To get the most out of this module, it is beneficial to:


- ✓ Be familiar with the underpinnings of classical test theory and its applications in assessment
- ✓ Understand the underlying theory, applications, and estimation of item response theory unidimensional models
- ✓ Have taken a graduate course in multivariate statistics, which includes factor analysis, principal components analysis, and multiple regression
- ✓ Have a working knowledge of running R

1.12 Resources

Resources



Reckase, M. D. (2009). *Multidimensional item response theory*. New York, NY: Springer.




Ackerman, T. A., Gierl, M. J., Walker, C. M. (2003). *Using multidimensional item response theory to evaluate educational and psychological tests*. Available in the ITEMS portal.

1.13 Module Citation

Module Citation

Ackerman, T., & Xie, Q. (2021). Multidimensional item response theory graphics (Digital ITEMS Module 23). *Educational Measurement: Issues and Practice*, 40(1), 105-106.



1.14 Main Menu

Main Menu

Theory

- 01 Response Data Dimensionality [10 Minutes]
- 02 Extending UIRT to MIRT [15 Minutes]
- 03 Plots for 2-D Items and Tests [45 Minutes]

Practice

- 04 Creating MIRT Plots in Rshiny [20 Minutes]
- 05 Data Activity [20 Minutes]
- 06 Quizzes [10 Minutes]

Help with Navigation

Navigation (Slide Layer)

Module Navigation

Home Button
Always brings you back to the **Main Menu**

Submenu Button
Always brings you back to last **Selection Point**

Player Functionalities
Direct access to all **slides, glossary, and resources**


2. Response Data Dimensionality

2.1 Cover: Response Data Dimensionality



Section 1:
Response Data Dimensionality
[10 Minutes]



2.2 Learning Objectives: Response Data Dimensionality



Learning Objectives

1. Know what we mean when we say a test is multidimensional
2. Identify some of the features of items and examinees that can cause multidimensionality
3. Identify a common approach practitioners use to examine the dimensionality of their tests
4. Examine dimensionality of simulated unidimensional and two-dimensional data



2.3 Understanding Dimensionality



Understanding Dimensionality

- A single item is always unidimensional because it measures only one skill or one composite of multiple skills
- Two or more items can be multidimensional if each item measures a different skill or different skill composites
- It is important to understand which skills or skill composite are being measured in order to articulate the meaning of the score scale, draw subsequent interpretations, and model the data appropriately



2.4 Test Dimensionality (I)



Test Dimensionality

- Test data represents the interaction between examinees and items
- Tests produce multidimensional response data when:
 - Items are capable of measuring multiple skills
 - Examinees differ in levels of skill proficiencies
- If all the items on a test only measure one skill, the test will yield only unidimensional response data, regardless of whether examinees have varying levels of proficiency on multiple skills



2.5 Dimensionality (II)



Test Dimensionality

- If the items on a test measure multiple skills but the examinees vary in their levels of proficiency on only one of the skills, the test will yield only unidimensional response data
- It is very important for practitioners to not only understand the skills or skill composites needed to correctly respond to an item, but also understand the skills or skill composite of their examinees


2.6 Causes of Dimensionality (I)





Causes of Dimensionality

Some of the potential causes of dimensionality that may enable **items** to measure multiple skills:

- Items measure different content
- Items measure different levels of cognitive reasoning
- Phrasing and wording of items can resonate more meaning for one group of examinees than others




2.7 Causes of Dimensionality (II)



  **Causes of Dimensionality**

Some of the potential causes of dimensionality that may cause **examinees** to vary in their levels of proficiency on a skill or skill composite:

- Different backgrounds or experiences due to socioeconomic status or educational opportunities
- Individual differences in interests or passions
- Different first languages
- Special needs requiring assistance for disabilities that may be physical, mental, or psychological



2.8 Assessing Dimensionality

  **Assessing Dimensionality**

- There are many ways to assess the dimensionality of response data
- We will examine a scree plot of eigenvalues from a factor analysis of the inter-item tetrachoric correlation matrix
- The goals of this method are two-fold:
 - a) Identify the dimensionality of the data to determine which IRT model is most appropriate
 - b) Articulate what skills the scales represent

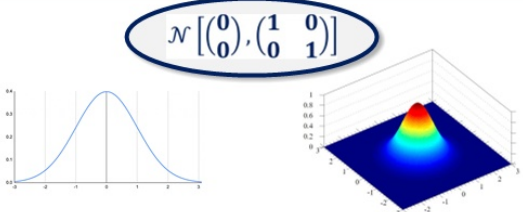
2.9 Example of Assessing Dimensionality

Example of Assessing Dimensionality

- To illustrate the distinction between scree plots for unidimensional and multidimensional data, we generated two 20-item tests for 1,000 examinees
- Dataset 1 used a unidimensional 2PL model for examinees from a $N(0,1)$ underlying ability distribution
- Dataset 2 used a two-dimensional compensatory model for examinees from an underlying bivariate normal distribution:

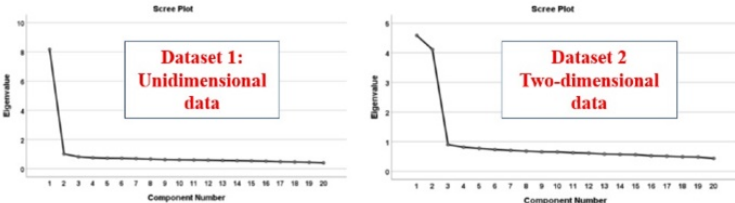
$\mathcal{N}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right]$

Dataset 1 vs. Dataset 2



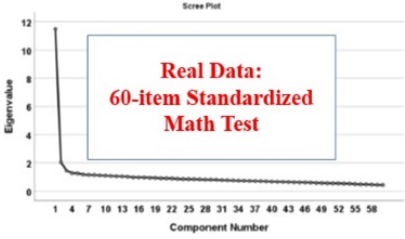
2.10 Scree Plots (I)

Scree Plots



- The inter-item correlation matrix for each dataset was factor analyzed and corresponding scree plots were created for both
- For Dataset 1, there is only one strong dominant dimension
- For Dataset 2, there are two dominant dimensions

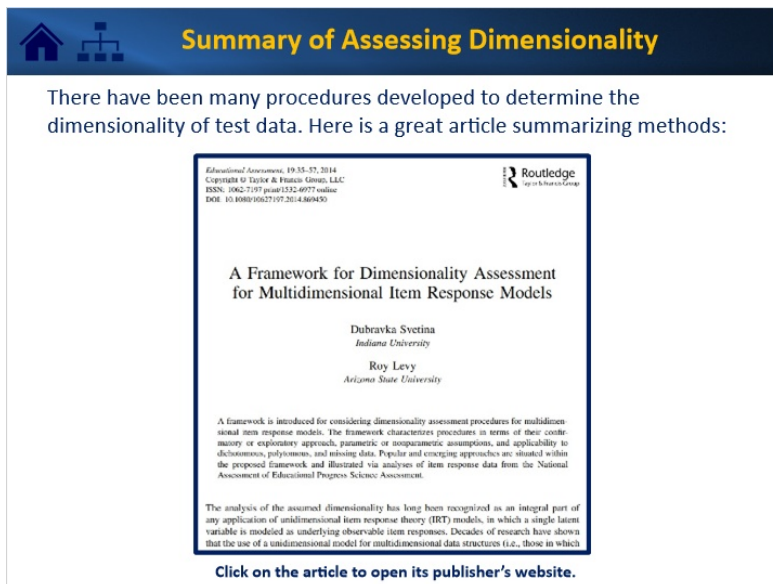
2.11 Scree Plots (II)



**Real Data:
60-item Standardized
Math Test**

- Factor analysis of a dataset from a standardized math test which had 60 items and 1500 examinees
- The scree plot indicates that these data are also unidimensional because of the dominant first factor

2.12 Summary of Assessing Dimensionality



There have been many procedures developed to determine the dimensionality of test data. Here is a great article summarizing methods:

Educational Assessment, 19(35-37), 2014
Copyright © Taylor & Francis Group, LLC
ISSN: 1062-7197 print/1532-6977 online
DOI: 10.1080/10627197.2014.869430

**A Framework for Dimensionality Assessment
for Multidimensional Item Response Models**

Dubravka Svetina
Indiana University

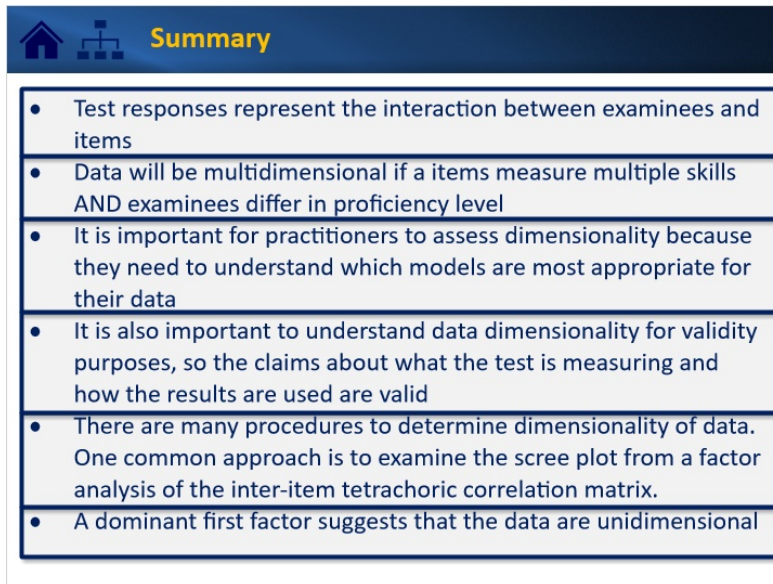
Roy Levy
Arizona State University

A framework is introduced for considering dimensionality assessment procedures for multidimensional item response models. The framework characterizes procedures in terms of their constructivist or exploratory approach, parametric or nonparametric assumptions, and applicability to dichotomous, polytomous, and missing data. Popular and emerging approaches are situated within the proposed framework and illustrated via analyses of item response data from the National Assessment of Educational Progress Science Assessment.

The analysis of the assumed dimensionality has long been recognized as an integral part of any application of unidimensional item response theory (IRT) models, in which a single latent variable is modeled as underlying observable item responses. Decades of research have shown that the use of a unidimensional model for multidimensional data structures (i.e., those in which

Click on the article to open its publisher's website.

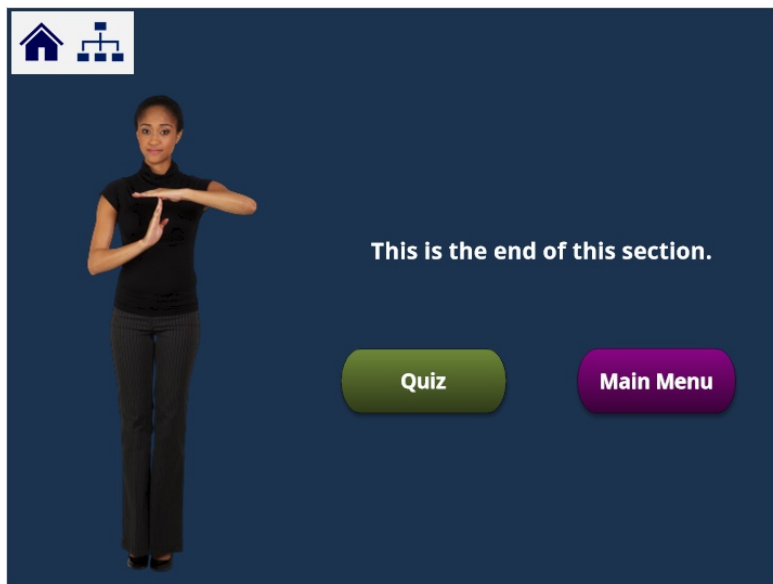
2.13 Summary: Response Data Dimensionality



The slide features a dark blue header with a home icon, a tree diagram icon, and the word "Summary" in yellow. Below the header is a white box with a blue border containing a bulleted list of six points.

