A Diagnostic Assessment of Teachers' Understandings of Rational Number

Laine Bradshaw

Assessment & Measurement, James Madison University

Jonathan Templin

Research, Evaluation, Measurement & Statistics, University of Georgia

Andrew Izsák

Mathematics Education, University of Georgia

Introduction

- *Diagnosing Teachers' Multiplicative Reasoning* (DTMR)*
  - » NSF funded grant (DRL-0903411)
- Goal is to create a test that will assess fine-grained components of teachers' reasoning multiplicatively with rational numbers
- The test would be used to
  - » Tailor professional development to teachers' needs
  - » Quantitatively study teachers' fine-grained abilities to reason multiplicatively
    - Quantify findings based on extensive qualitative research base
    - Generalize to larger populations

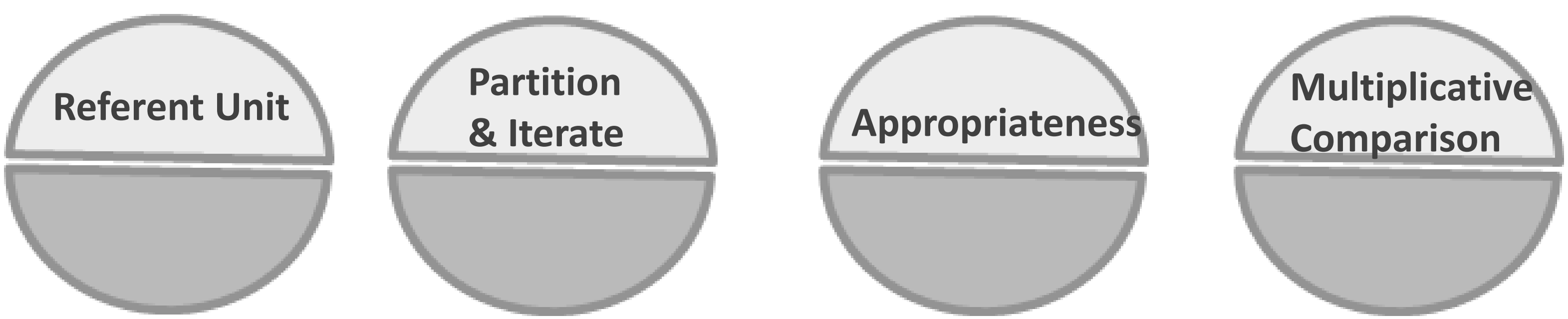
*Izsák, A., Lobato, J., Druken, B., Orrill, C., Jacobson, E., & Bradshaw, L. (2010, July). *Applying cognitive diagnosis models to measure middle grades teachers' multiplicative reasoning.* Paper presented at the annual International Meeting of the Psychometric Society in Athens, GA.

Big Picture

- Most psychometric models are designed to measure a unidimensional continuous trait or ability
- Examples of continuous traits
  - » Student's "math" ability at the 8^{th} grade level
  - » In-service teachers' mathematical knowledge for teaching number and operations
    - ▪ The content area of focus for this study
- As a result, many tests are designed to measure a unidimensional ability
- This project took a different approach
  - » A multidimensional diagnostic approach

Diagnosing Multiplicative Reasoning

- Instead of measuring an overall ability to reason multiplicatively with fractions, we can break that continuous trait down into more fine-grained cognitive facilities or *attributes*:
  - » Ability to identify appropriate **referent units** for numbers
  - » Ability to **partition** quantities and **iterate** unit fractions
  - » Ability to identify **appropriate** arithmetic operations
  - » Ability to make **multiplicative comparisons**
- We treat these attributes as categorical
  - » Dichotomous (have two categories)
  - » Mastery of an attribute (= 1) or non-mastery of an attribute (= 0)



Diagnosis from a Psychometric Model

- Diagnostic classification models (DCMs) are a set of statistical tools that provide diagnostic feedback
- DCMs are well-aligned with educational assessment needs
  - » We are trying to make decisions about examinees
- A diagnosis is a decision
  - » Is a student a master of a given standard?
  - » Does a teacher need professional development on a given concept?
- DCMs provide diagnoses by directly classifying examinees into groups according to categorical latent traits
  - » Other psychometric model families rank-order examinees on continuous traits

Groups According to Attribute Mastery

- The groups are based on patterns of mastery according to the set of attributes
- A classification of each individual skill results in a classification into one of these 16 patterns

2^A possible patterns or groups:

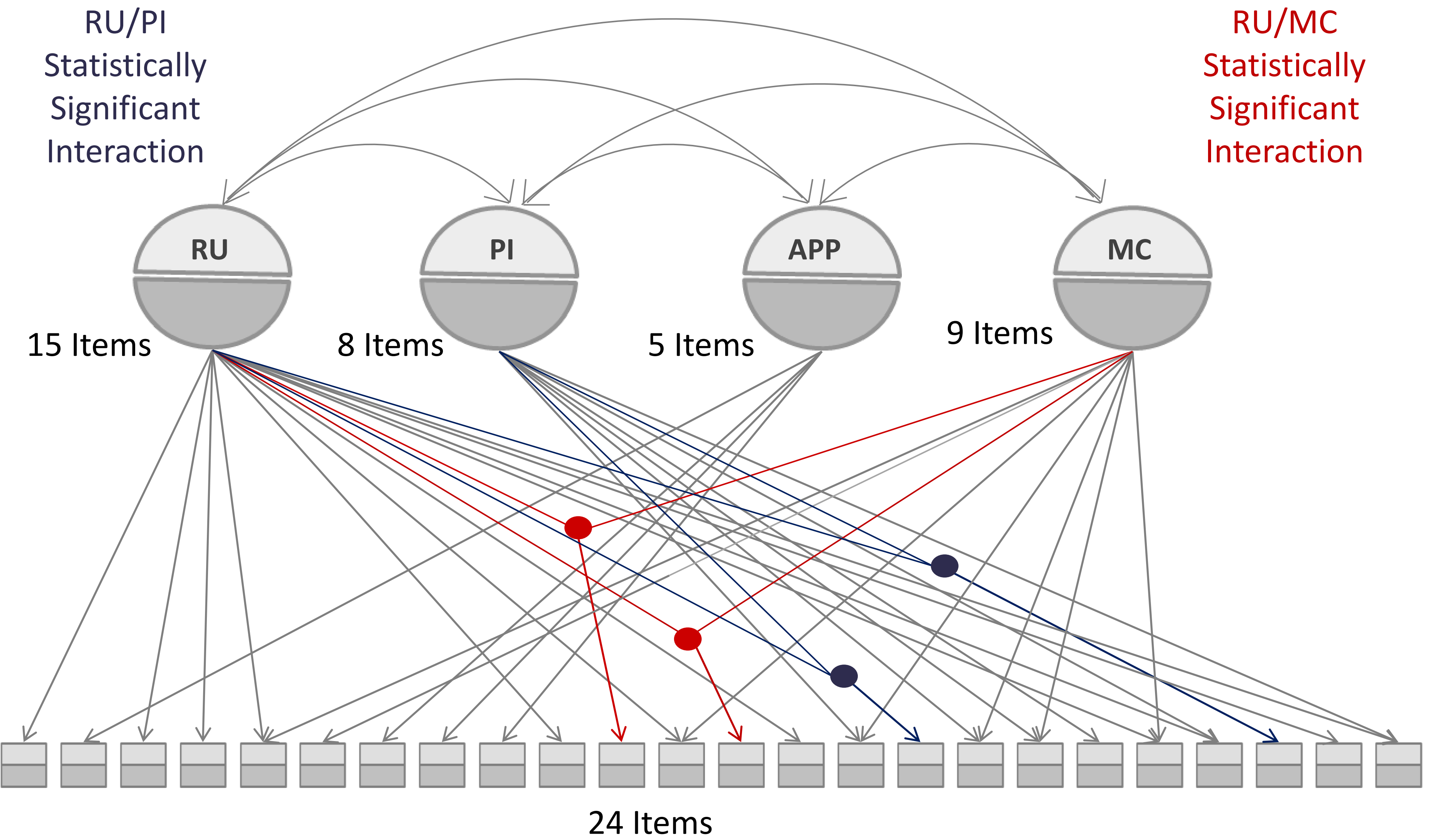
Pattern	RU	PI	APP	MC
1	0	0	0	0
2	0	0	0	1
3	0	0	1	0
4	0	0	1	1
5	0	1	0	0
6	0	1	0	1
7	0	1	1	0
8	0	1	1	1
9	1	0	0	0
10	1	0	0	1
11	1	0	1	0
12	1	0	1	1
13	1	1	0	0
14	1	1	0	1
15	1	1	1	0
16	1	1	1	1

Designing Diagnostic Tests

- Diagnostic tests are written so that each item measures one or more of the attributes
- The attributes measured by each item are recorded in a **Q-matrix**
  - » Describes whether an item measures an attribute (q = 1) or not (q = 0)
  - » Mapping is established by content experts
    - Confirmed by item response interviews
- First several items on DTMR test:

	RU	PI	APP	MC
Item 1	1	0	0	0
Item 2	0	0	1	0
Item 3	1	0	0	0
Item 4	1	0	0	1

A Model of the DTMR Diagnostic Test



Log-linear
Cognitive Diagnosis Model

Log-linear Cognitive Diagnosis Model (LCDM)

- The **L**og-linear **C**ognitive **D**iagnosis **M**odel* (LCDM)
  - Parameterizes DCMs using a linear model framework
  - Can be compensatory or non-compensatory at the item level
  - Can be estimated using Mplus
- The item response is predicted as a function of the set of attributes that is measured by that item
  - Mastering additional attributes should increase the probability of answering the item correctly
- Attributes are categorical latent variables
  - Linear predictor is like ANOVA
- Responses are (typically) binary: correct or incorrect
  - Logit link function like logistic regression

*Henson, R., Templin, J., & Willse, J. (2009). Defining a family of cognitive diagnosis models using log linear models with latent variables. *Psychometrika, 74*, 191-210.

