On the Robustness of the One-parameter Log-linear Cognitive Diagnosis Model

Matthew J. Madison, Ashley Heady,
Beltran Pantoja, Nancy Alila, Oluwatosin Adeleye
University of Georgia

Access slides and paper at www.matthewmadison.com/research

2026 Meeting of the National Council on Measurement in Education

Talk Overview

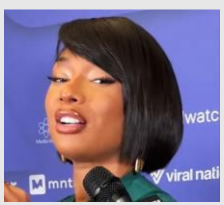
- Background
  - » Introduction to DCMs and the 1-PLCDM
  - » DCM empirical blueprints
  - » Purpose of this study

- Simulation study
  - » Design
  - » Results

- Conclusions

Diagnostic Classification Models

- DCMs use item responses to place students into groups according to proficiency or non-proficiency of attributes

Student	Addition	Subtraction	Multiplication	Division
	✓	—	✓	—
	✓	✓	✓	✓

- Defining features: multidimensionality and diagnostic interpretations

General and Constrained DCMs

- General model: log-linear cognitive diagnosis model (LCDM)
  - » Subsumes many other DCMs (e.g., DINA, CRUM)
  - » Flexible parametrization
  - » Allows for top-down approach to model building

- Constrained model: one-parameter LCDM (1-PLCDM)
  - » Special case of LCDM where attribute main effects are constrained
  - » Analogous to 1-PL IRT model
  - » Nice measurement properties (sufficiency, invariant item ordering)

A One-Parameter DCM

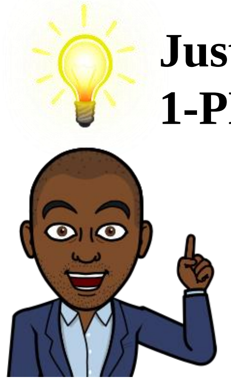
- One-parameter log-linear cognitive diagnosis model (***1-PLCDM***)
  - Combines a general DCM (LCDM) and the 1-PL IRT model

- Logit response function for the LCDM

$$logit(X_i = 1) = \boxed{\lambda_{0,i}} + \boxed{\lambda_{1,i}} \cdot \alpha$$

- Logit response function for the 1-PLCDM

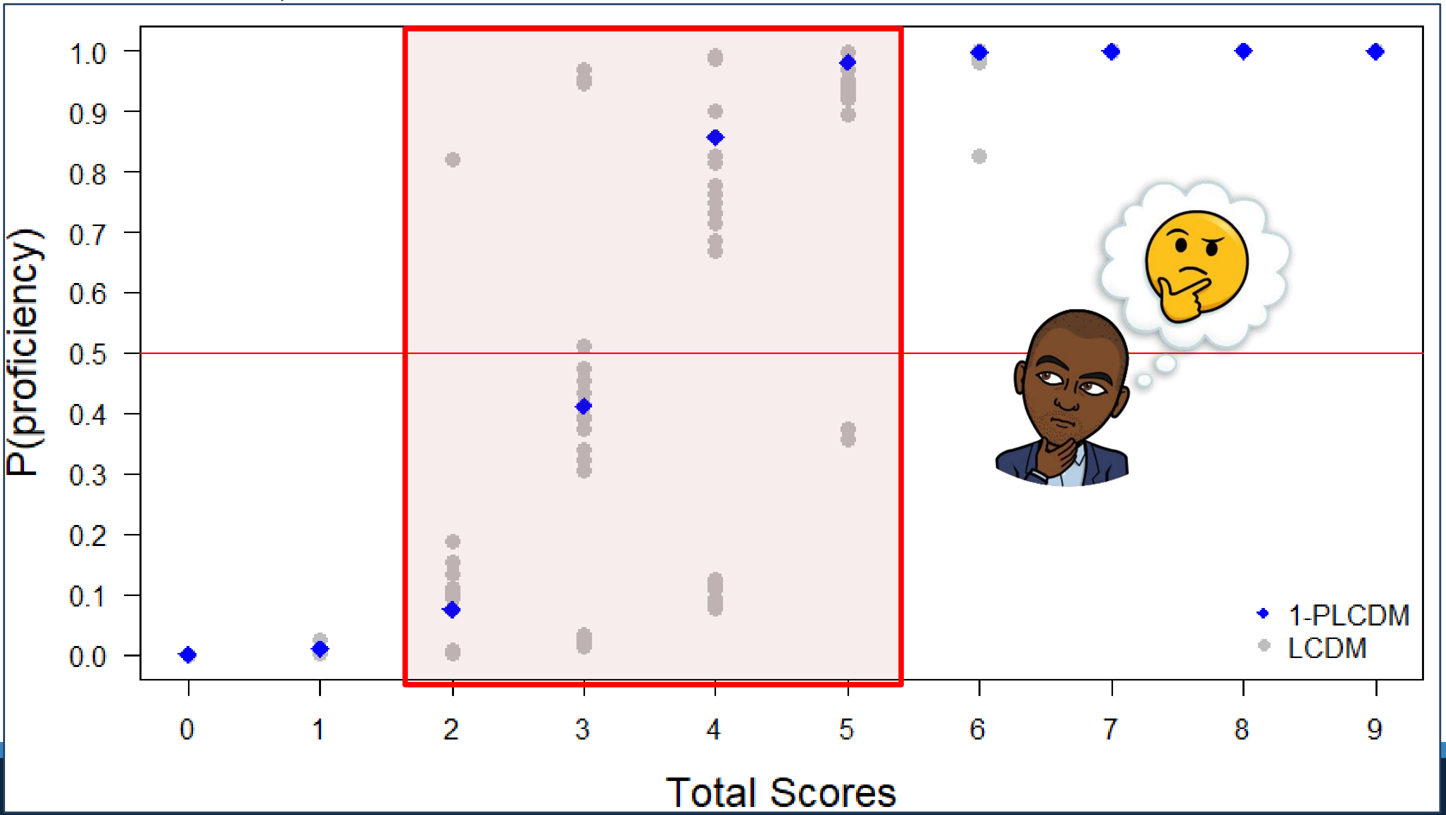
$$logit(X_i = 1) = \boxed{\lambda_{0,i}} + \boxed{\lambda_{1} \cdot} \alpha$$