- Test responses represent the interaction between examinees and items
- Data will be multidimensional if a items measure multiple skills AND examinees differ in proficiency level
- It is important for practitioners to assess dimensionality because they need to understand which models are most appropriate for their data
- It is also important to understand data dimensionality for validity purposes, so the claims about what the test is measuring and how the results are used are valid
- There are many procedures to determine dimensionality of data. One common approach is to examine the scree plot from a factor analysis of the inter-item tetrachoric correlation matrix.
- A dominant first factor suggests that the data are unidimensional

2.14 Bookend: Response Data Dimensionality



The slide features a dark blue background. In the top left corner, there is a white box containing a home icon and a tree diagram icon. On the left side, a woman in a black top and pants is performing a sign language gesture. In the center, the text "This is the end of this section." is displayed. At the bottom, there are two buttons: a green "Quiz" button and a purple "Main Menu" button.

3. Extending UIRT to MIRT

3.1 Cover: Extending UIRT to MIRT



3.2 Learning Objectives: Extending UIRT to MIRT

The slide has a dark blue header with a home icon, a tree icon, and the title "Learning Objectives" in yellow. Below the header is a target icon with an arrow in the bullseye. There are four rounded rectangular boxes containing the following objectives:

1. Understand how to extend UIRT models to 2-D IRT models
2. Understand how the UIRT ICC becomes a 2-D ICS
3. Explain why the M2PL model is known as a compensatory model
4. Examine how unidimensional polytomous models can be extended to their 2-D counterparts

3.3 The 2PL IRT Model

The 2-PL IRT Model

In the unidimensional 2PL model, the probability of examinee j responding correctly to item i can be expressed as:

$$P_{ij} = \frac{1}{1 + e^{-1.7a_i(\theta_j - b_i)}}$$

Discrimination parameter

Latent ability parameter

Difficulty parameter

3.4 Item Characteristic Curves for UIRT 2PL

Item Characteristic Curve (UIRT-2PL)

- Graphically, the 2PL model for an individual item can be illustrated as an item characteristic curve (ICC)
- In this model, the steepness of the ICC relates directly to the a parameter (discrimination) and the location relates to the b parameter (difficulty)

ITEM 1: $a = 2.0$ $b = 0.0$
ITEM 2: $a = 1.0$ $b = 0.0$

Item 1 is more discriminating (i.e., a steeper ICC)

Item 2 is easier at this ability

Item 1 is easier at this ability

3.5 The 2-Dimensional Compensatory Model

🏠
The Two-dimensional Compensatory Model

- In the multidimensional 2PL model (M2PL), the probability of examinee j responding correctly to item i can be expressed as:

$$P_{ij} = \frac{1}{1 + e^{-1.7(a_{1i}\theta_{1j} + a_{2i}\theta_{2j} + d_i)}}$$

Discrimination parameters

Latent ability parameters

Difficulty parameter

- There are discrimination parameters a_1 and a_2 and ability parameters θ_1 and θ_2
- Regardless of the number of dimensions, there is only ever one difficulty parameter, d

3.6 2-D Compensatory Item Surface

🏠
2-D Compensatory Item Surface

- Under the 2-D compensatory model, items can be illustrated with item characteristic surfaces (ICSs)
- The steepness of the ICS relates directly to the a parameters
- The relative sizes of the a s indicate which composite of θ_1 and θ_2 is being best measured. If $a_1 = a_2$, θ_1 and θ_2 are being measured equally well. If $a_1 > a_2$, θ_1 is being given greater emphasis, and if $a_1 < a_2$, θ_2 is being given greater emphasis.

Item Characteristic Surface
A1=0.80 A2=1.40 D = -0.30

3.7 Compensatory Surface Contour

Compensatory Surface Contour

- Equiprobability contours are always parallel for the compensatory model
- Every (θ_1, θ_2) combination on the same contour has the same response probability
- When $a_1 = a_2$, compensation between abilities is optimal: examinees who are high on θ_1 but low on θ_2 have the same probability of answering correctly as an examinee with the opposite profile because being high on one ability compensates for being low on the other.

3.8 Other Unidimensional Models Extended to 2D

Other Unidimensional Models Extended to 2-D

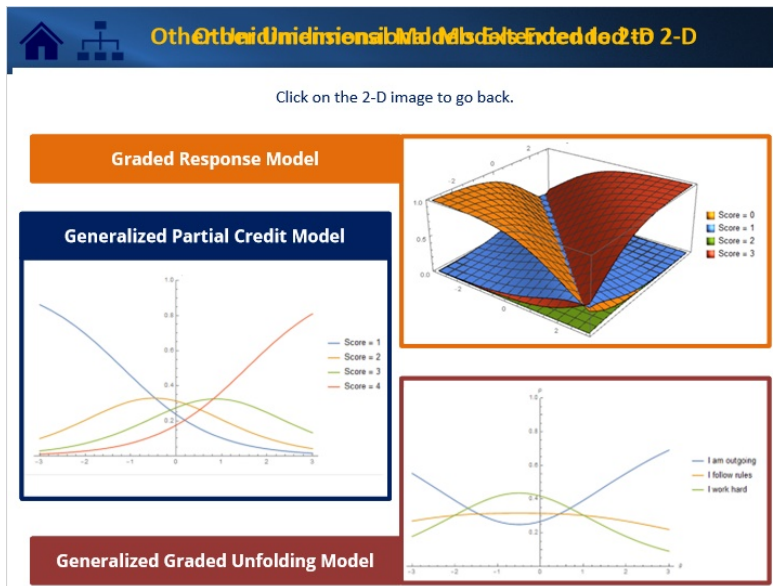
All common unidimensional models can be extended to two dimensions. Click the graphs to see how!

Graded Response Model

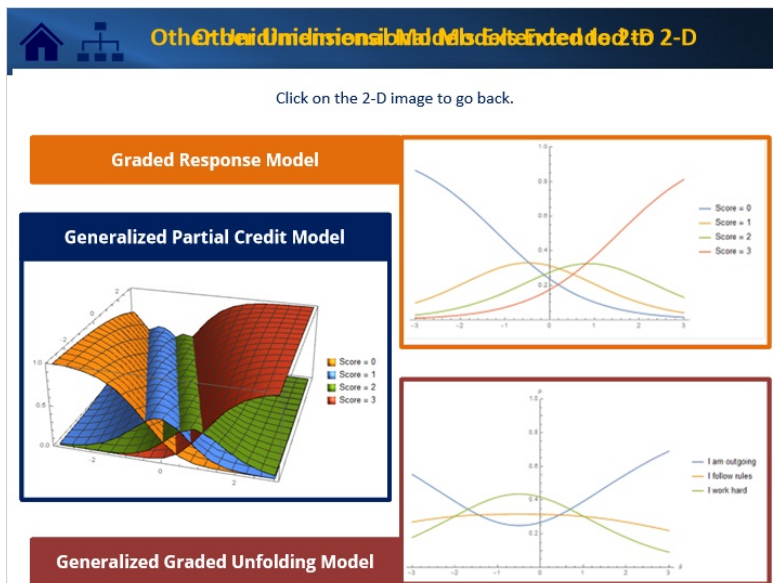
Generalized Partial Credit Model

Generalized Graded Unfolding Model

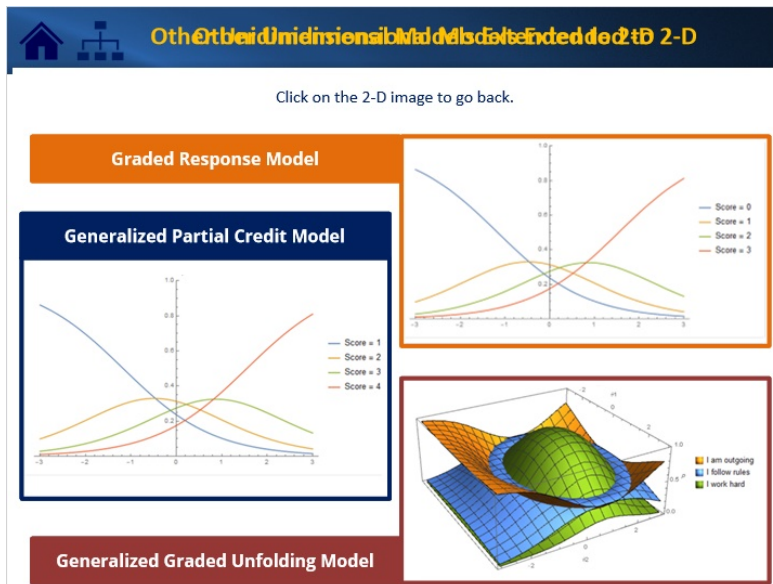
2DGRM (Slide Layer)



2DGPCM (Slide Layer)



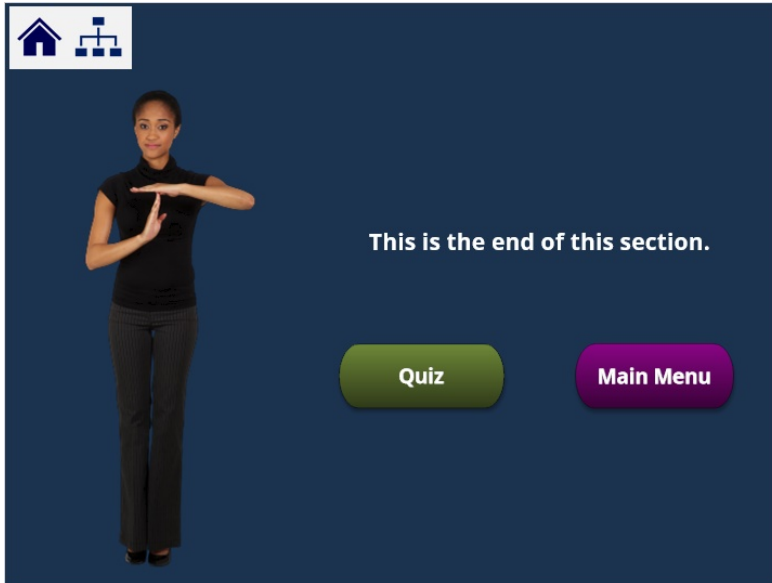
2DGGUM (Slide Layer)