Notation

e	**Examinee**
i	**Item**
α (Alpha)	**Attribute (Categorical Latent Trait)**
λ (Lambda)	**Loading (Coefficient)**

Example Item

- This item is analogous to Item 22 on the DTMR test
  - » Measures Referent Unit (Attribute 1) and Partitioning and Iterating (Attribute 2)

Ms. Roland gave her students the following problem to solve:

> *Candice has 4/5 of a meter of cloth. She uses 1/8 of a meter for a project. How much cloth does she have left after the project?*

She had students use the number line so that they could draw the lengths. Which of the following diagrams shows the solution? Assume all intervals are subdivided equally.

a)

b)

c)

d)

e)

LCDM Example Item Response Function

- Referent unit (α_1) and partitioning and iterating (α_2) are measured
  - Q-matrix entries:

	RU	PI	APP	MC
Item 22	1	1	0	0

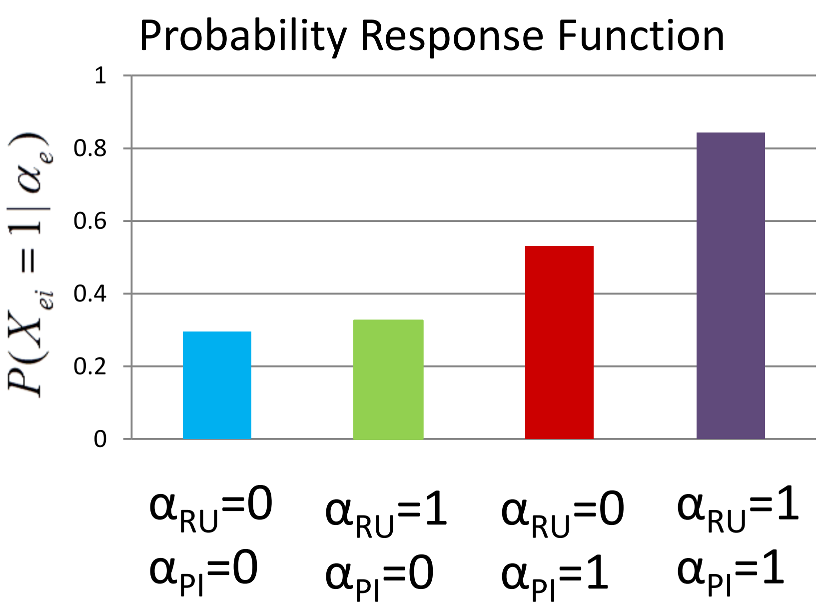
  - LCDM item response function:

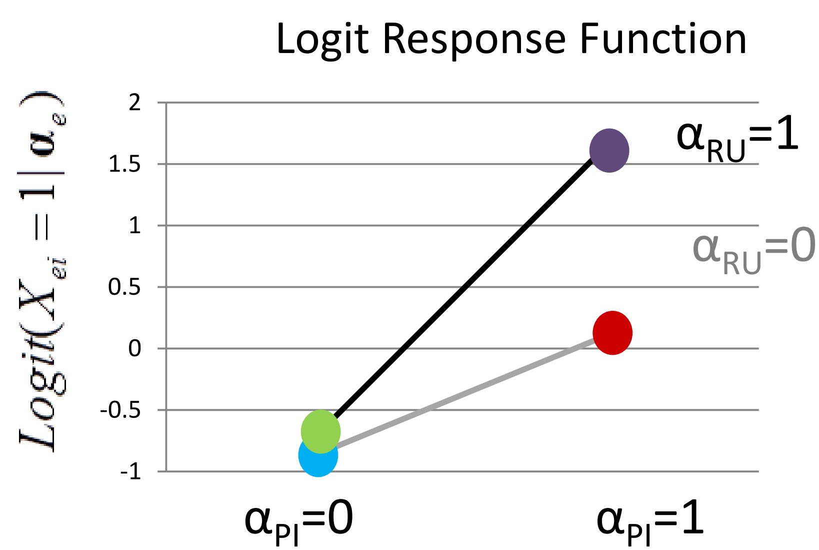
$$\log \frac{P(X_{ei}=1|\alpha_e)}{P(X_{ei}=0|\alpha_e)} = \underbrace{\boxed{\lambda_{i,0}}}_{\text{Intercept (Guessing)}} + \underbrace{\boxed{\lambda_{i,1(1)}}}_{\text{Main Effect (RU)}}(\alpha_{e1}) + \underbrace{\boxed{\lambda_{i,1(2)}}}_{\text{Main Effect (PI)}}(\alpha_{e2}) + \underbrace{\boxed{\lambda_{i,2(12)}}}_{\text{Interaction (Between RU and PI)}}(\alpha_{e1} \cdot \alpha_{e2})$$

LCDM Example Item Response Function

$$\log \frac{P(X_{ei}=1|\alpha_e)}{P(X_{ei}=0|\alpha_e)} = -.871+.146(\alpha_{e1})+.991(\alpha_{e2})+1.415(\alpha_{e1}\cdot\alpha_{e2})$$

- On the logit scale, we can see the main effects are positive and the interaction is positive (similar to ANOVA methods)
- Item parameters provide construct validation
  - Is the item actually measuring the attribute?