Just like the 1-PL!

Attribute Sum Score Sufficiency

- Translated for DCMs: the attribute sum score is sufficient for posterior probabilities of proficiency (and resulting classifications)



Madison, et al. (NCME, 2026)

General and Constrained DCMs

- General model: log-linear cognitive diagnosis model (LCDM)
  - » Subsumes many other DCMs (e.g., DINA, CRUM)
  - » Flexible parametrization
  - » Allows for top-down approach to model building

- Constrained model: one-parameter LCDM (1-PLCDM)
  - » Special case of LCDM where attribute main effects are constrained
  - » Analogous to 1-PL IRT model
  - » Nice measurement properties (sufficiency, invariant item ordering)

Test Blueprints

- Test blueprints guide test development efforts
  - » Indicate content domains, topic weights, rigor, etc.

Table 1. EOC Mathematics domain weight distributions.

Domain	NC Math 1
Number and Quantity and Algebra	36 - 40%
Functions	32 - 36%
Geometry	8 - 12%
Statistics and Probability	18 - 20%
Total	100%

Test Blueprints

- Test blueprints guide test development efforts
  - » Indicate content domains, topic weights, rigor, etc.

TCAP Assessment Blueprint

Math – Grade 4

This blueprint describes the content and structure of an assessment and defines the ideal percentage and number of operational test items by reporting category for the Tennessee Comprehensive Assessment Program (TCAP). For more information on this assessment, please see the Assessment Overviews here.

Computation with Whole Numbers
27 – 33% of Items

Standard
4.OA.A
4.NBT.B

Fractions
30 – 38% of Items

Standard
4.NF.A
4.NF.B
4.NF.C

Number Relationships and Patterns
17 – 20% of Items

Standard
4.OA.B
4.OA.C
4.NBT.A

Geometric and Measurement Concepts
15 – 23% of Items

Standard
4.MD.A
4.MD.B
4.MD.C
4.G.A

Test Blueprints

- Test blueprints guide test development efforts
  - » Indicate content domains, topic weights, rigor, etc.

Geometric measurement: understand concepts of angle and measure angles.

5. Recognize angles as geometric shapes that are formed wherever two rays share a common endpoint, and understand concepts of angle measurement:
   a. An angle is measured with reference to a circle with its center at the common endpoint of the rays, by considering the fraction of the circular arc between the points where the two rays intersect the circle.

   b. An angle that turns through n one-degree angles is said to have an angle measure of n degrees.