3.9 Summary: UIRT and MIRT Concepts

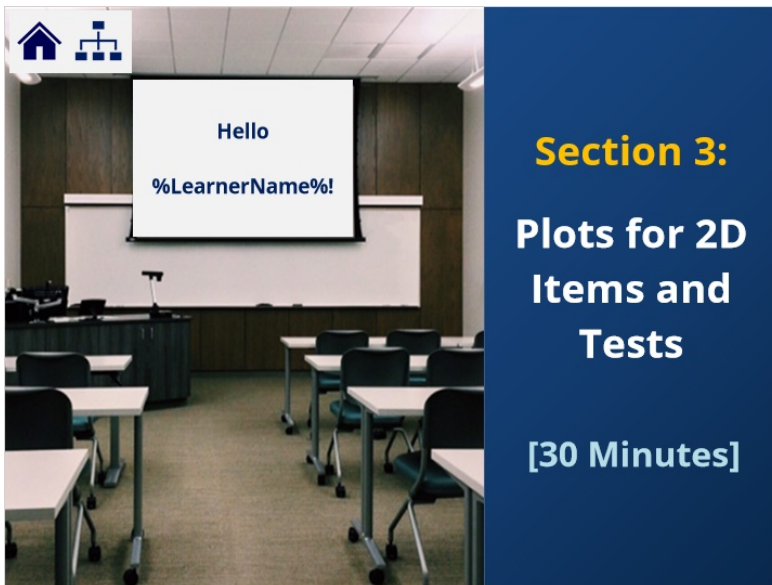
- Summary
- M2PL model requires a discrimination parameter and ability parameter for each dimension, but only one difficulty parameter
 - Under the M2PL, the product of the discrimination and the ability parameters for each dimension are additive in the logit, so being high on one ability can compensate for being low on the other
 - Compensation is optimal when $a_1 = a_2$ or when the model weights each ability equally
 - Equiprobability contours of the compensatory model are always parallel. All examinees lying on the same equiprobability contour have the same response probability.
 - Larger a values lead to steeper ICSSs, making the equiprobability contours closer together
 - 2-D extensions are possible for other common IRT models as well

3.10 Bookend: Extending UIRT to MIRT





4. Plots for 2D Items and Tests


4.1 Cover: Plots for 2D Items and Tests



4.2 Learning Objectives: Plots for 2D Items and Tests





Learning Objectives





1. Understand how 2-D response surfaces are represented as item vectors and 1-D test characteristic curves generalize to 2-D test characteristic surfaces
2. Understand how item and test information in a 2-D latent space is calculated
3. Understand how test information is best captured using a clamshell plot and how a test information number plot profiles which composites are best measured
4. Understand score scale consistency and how to interpret a centroid plot when a scale is not consistent

4.3 Topic Selection



Plots for 2-D Items and Tests



Click on each circle to learn more 

Section Summary

4.4 Bookmark: Basics of 2-D Plots



4.5 Item Vectors (I)

Item Vectors

- Plotting multiple ICSs can be very cumbersome
- Perhaps the best representation of 2-D tests is to represent each item as a vector in the latent ability plane

Guidelines for process:

1. All vectors lie on lines that pass through the origin
2. Vectors can lie only in the first and third quadrants because a parameters are constrained to be positive
3. Vectors representing easy items lie in the third quadrant; vectors representing difficult items lie in the first quadrant

- To create a vector, you need to know the length, the origin, and the angle with the θ_1 -axis

4.6 Item Vectors (II)

Item Vectors

- The length of the vector indicates how discriminating the item is. This value is called MDISC.

$$\text{MDISC} = \sqrt{a_1^2 + a_2^2}$$

- The tail of the vector lies on the $p = .50$ equiprobability contour. The signed distance from the origin to this contour is denoted as MDIFF

$$\text{MDIFF} = \frac{-d}{\text{MDISC}}$$

- The angular direction indicates the composite of ability that the item is best measuring

$$\alpha = \cos^{-1} \left(\frac{a_1}{\text{MDISC}} \right)$$

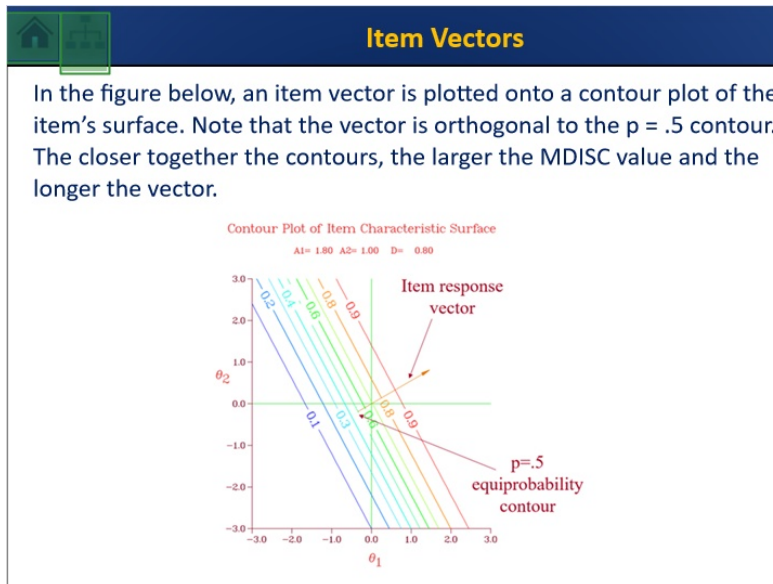
4.7 Item Vectors (III)

Item Vectors

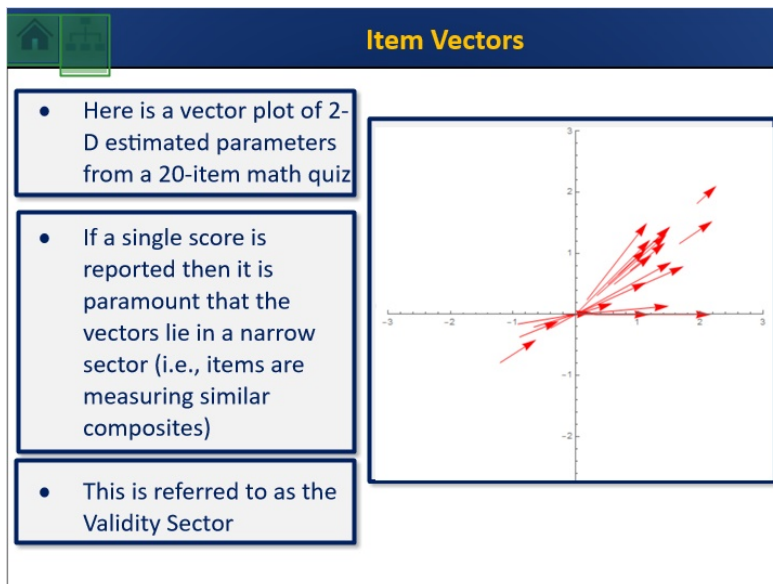
Vectors are actually projections of the direction of maximum discrimination or slope onto the latent ability plane

The diagram illustrates the relationship between the response surface and the latent ability plane. The response surface is a grid of points representing the probability of a correct response for different ability levels. The direction of maximum slope is the direction of steepest increase in probability. The item vector is the vector from the origin to the point on the response surface that corresponds to the item's characteristics. The projected item vector is the projection of the item vector onto the latent ability plane.

4.8 Item Vectors (III)



4.9 Item Vectors (IV)



4.10 Item Vectors (V)

HomeItem Vectors

By color coding the vectors to match different content areas we can determine

- Are items from a certain content area more discriminating or more difficult?
- Do different items from different content areas measure different ability composites?
- How similar are the vector profiles for different yet "parallel" forms?

4.11 Example of Item Vectors for the 101-Item LSAT

HomeExample of Item Vectors for the 101-item LSAT

- Here is an example of color-coding vectors from a 101-item LSAT test with three categories of items
- Note how each category lies in a narrow sector
- By color-coding we can also get a better sense of how to define our latent abilities from our table of specifications for each content area



Total LSAT Test

θ_2

θ_1

Analytical Reasoning
Logical Reasoning
Reading Comprehension

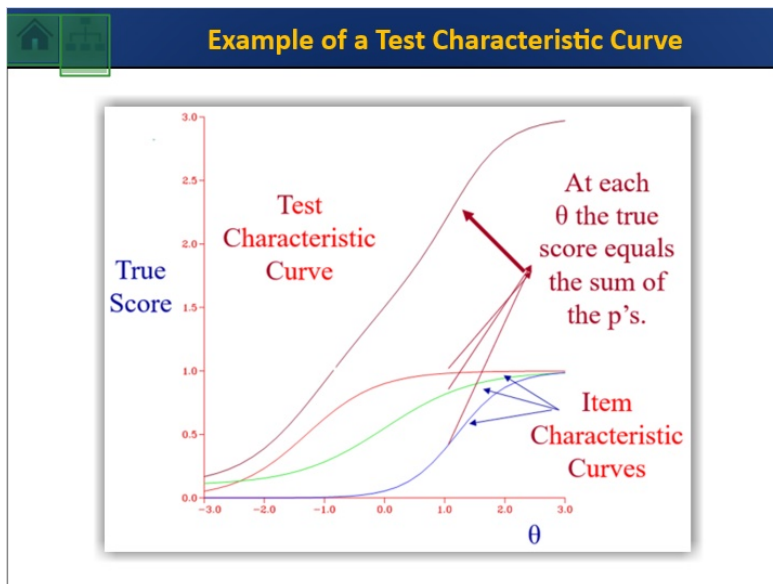
4.12 Test Characteristic Curves

Test Characteristic Curves

- In unidimensional IRT models, we can “sum up” the probabilities at each theta from ICCs to obtain a test characteristic curve (TCC)
- This curve enables us to determine the expected score (e.g., true score) of each ability on the θ -axis
- This curve has many functions, including mapping the θ scale onto the true score scale and equating across test forms.

4.13 Example of a Test Characteristic Curve



4.14 Test Characteristic Surface

Test Characteristic Surface

- We can extend the concept of the TCC to the 2-D case, in which case we will have a Test Characteristic Surface (TCS)
- This allows us to determine the expected score, or true score, for each (θ_1, θ_2) ability

Everyone along the same contour would have the same true score.

4.15 Test Characteristic Surface (II)

Test Characteristic Surface (MIRT-2PL)

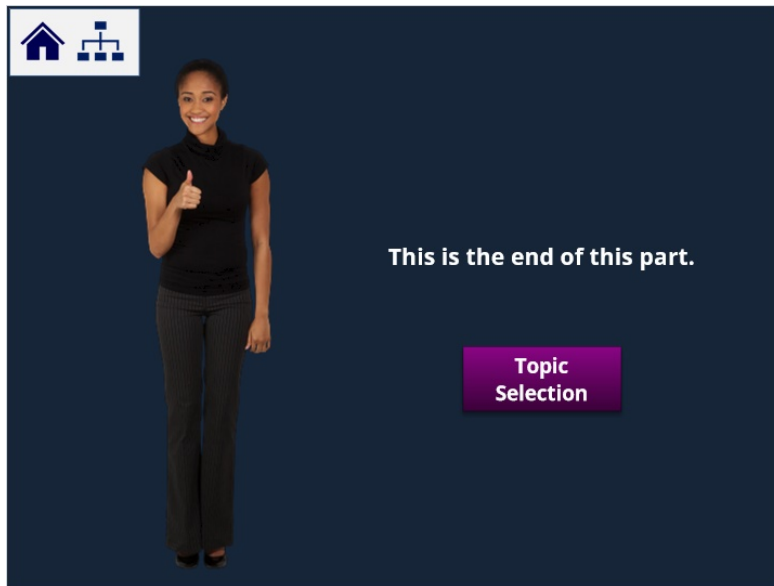
- One important insight afforded by true score surfaces is that one can take the difference between the surfaces for two forms after their parameters are on the same scale
- This allows us to see how multidimensionally parallel test forms are at right is the difference surface for Form 39F – 39B

Examinees in this region would have a greater true score on Form 39B

Examinees in this region would have a greater true score on Form 39F

The zero-difference plane represents no true score surface differences

4.16 Bookend: Basics of 2-D Plots



4.17 Bookmark: Test Information Plots



4.18 Test Information Plots (I)

HomeTest Information Plots

- In unidimensional IRT models the formula for computing information for item i can be expressed as:
$$I_i(\theta) = 2.89P_iQ_i$$
 for the 1PL,
$$I_i(\theta) = 2.89a_i^2P_iQ_i$$
 for the 2PL, and
$$I_i(\theta) = 2.89a_i^2Q_i \frac{(P_i - c_i)^2}{(1 - c_i)^2P_i}$$
 for the 3PL
- Note that when $c_i = 0$, the 3PL formula is the same as the 2PL, and when $a_i = 1$, the 2PL is the same as the 1PL

4.19 Test Information Plots (II)

HomeTest Information Plots

- To compute the information for the 2-D M2PL model, one would have to use the following formula:
$$I_\alpha(\theta) = P(\theta)Q(\theta) \left(\sum_{v=1}^2 \alpha_v \cos \alpha_v \right)^2$$
- where $P(\theta)$ is the M2PL model, $Q(\theta)$ is $1 - P(\theta)$, and α is the composite direction for which you want to compute the information.