DTMR Preliminary Results

Results Overview

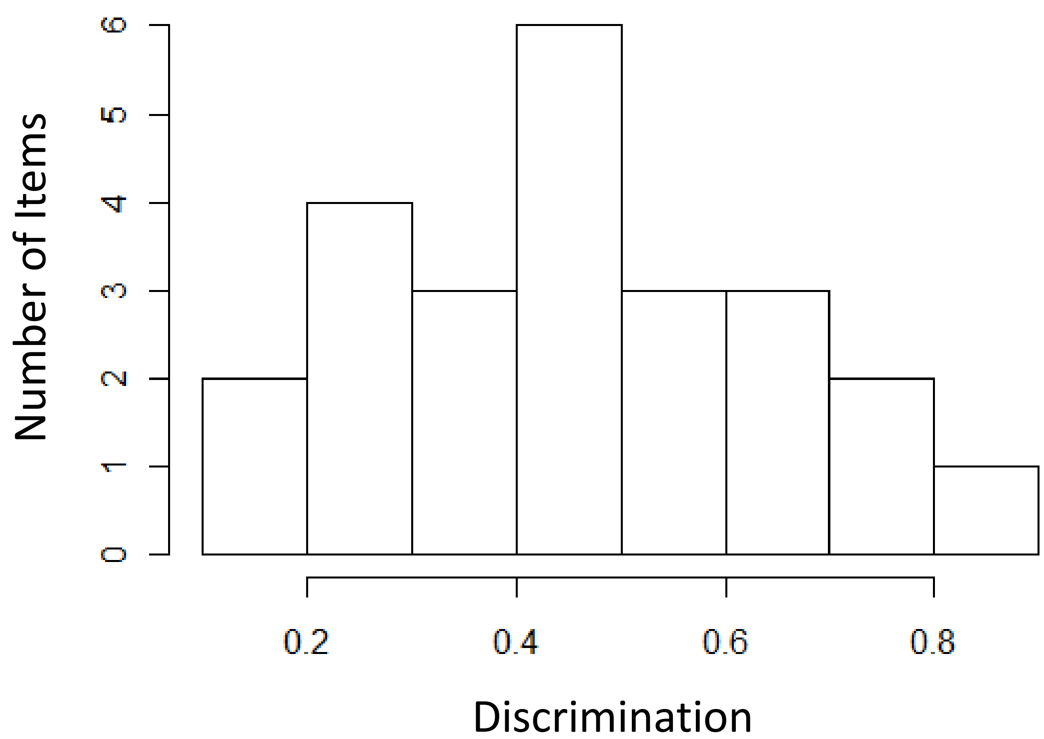
- Data Collection
- Items
  - » How well did they function?
- Attribute Patterns
  - » How many teachers are masters of each attribute?
  - » What are the attribute mastery probabilities for a single teacher?
- Attribute Correlations
  - » How highly correlated are the attributes?
  - » Are any attributes dependent on another?