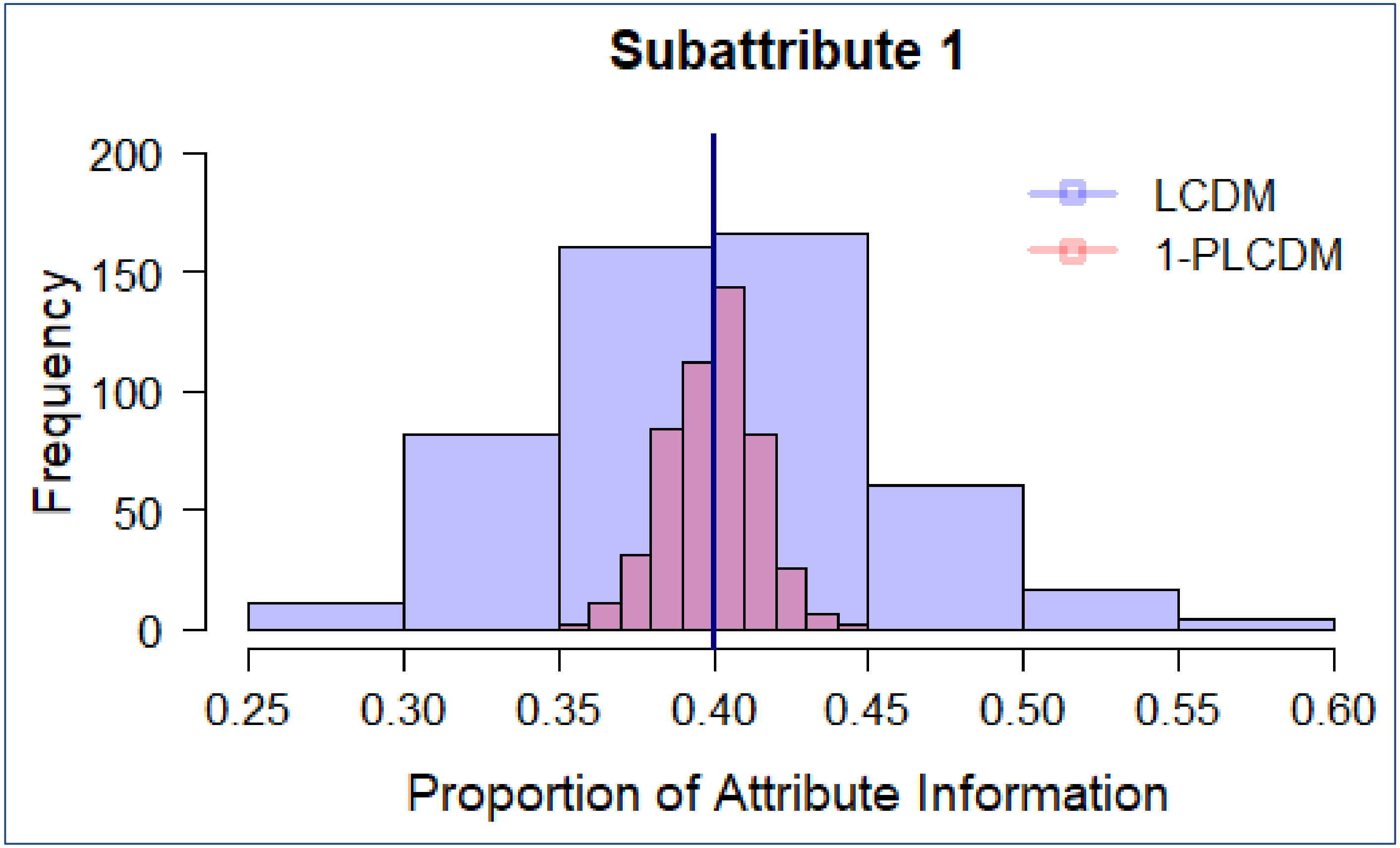
Test Blueprints

- Test blueprints guide test development efforts
  - » Indicate content domains, topic weights, rigor, etc.

- For test score validity, the empirical blueprint must match the prespecified blueprint

- Proportional item allocation is a standard practice
  - » If fractions has a prespecified weight of 40%, then it will get 8/20 items.
- For DCMs, proportional item allocation does not work
  - » Jurich and Madison (2023) – proportion of attribute information
  - » Madison and Alila (2025) – does work for the 1-PLCDM

Test Blueprints



General and Constrained DCMs

- General model: log-linear cognitive diagnosis model (LCDM)
  - » Subsumes many other DCMs (e.g., DINA, CRUM)
  - » Flexible parametrization
  - » Allows for top-down approach to model building

- Constrained model: one-parameter LCDM (1-PLCDM)
  - » Special case of LCDM where attribute main effects are constrained
  - » Analogous to 1-PL IRT model
  - » Nice measurement properties (sufficiency, invariant item ordering)
  - » **<u>Blueprint matching guaranteed</u>***

Purpose of Study

- In order to achieve these properties and blueprint matching, the 1-LCDM makes strong assumptions

1. Main effects are constrained to be equal
2. Attributes are assumed independent (correlation = 0)
3. Q-matrix must be simple structure (no complex items)

- **Purpose of this study is to examine the ability of the 1-PLCDM to match blueprints when these assumptions are violated**