4.20 Test Information Plots (III)

Home
Test Information Plots

- For example, assume $a_1 = 1$, $a_2 = 1$ and $d = 0$. To calculate the information in a 45° angle at the point $(0,0)$, first calculate $P(0,0)$ and $Q(0,0)$.

- Under the M2PL, $P(0,0) = .5$ and $Q(0,0) = .5$, so:

$$I_{45}(1,1) = P(\theta)Q(\theta)(\sum_{v=1}^m \alpha_v \cos \alpha_v)^2$$

$$I_{45}(1,1) = (.5)(.5)((1.0)(\cos 45) + (1.0)(\cos 45))^2$$

$$I_{45}(1,1) = (.5)(.5)[(1.0)(.707) + (1.0)(.707)]^2$$

$$I_{45}(1,1) = .499$$

- Continuing to move away from the origin in the same direction, $P(1,1) = .967$ and $Q(1,1) = .033$, yielding $I_{45}(1,1) = .064$. If we move out further to $(2,2)$, $I_{45}(2,2) = .002$. Information decreases because the ICS flattens out in this direction.

4.21 Test Information Plots (IV)

Home
Test Information Plots

- Assume we are still working with the same item but now want to calculate $I_{30}(0,0)$:

$$I_{30}(0,0) = P(\theta)Q(\theta)(\sum_{v=1}^m \alpha_v \cos \alpha_v)^2$$

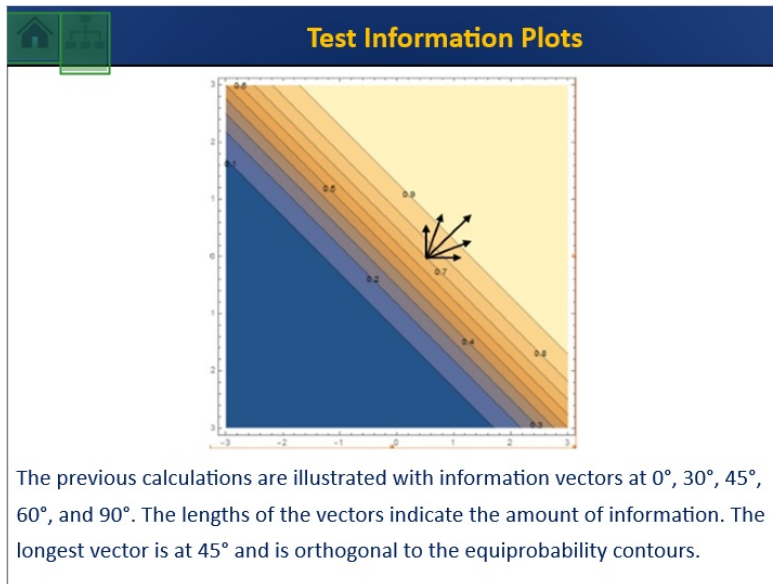
$$I_{30}(0,0) = (.5)(.5)((1.0)(\cos 30) + (1.0)(\cos 60))^2$$

$$I_{30}(0,0) = (.5)(.5)[(1.0)(.866) + (1.0)(.5)]^2$$

$$I_{30}(0,0) = .466$$

- Further, $I_0(0,0) = .25$. The angle of maximum information when $a_1 = a_2$ is at 45° , or orthogonal to the equiprobability contours. As we move away from this direction, the surface is not as steep, the item is not as discriminating, and the information decreases.

4.22 Test Information Plots



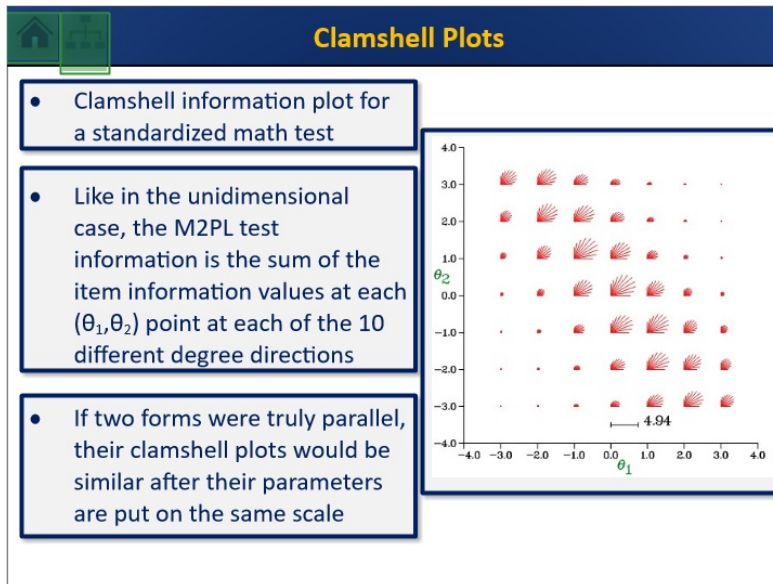
4.23 Clamshell Plots (I)

The figure is a slide titled "Clamshell Plots". It contains a list of three steps for creating a clamshell plot. Each step is enclosed in a blue-bordered box. The text describes the process of computing information at 49 points on a 7x7 grid, evaluating ten directions at each point, and the resulting clamshell-like vectors.

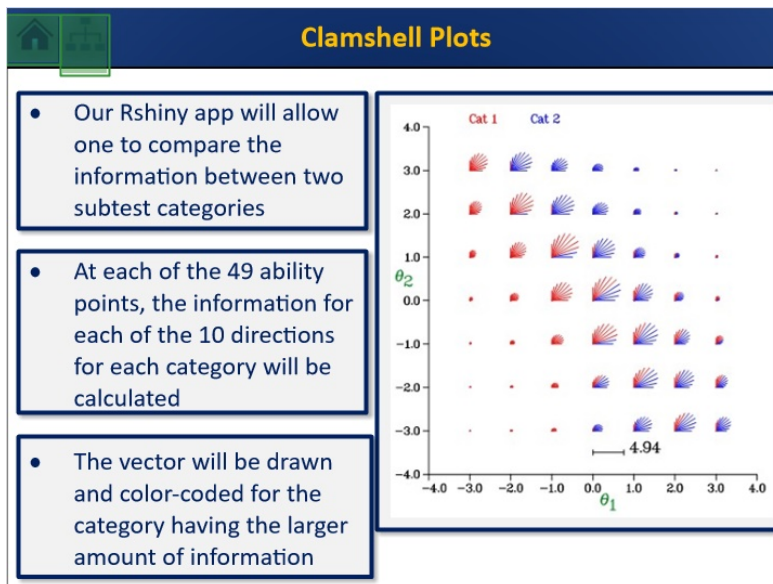
Reckase and McKinley (1991) show a way to illustrate the amount of information for the M2PL model using a “clamshell” plot. To create this plot:

1. The amount of information is computed as 49 uniformly spaced points on a 7x7 grid on the (θ_1, θ_2) plane
2. At each point, the amount of information the test provides is computed for ten different directions (or ability composites) from 0 to 90 degrees in 10-degree increments
3. The resulting ten vectors at each ability point resemble a clamshell

4.24 Clamshell Plots (II)



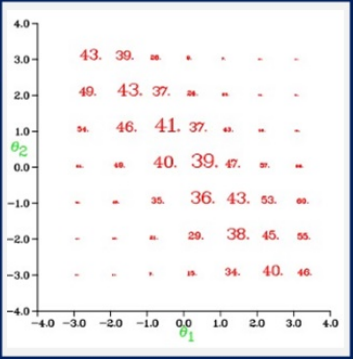
4.25 Clamshell Plots (III)



4.26 Maximum Information


HomeMaximum Information

- Another way to display the information of a 2-D test is with a number plot
- This plot shows the direction of the maximum information at each of the 49 grid points with font size proportional to information size
- In this example, information at the point (0,0) is maximized at 39°. This is also the location and direction of the maximum value of information for the ability plane.



4.27 Bookend: Examining Score Scale Consistency

HomeTree



This is the end of this part.

Topic Selection

4.28 Bookmark: Examining Score Scale Consistency



4.29 Score Scale Consistency (I)

Score Scale Consistency

- When a test is unidimensional, we automatically assume that all the items are measuring the same skill or the same composite of multiple skills
- When a test is two-dimensional, we assume that items are measuring different composite of the two latent skills
- It is important to ask whether there is any confounding of difficulty and dimensionality
 - Example: do easy items measure mainly θ_1 and difficult items measure mainly θ_2 ?
 - If so, low scores would represent different levels of θ_1 -ability and high scores different levels of θ_2 -ability.

4.30 Score Scale Consistency (II)

Score Scale Consistency

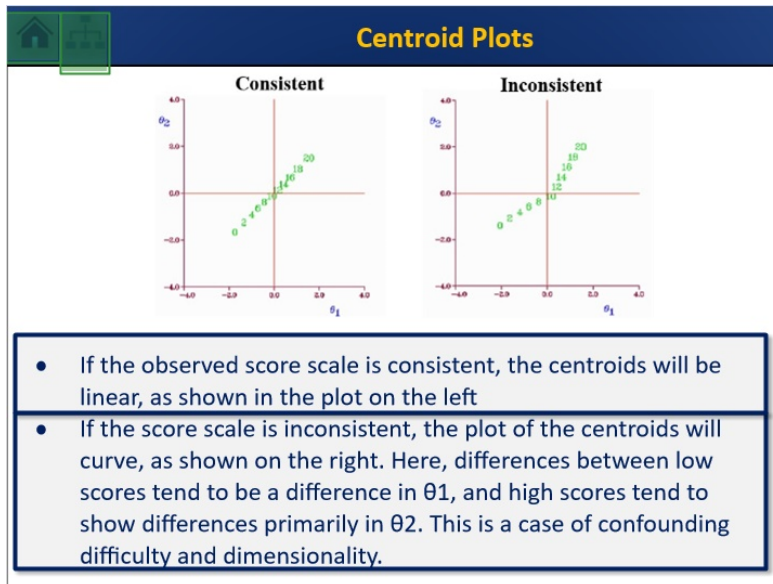
- Consider an algebra test with some story problems
- Easy items could be measuring θ_1 , algebraic reasoning
- Story problems, which always tend to be more difficult, require reading and understanding a real-life situation and then applying algebraic reasoning.
- These items could be measuring more reading ability, which would be θ_2

4.31 Conditional Estimation Using Centroids

Conditional Estimation Using Centroids

- To gain insight about score scale consistency, we create a centroid plot
- The purpose of this plot is to examine the means of the conditional distributions for each score category:
$$(\bar{\theta}_1, \bar{\theta}_2 | X = x)$$
- The number of the score scale will be printed at the $(\bar{\theta}_1, \bar{\theta}_2)$ of the distribution for all the people with that observed score
- For an interpretation of the score scale to be consistent, the centroid should form a linear pattern

4.32 Centroid Plots





4.33 Bookend: Test Information

This is the end of this part.