Data Collection

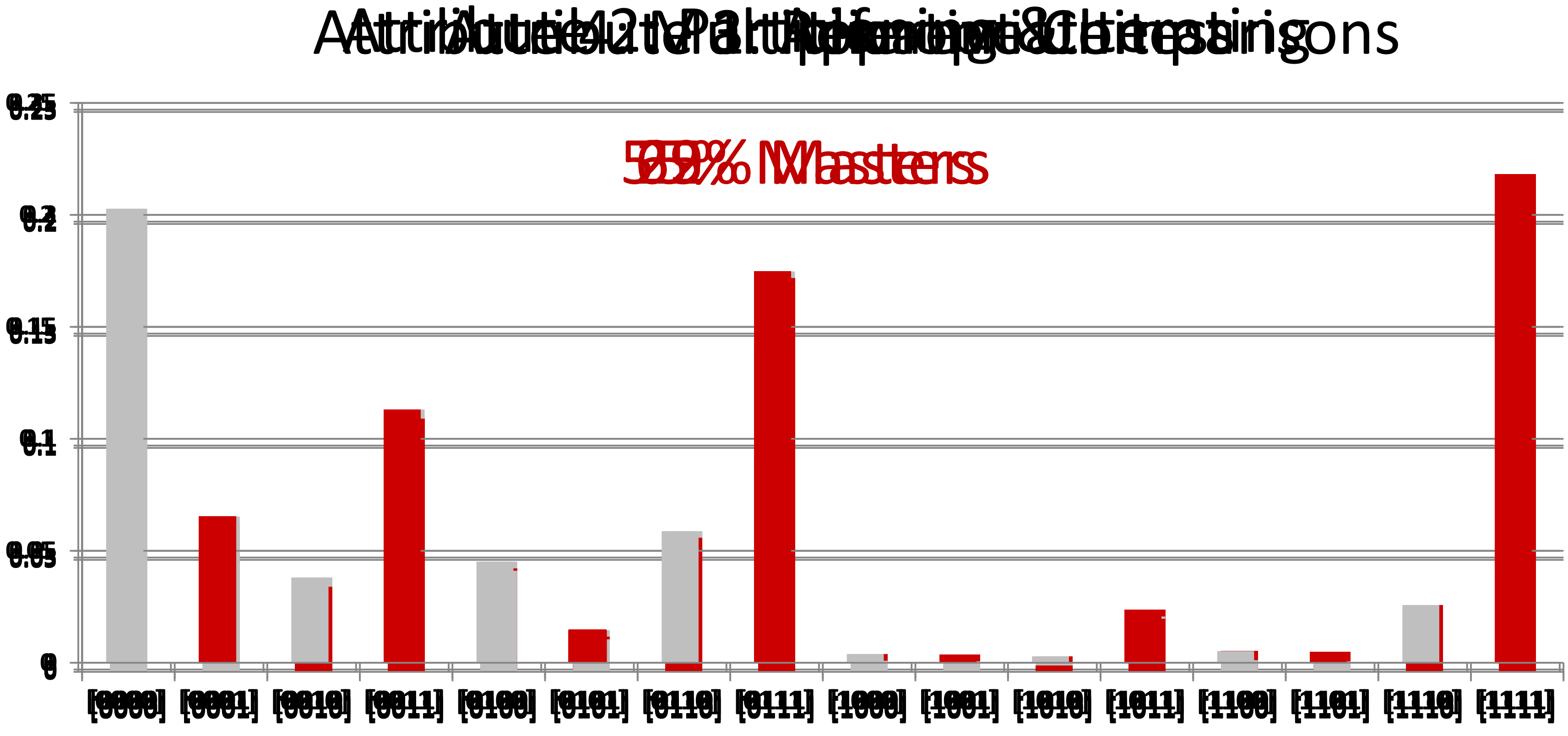
- National sample of 692 in-service middle grades mathematics teachers
- Sample stratified by
  - » Region of the country (4 levels)
    - ▪ Northeast, Midwest, South, West
  - » Urban-centric locale (12 levels)
    - ▪ City or suburb
      - ♦ Small, medium, large
    - ▪ Town or rural
      - ♦ Fringe, distant, remote
- Response rate: ≈20%
  - » Received 692 of 5400 teachers (so far)

Item Discrimination

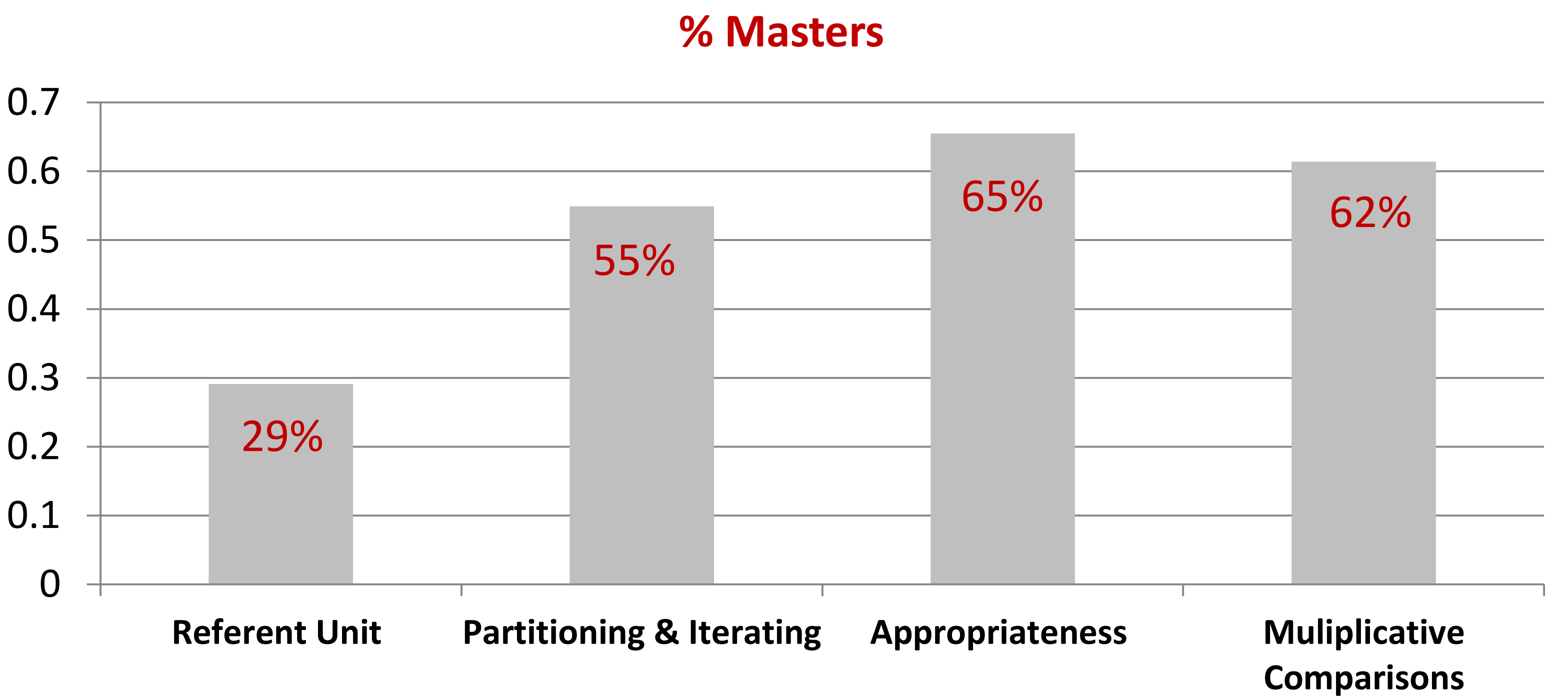
- For "good" items, masters of the attribute(s) answer the item correctly and non-masters answer the item incorrectly
  - » This would yield high discrimination, or differences in the probability masters and non-masters answer the item correctly