Topic Selection



4.34 Summary: Plots for 2D Items and Tests (I)



Summary

- With the M2PL, best to represent items as vectors
 - Length of vector = MDISC = discrimination
 - Location of tail in relation to origin = MDIFF = difficulty
 - Angle of vector with θ_1 axis = α = composite of (θ_1, θ_2) best measured
 - Tail of vectors lie on, and are orthogonal to, $P = .5$ probability contour
- Color-coding helps in understanding variation of composites
- The TCS helps determine the true score for all (θ_1, θ_2) combinations in the latent ability plane.
 - All examinees lying on the same true score contour will be expected to score the same on a test
- One form's TCS can be subtracted from another's to determine how parallel the two forms are
- The TCS for the M2PL model changes in slope/steepness depending on the direction one travels across the surface

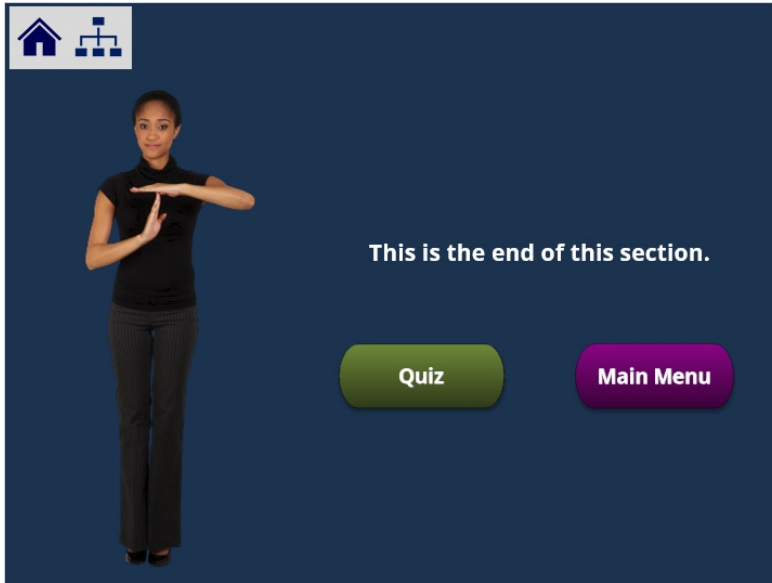
4.35 Summary: Plots for 2D Items and Tests (II)



Summary

- To represent the amount of information across the 2-D latent ability plane we create a "clamshell" plot
- The clamshells tell us which composites are being best measured across the ability plane
- Practitioners need to be cautious about interpreting the two-dimensional test results if they just report a single observed or a single scaled score
- To examine whether the score scale is consistent in interpretation we construct a centroid plot
- Problems arise when there is a confounding and easy item measure primarily one ability and difficult items measure primarily the other ability

4.36 Bookend: Plots for 2D Items and Tests





5. MIRT Plots in RShiny


5.1 Cover: MIRT Plots in Rshiny



5.2 Learning Objectives: MIRT Plots in RShiny





Learning Objectives



1. Access the MIRTGraph RShiny app online
2. Upload the item parameter file into MIRTGraph
3. Review how to plot items and tests in MIRTGraph


5.3 Access to RShiny



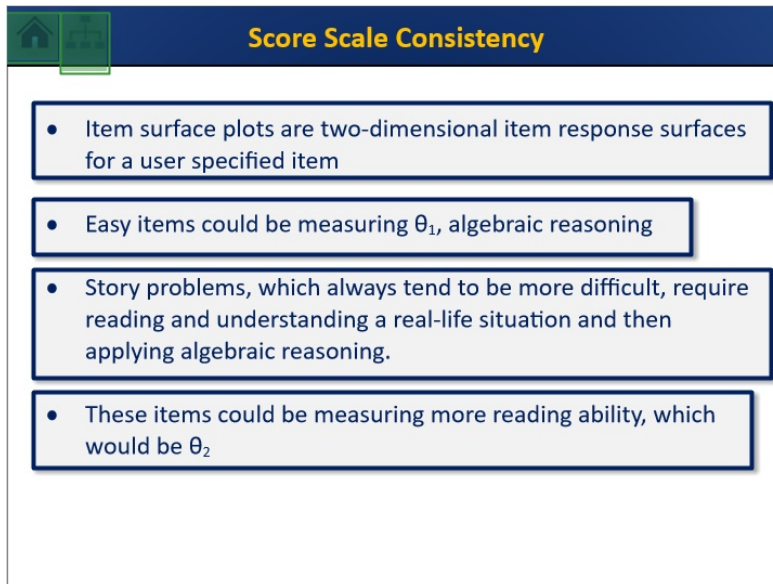
Access to Rshiny

The *MIRTGraph Rshiny* app can be accessed via the following link:

<https://qing-xie-uiowa.shinyapps.io/MIRTGraph/>



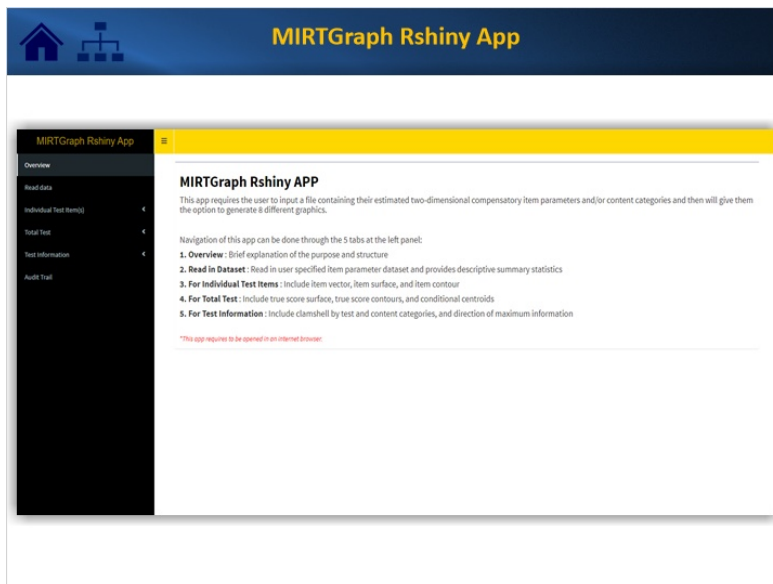
5.4 FORMATTT



Score Scale Consistency

- Item surface plots are two-dimensional item response surfaces for a user specified item
- Easy items could be measuring θ_1 , algebraic reasoning
- Story problems, which always tend to be more difficult, require reading and understanding a real-life situation and then applying algebraic reasoning.
- These items could be measuring more reading ability, which would be θ_2

5.5 MIRTGraph RShiny App



MIRTGraph Rshiny App

Overview

Read Data

Individual Test Items

Total Test

Test Information

Auth Trail

MIRTGraph Rshiny APP

This app requires the user to input a file containing their estimated two-dimensional compensatory item parameters and/or content categories and then will give them the option to generate 4 different graphics.

Navigation of this app can be done through the 5 tabs at the left panel:

1. **Overview**: Brief explanation of the purpose and structure
2. **Read in Dataset**: Read in user specified item parameter dataset and provides descriptive summary statistics
3. **For Individual Test Items**: Include item vector, item surface, and item contour
4. **For Total Test**: Include true score surface, true score contours, and conditional centroids
5. **For Test Information**: Include clamshell by test and content categories, and direction of maximum information

*This app requires to be opened in an internet browser.

5.6 Read in Data (I)

The screenshot shows the 'Read in Data' interface of the MIRTGraph Rahiny App. The main heading is 'Read in dataset'. Below it, there is a section titled 'Uploading Files' with a 'Choose CSV File' button and a 'Browse...' link. A dropdown menu is open, showing options for 'Header', 'Separator', 'Comma', 'Semicolon', 'Tab', 'Quote', 'None', 'Double Quote', and 'Single Quote'. The 'Data' and 'Summary' tabs are visible at the top right of the main content area.

5.7 Read in Data (II)

The screenshot shows the 'Read in Data' interface of the MIRTGraph Rahiny App. The main heading is 'Read in dataset'. Below it, there is a section titled 'Uploading Files' with a 'Choose CSV File' button and a 'Browse...' link. A dropdown menu is open, showing options for 'Header', 'Separator', 'Comma', 'Semicolon', 'Tab', 'Quote', 'None', 'Double Quote', and 'Single Quote'. The 'Data' and 'Summary' tabs are visible at the top right of the main content area. A table is displayed with the following data:

	A1	A2	B	Cost
1	0.79	1.26	-0.9	1
2	0.99	1.08	-1.2	1
3	0.96	0.94	1	1
4	0.87	0.87	-0.87	2
5	0.83	0.79	-1.06	1
6	0.91	0.99	-1.53	1
7	0.8	0.48	-0.61	1
8	0.8	0.87	-0.6	2
9	1.44	0.15	1.24	1
10	1.51	1.1	-0.89	2

Showing 1 to 10 of 20 entries. Previous 1 2 Next

5.8 Read in Data (III)

Read in dataset

Uploading Files

Choose CSV File

Browse... (data.csv)

Other options

Header

Separator

Comma

Semicolon

Tab

Quote

None

Double Quote

Single Quote

Par	Mean	Std	Median	Min	Max
a1	0.94	0.46	0.85	0.31	2.37
a2	0.76	0.48	0.78	0.09	1.76
D	-0.20	1.08	-0.60	-1.33	2.49

5.9 Item Vector Plots

Item vector

Please enter Item ID that you want to plot:

1,1,1

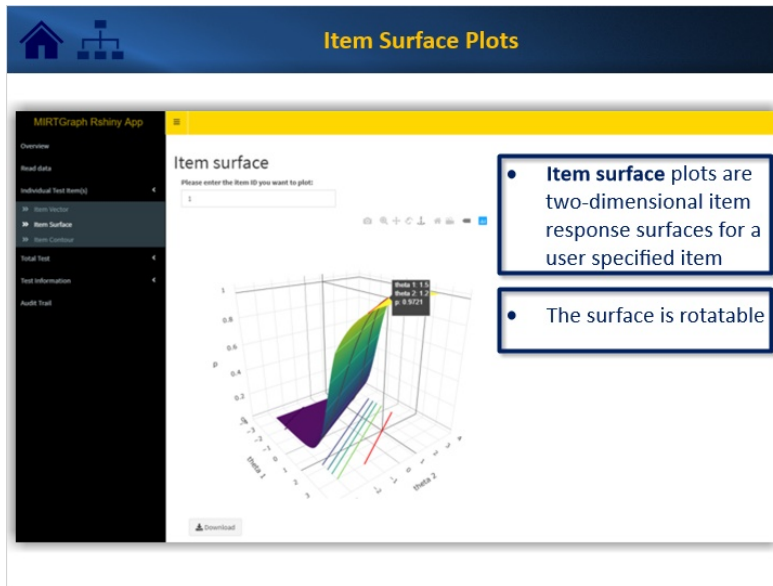
theta 2

theta 1

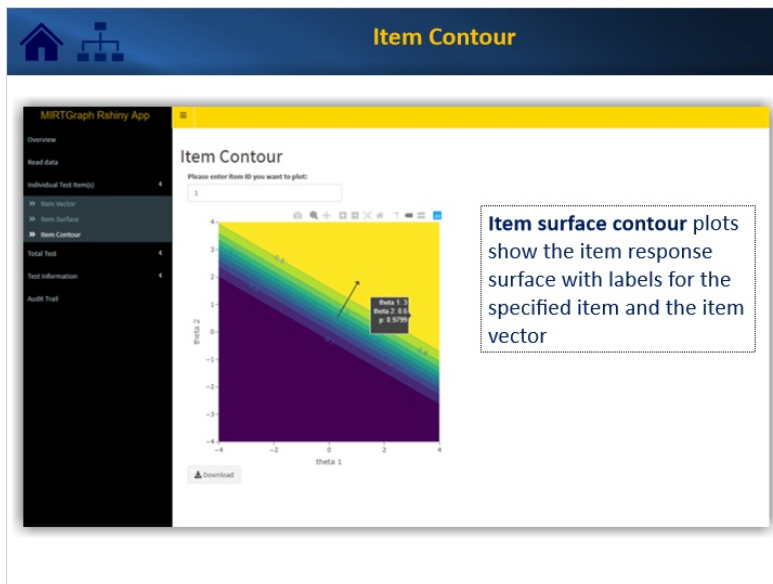
Download

- The tail of the vector lies on, as is orthogonal to, the $p = .5$ equiprobability contour
- The length of the vector is equal to MDISC
- The angle with the positive θ_1 axis is equal to the arccosine of $a1/\sqrt{MDISC}$

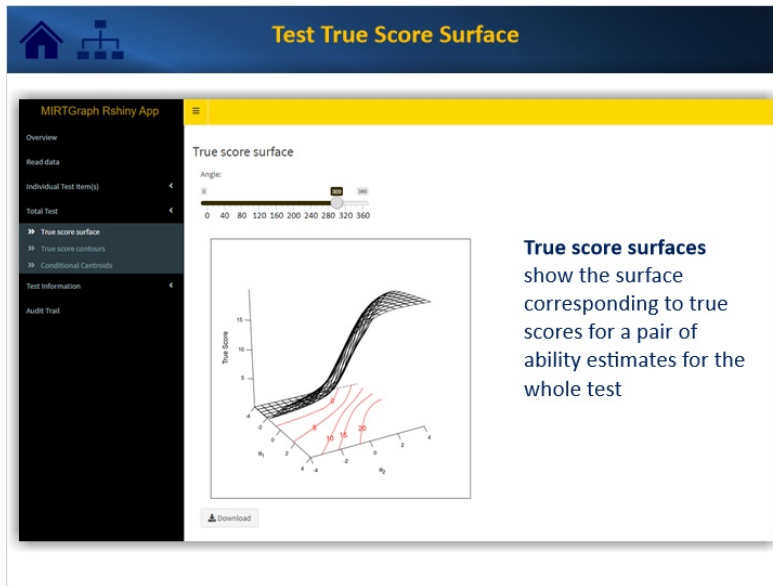
5.10 Item Surface Plots



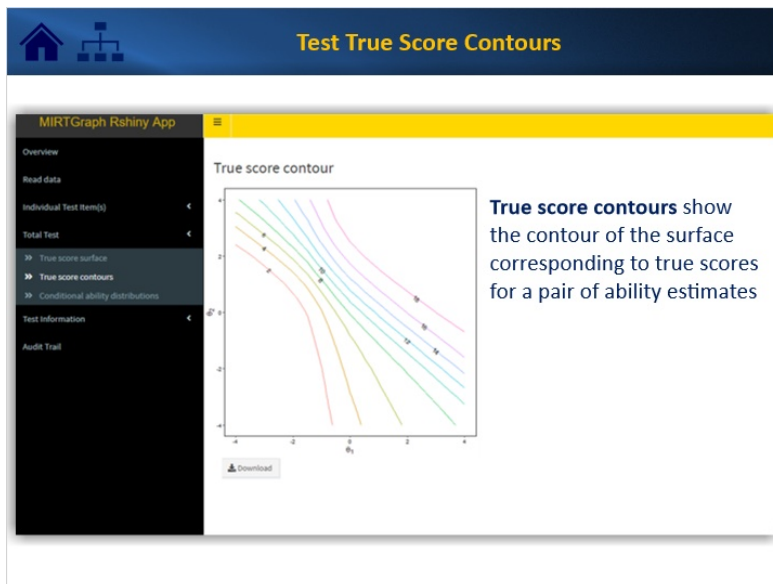
5.11 Item Contour Plots



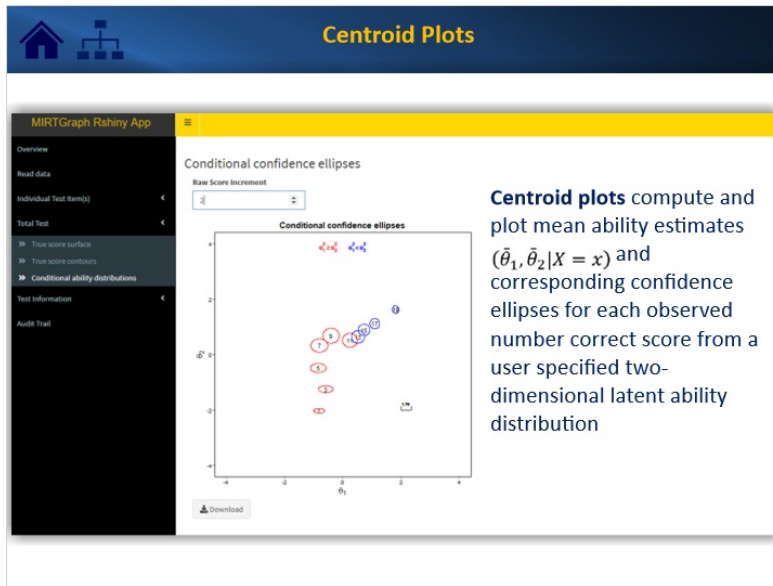
5.12 Test True Score Surfaces



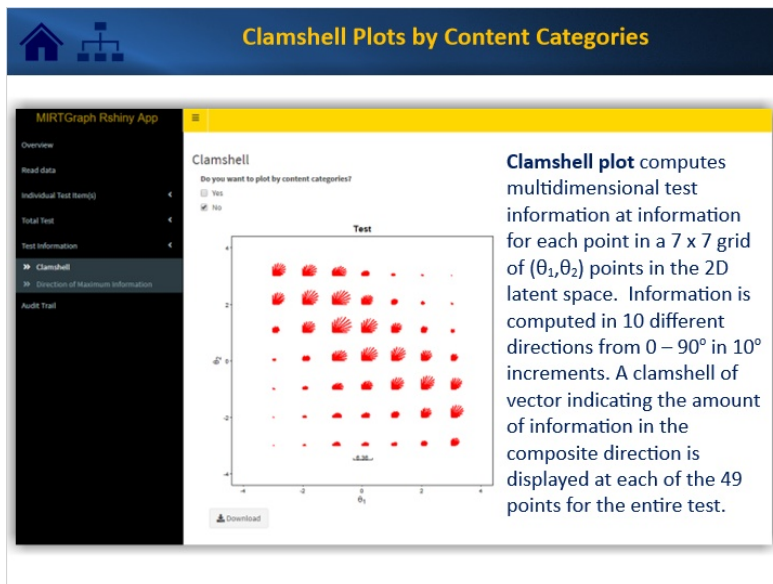
5.13 Test True Score Contours



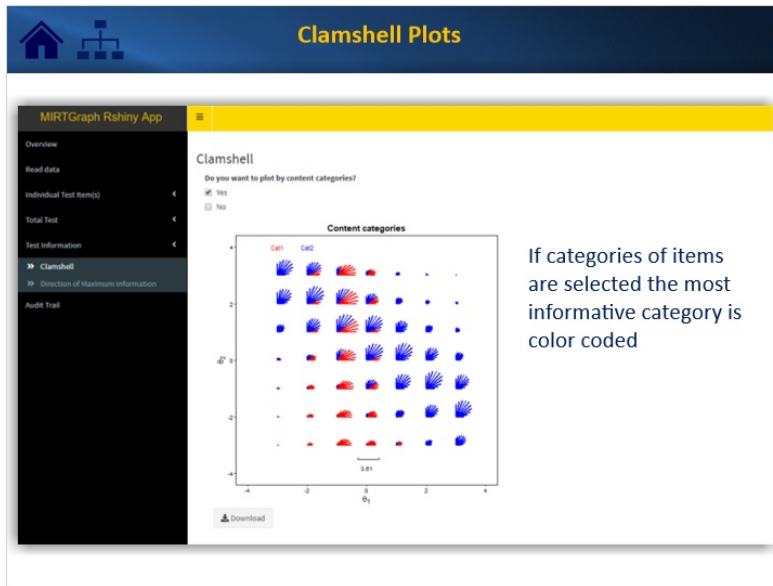
5.14 Centroid Plots



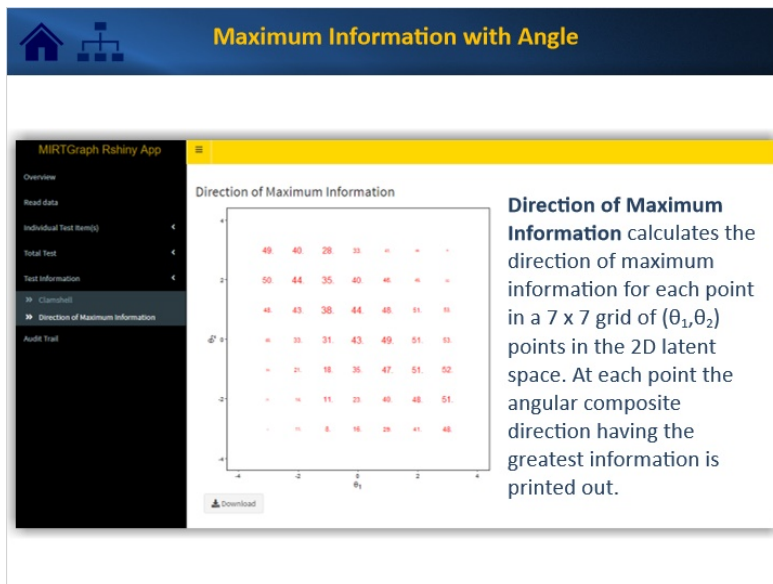
5.15 Clamshell Plots by Content Categories