Attribute Patterns of Mastery



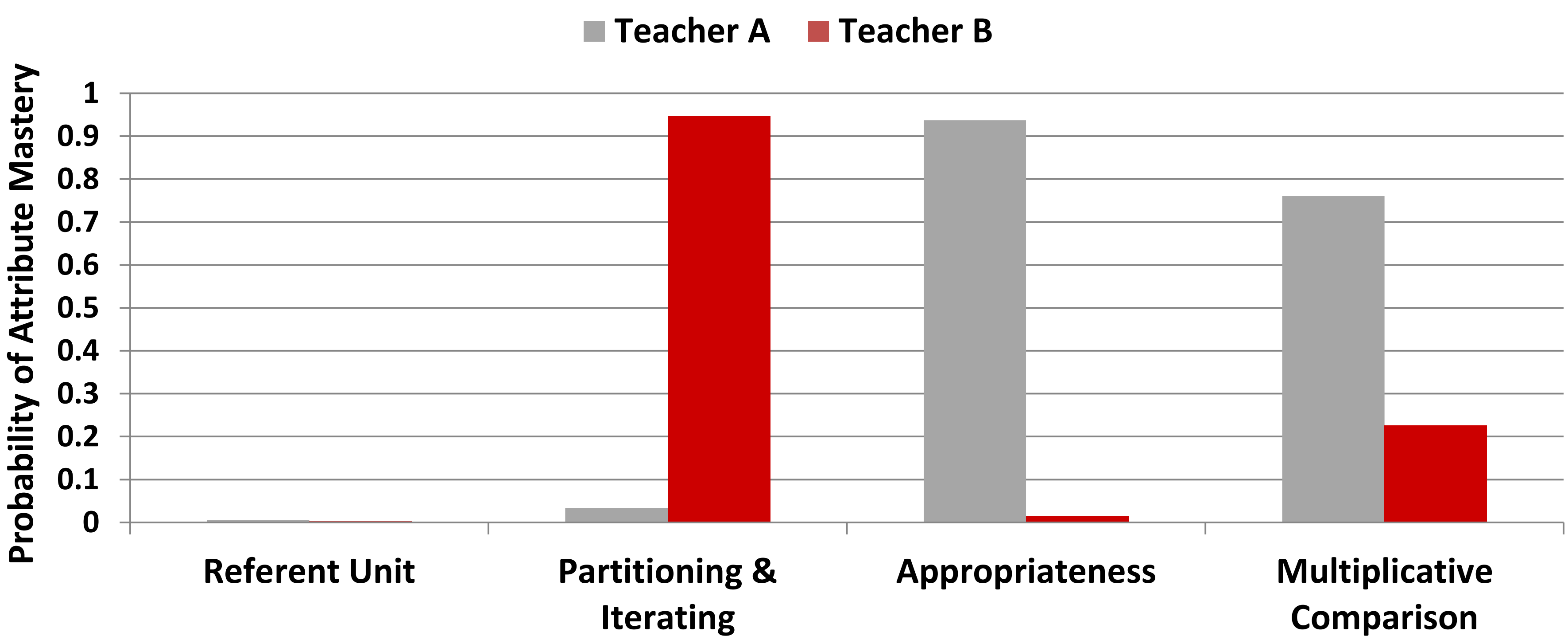
Individual Attribute Mastery



- Information useful for
  - Tailoring professional development
    - Many teachers may benefit from professional development on referent unit
  - Understanding base-rates of attribute mastery in the population of in-service teachers
    - Quantitative Research