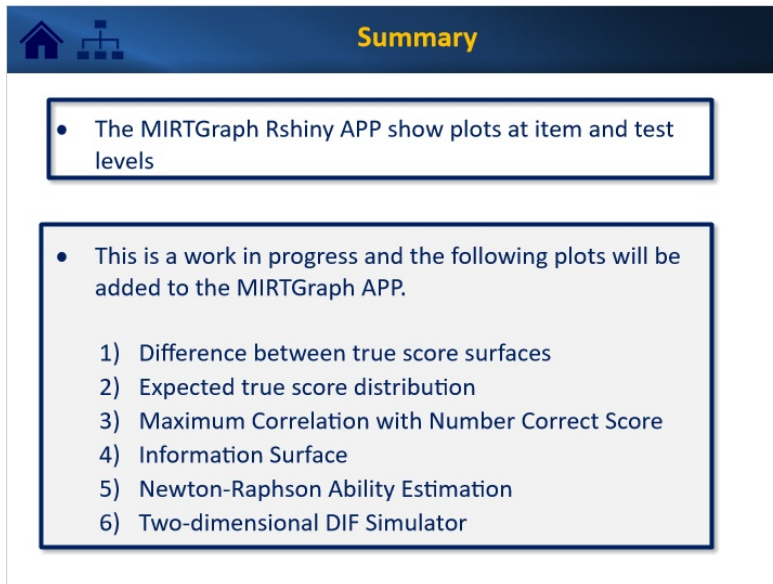
5.16 Clamshell Plots



5.17 Maximum Information with Angle



5.18 Summary

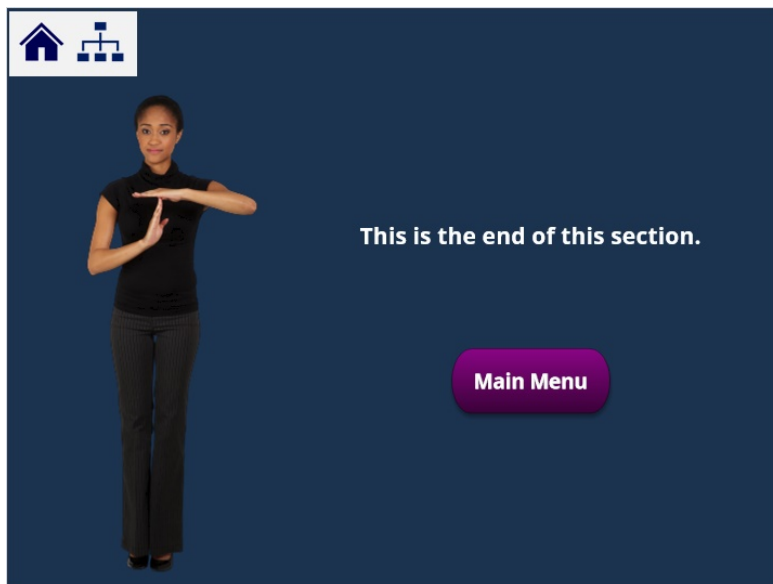


The slide features a dark blue header with a home icon and a tree diagram icon on the left, and the word "Summary" in yellow on the right. Below the header, there are two white boxes with blue borders containing text. The first box contains a single bullet point. The second box contains a bullet point followed by a numbered list of six items.

- The MIRTGraph Rshiny APP show plots at item and test levels

- This is a work in progress and the following plots will be added to the MIRTGraph APP.
 - 1) Difference between true score surfaces
 - 2) Expected true score distribution
 - 3) Maximum Correlation with Number Correct Score
 - 4) Information Surface
 - 5) Newton-Raphson Ability Estimation
 - 6) Two-dimensional DIF Simulator

5.19 Bookend: Plots for 2D Items and Tests



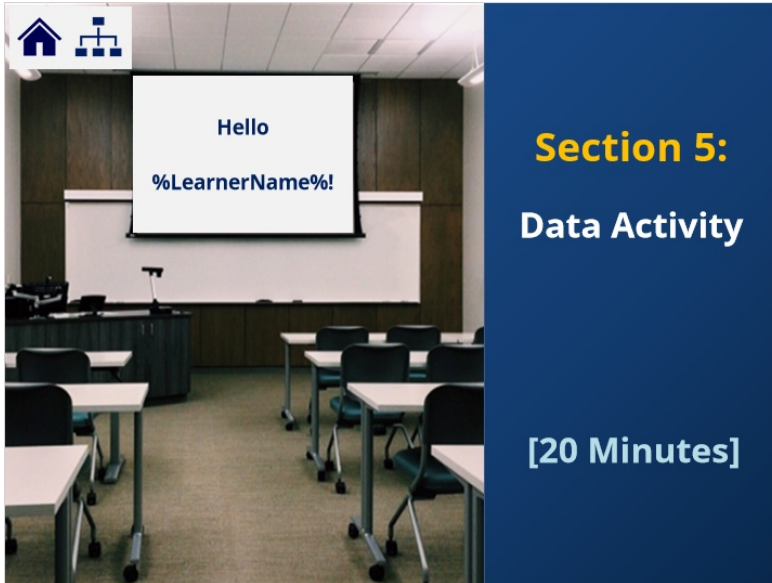
The slide features a dark blue background. In the top left corner, there is a white box containing a home icon and a tree diagram icon. On the left side, a woman in a black top and pants is performing a sign language gesture. To the right of the woman, the text "This is the end of this section." is displayed. Below the text, there is a purple rounded rectangular button with the text "Main Menu" in white.

This is the end of this section.

Main Menu

6. Data Activity

6.1 Cover: Data Activity



6.2 Access to RShiny

The slide has a dark blue header with a home icon and a tree icon on the left, and the text "Access to Rshiny" in yellow on the right. Below the header, the text reads: "The MIRTGraph Rshiny app can be accessed via the following link:" followed by a green box containing the URL <https://qing-xie-uiowa.shinyapps.io/MIRTGraph/> and the Shiny logo.

6.3 Data Activity using RShiny

Data Activity using RShiny

Given four items with parameters...

- a1 = 1.7, a2 = 0.4, d = -.6, Content = 1
- a1 = 0.4, a2 = 0.9, d = .8, Content = 2
- a1 = 0.9, a2 = 1.1, d = -.2, Content = 1
- a1 = 1.1, a2 = 0.5, d = 1.2, Content = 2

...create the following plots using MIRTGraph

- a) Item vector
- b) Test characteristic surface
- c) Centroid
- d) Clamshell
- e) Direction of maximum information plot

Solution **Section End**

6.4 Read in Dataset

Read in Dataset

MIRTGraph Rshiny App

Read in dataset

Uploading Files

Choose CSV File

Upload CSV File

Separator: Comma

Quote: Double Quote

	a1	a2	D	Cont
1	1.7	0.4	-0.6	1
2	0.4	0.9	0.8	2
3	0.9	1.1	-0.2	1
4	1.1	0.5	1.2	2

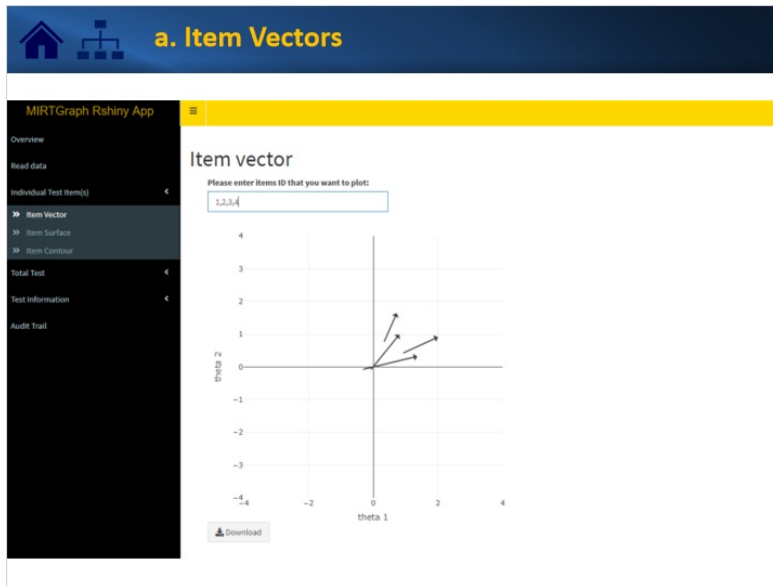
Showing 1 to 4 of 4 entries

Previous 1 Next

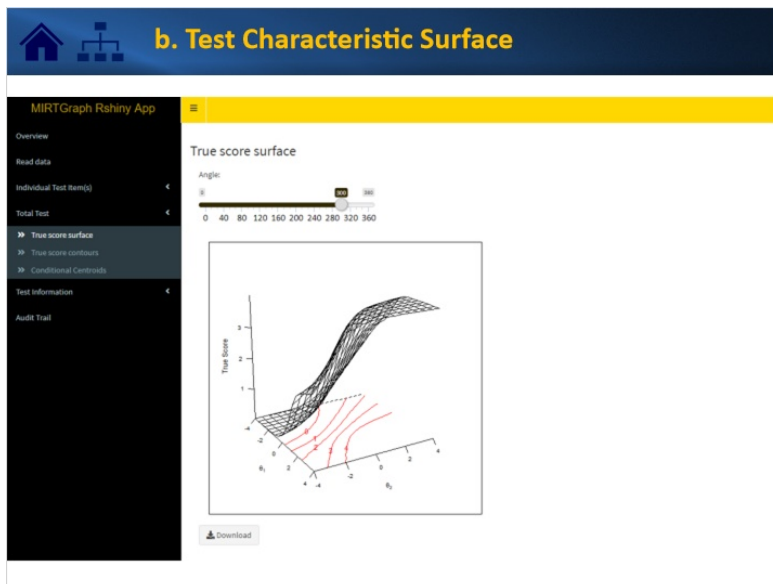
Your .csv file should look like this

	A	B	C	D
1	a1	a2	D	Cont
2	1.7	0.4	-0.6	1
3	0.4	0.9	0.8	2
4	0.9	1.1	-0.2	1
5	1.1	0.5	1.2	2
6				
7				

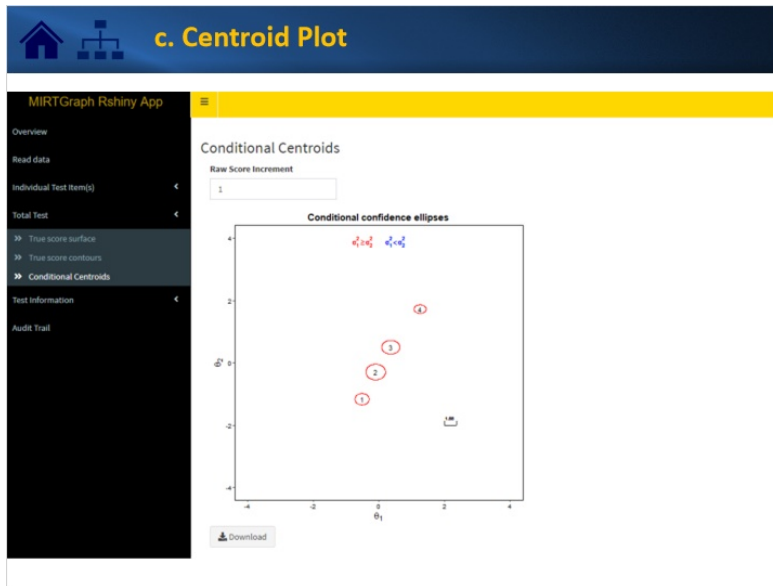
6.5 Item Vectors



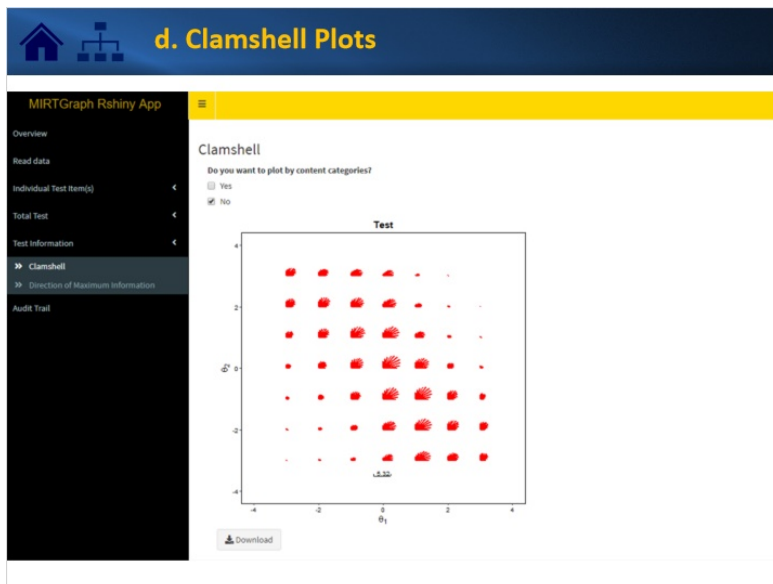
6.6 Test Characteristic Surface



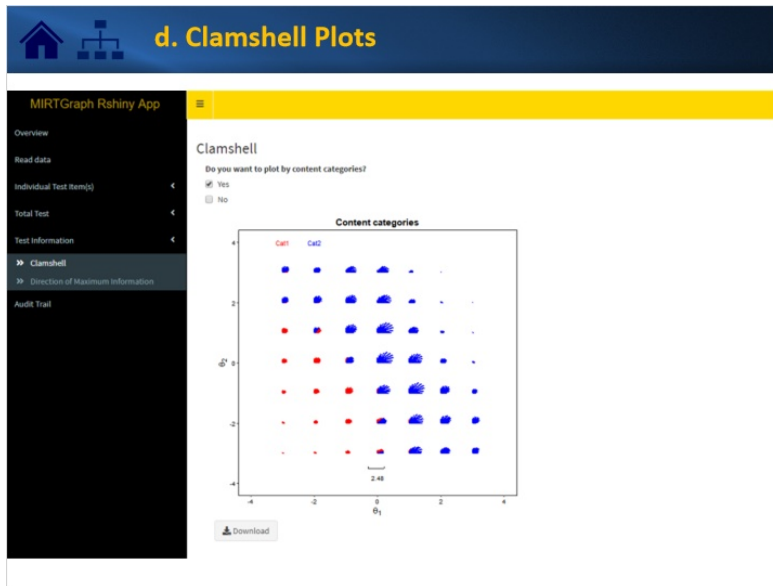
6.7 Centroid Plot



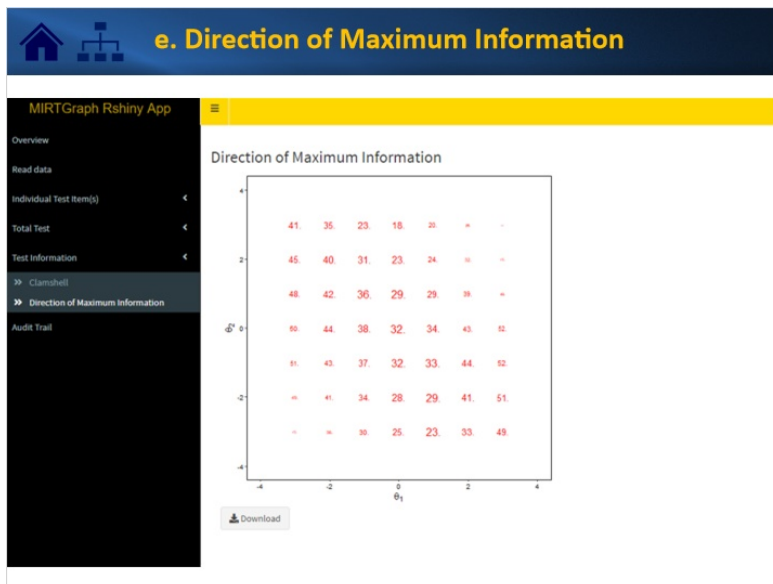
6.8 Clamshell Plots (I)



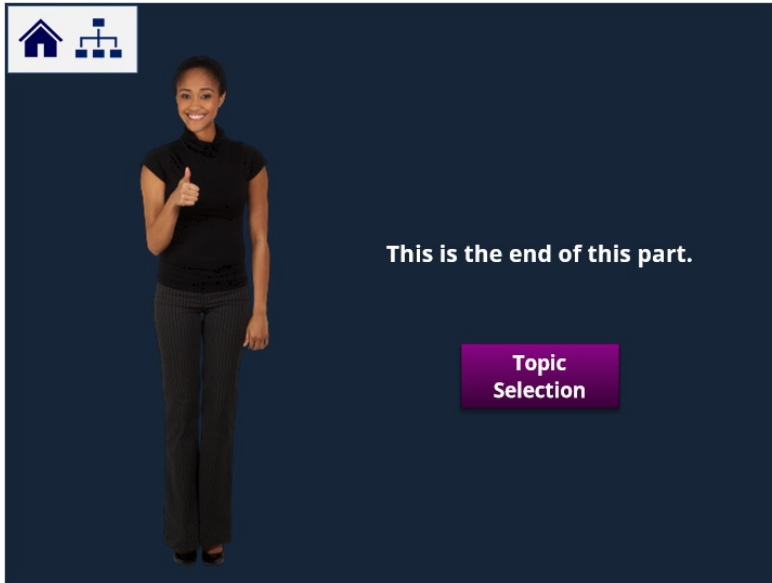
6.9 Clamshell Plots (II)



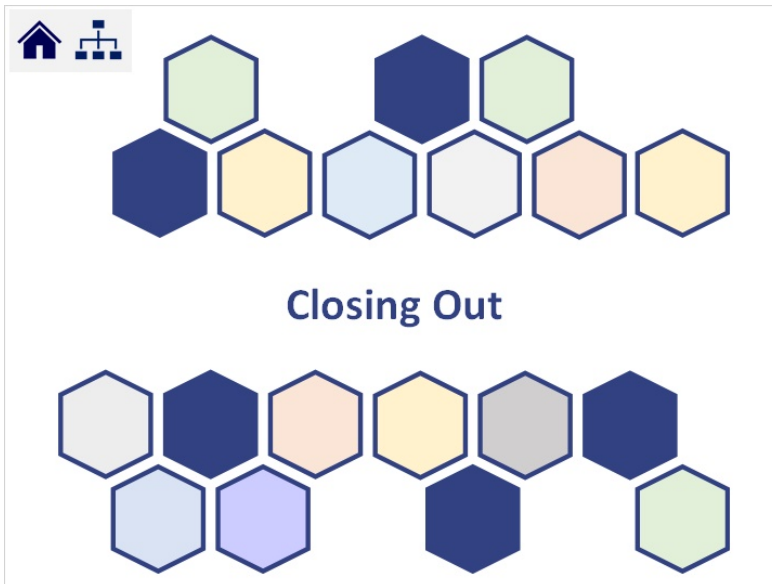
6.10 Direction of Maximum Information



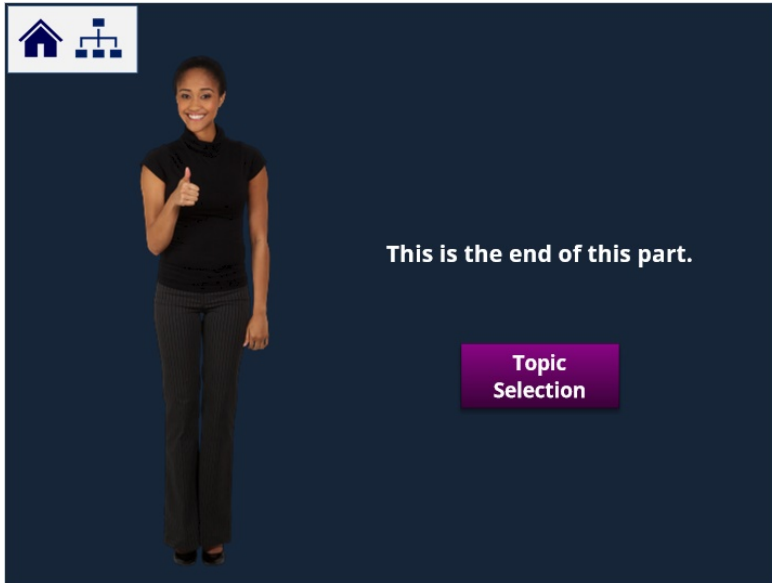
6.11 Bookend: Data Activity Solutions



6.12 End of Section

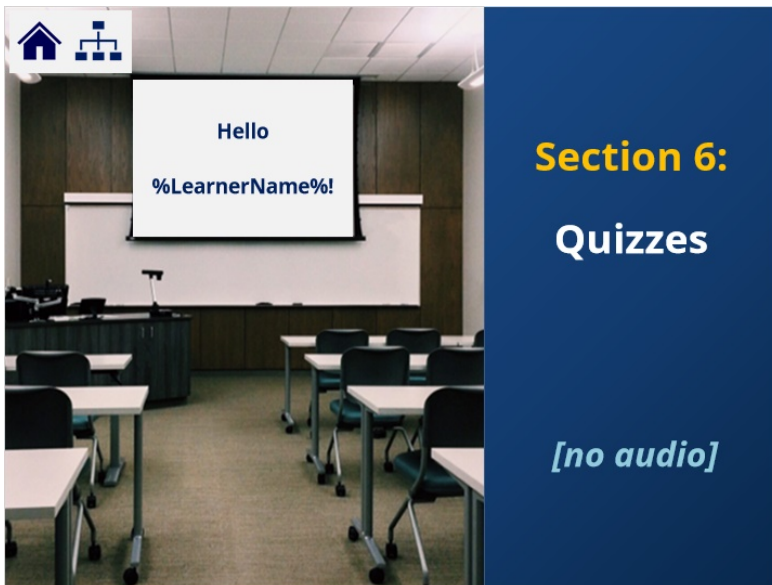


6.13 Bookend: Data Activity

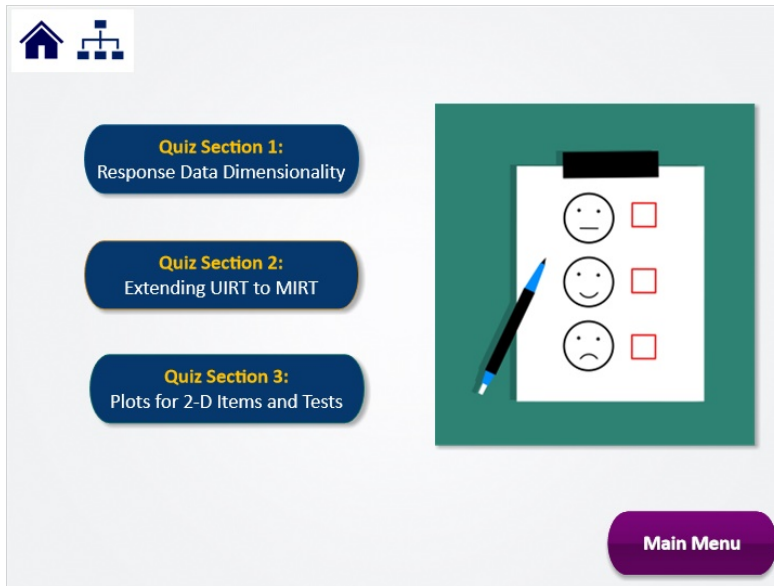


7. Quizzes

7.1 Cover: Quizzes



7.2 Quiz Selection



The interface features a navigation bar at the top left with a home icon and a tree structure icon. Below this, three blue buttons are stacked vertically, each with a yellow title and white text. To the right of these buttons is a green square containing a white clipboard with a black pen, three smiley faces, and three red checkboxes. At the bottom right, there is a purple button with white text.

Quiz Section 1:
Response Data Dimensionality

Quiz Section 2:
Extending UIRT to MIRT

Quiz Section 3:
Plots for 2-D Items and Tests

Main Menu