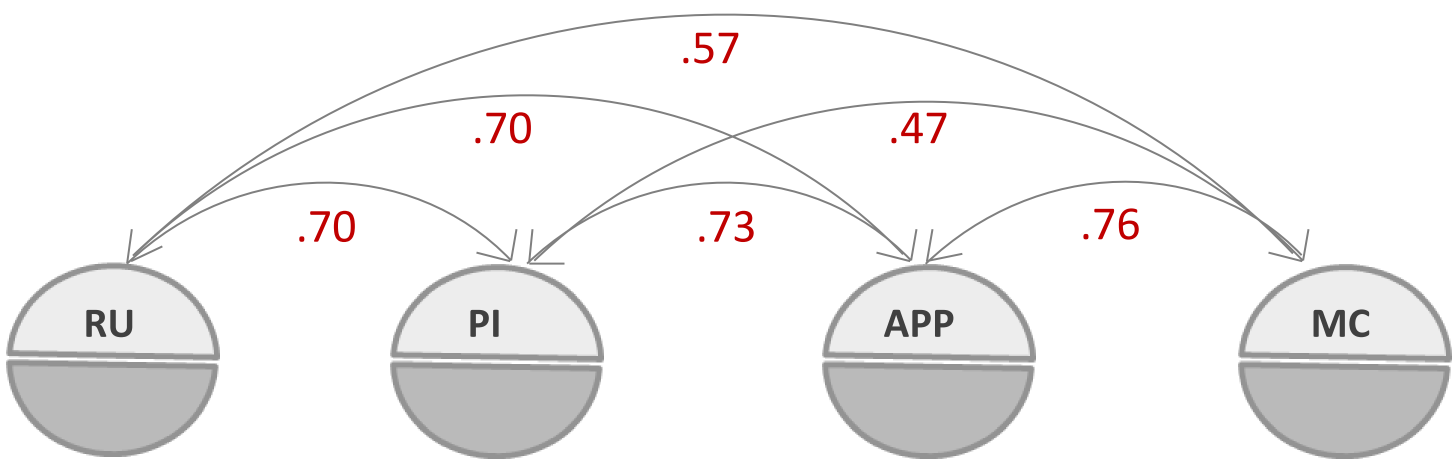
Teacher-level Individual Attribute Feedback

- Comparison of total scores and DCM diagnosis:
  - » Teacher A and Teacher B both answered 8 out of 24 items correctly
  - » Teacher A has attribute pattern [0100]
  - » Teacher B has attribute pattern [0011]
    - Need different types of professional development



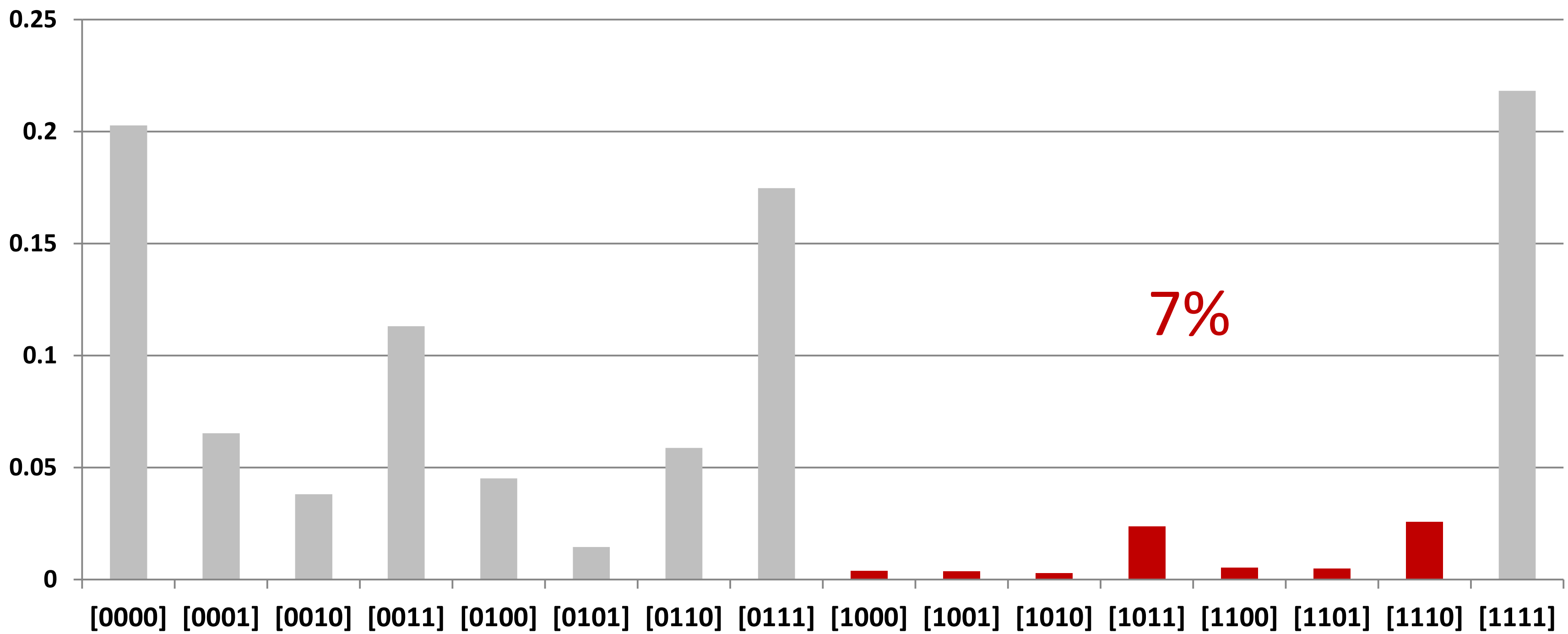
Attribute Correlations

- The attribute patterns are reflections of the correlations among the latent variables
  - » Tetrachoric correlations (between categorical variables)
  - » The relationships among the attributes are parameterized through a log-linear structural model



Attribute Patterns of Mastery

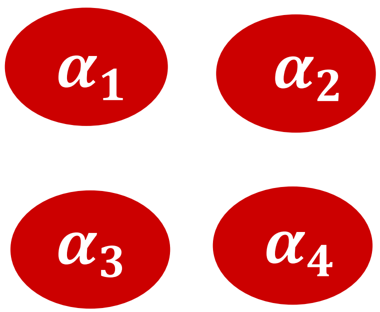
- Observed patterns make you wonder if Attribute 1, Referent Unit, is dependent upon other attributes
  - 7% of students mastered Referent Unit without being a master of all other 3
  - 3.4% mastered Referent Unit (α_1) without Partitioning & Iterating (α_2)
  - 1.7% mastered Referent Unit (α_1) without Appropriateness (α_3)
  - 3.8% mastered Referent Unit (α_1) without Multiplicative Comparison (α_4)



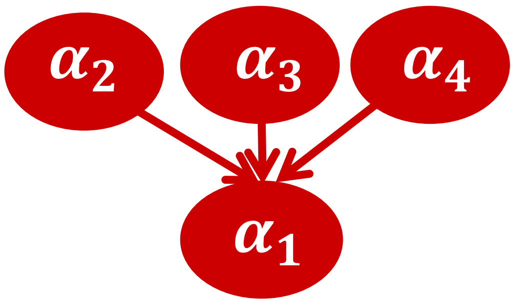
Attribute Hierarchies

- We tested the following hierarchies using the Hierarchical Diagnostic Classification Model* (HDCM)
  - All hierarchies fit significantly worse (p<.001) than the no hierarchy

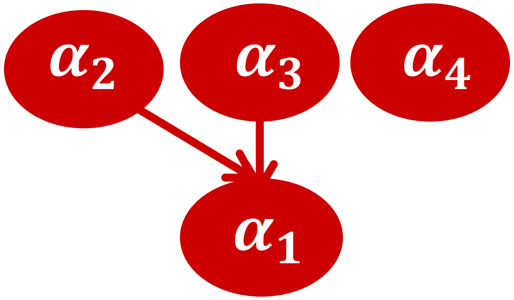
No Hierarchy



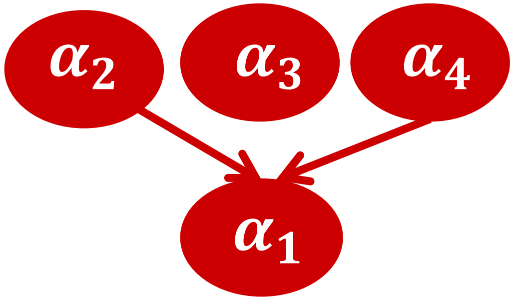
Hierarchy 1



Hierarchy 2

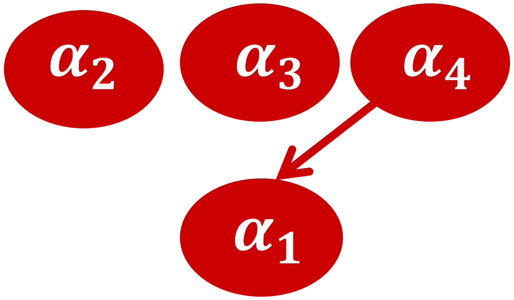


Hierarchy 3



…

Hierarchy 7



*Templin, J., & Bradshaw, L. (2011). *Hierarchical diagnostic classification models: A family of models for estimating and testing attribute hierarchies.* Manuscript under review.

Not Just New to Mathematics Education…

- This presentation has focused on the contributions of this project to mathematics education research base and use in mathematics education
- This project is also one of the first efforts to **prospectively** diagnose attributes
  - » Further unique in that the attributes are cognitive in nature and very fine-grained
  - » Helpful to have a model of how to do this in practice for the field of psychometrics
    - Especially since "it" worked!

Thank you!

If you have questions or comments,
please feel free to email me:

bradshlp@jmu.edu