Simulation Study Design

- Four attributes, 24 items
- Attribute 1 had two sub-attributes with weights of .40 and .60
  - Allocated 2/5 and 3/5 items
- Attribute 2 had two sub-attributes with weights of .60 and .40
  - Allocated 3/5 and 2/5 items
- Attribute 3 had two sub-attributes with weights of .33 and .66
  - Allocated 2/6 and 4/6 items
- Attribute 4 had three sub-attributes with weights of .25, .25, and .50
  - Allocated 2/8, 2/8, and 4/8 items

- Blueprint error quantified as the sum of absolute deviations

$$\sum_{a=1}^{4}\sum_{s=1}^{S}\left|preblue_{a_s} - empblue_{a_s}\right|$$

Simulation Study Design

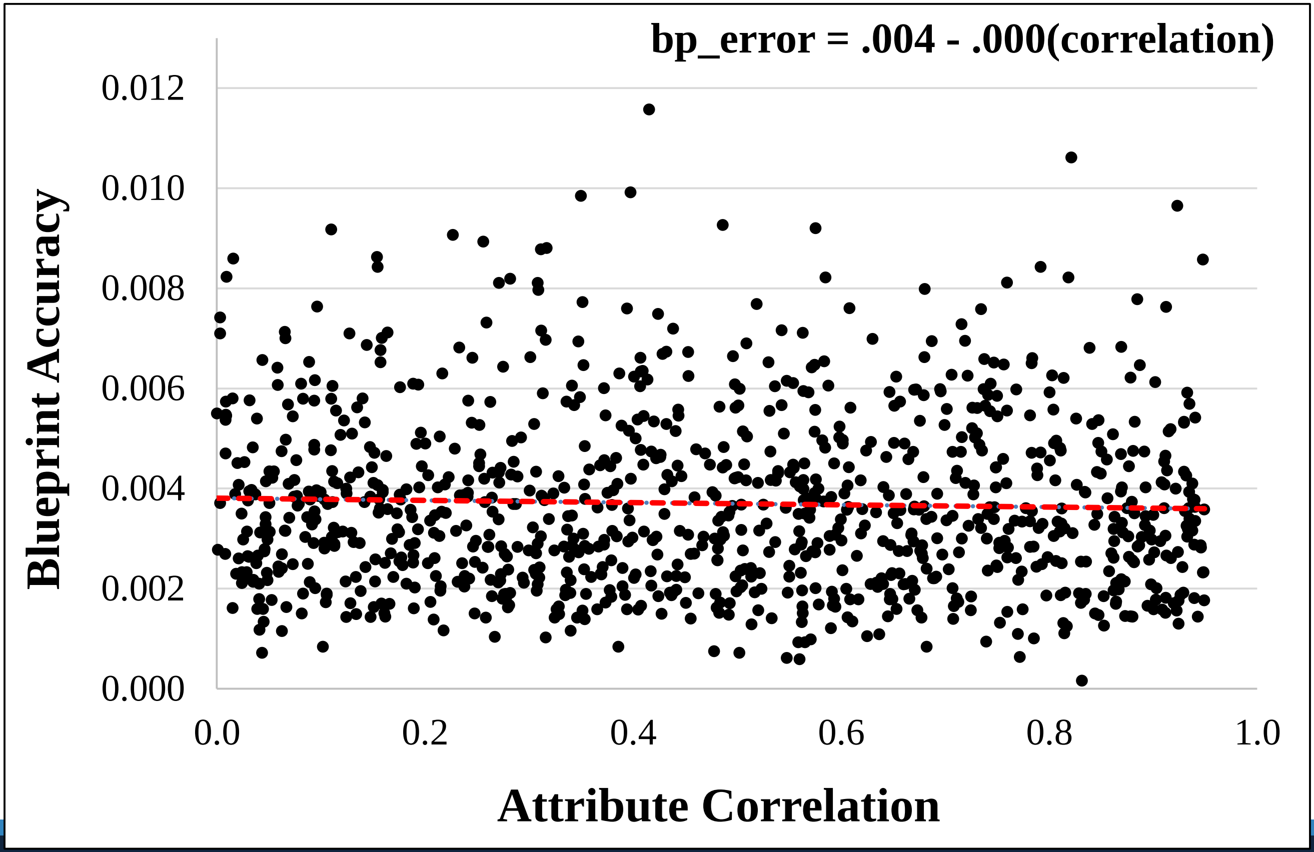
- Simulation #1: attribute correlations ~ $U(0, .95)$

- Simulation #2: main effects ~ $N(2, u)$, $u \sim U(0,1)$

- Simulation #3: Q-matrix complexity
  - » Complexity defined as the number of items measuring two attributes
  - » Complexity ~ $U(0, .80)$
  - » First estimated the CRUM, used the primary (i.e., largest) attribute main effect to simplify the Q-matrix
  - » Calibrated the 1-PLCDM using the simplified Q-matrix

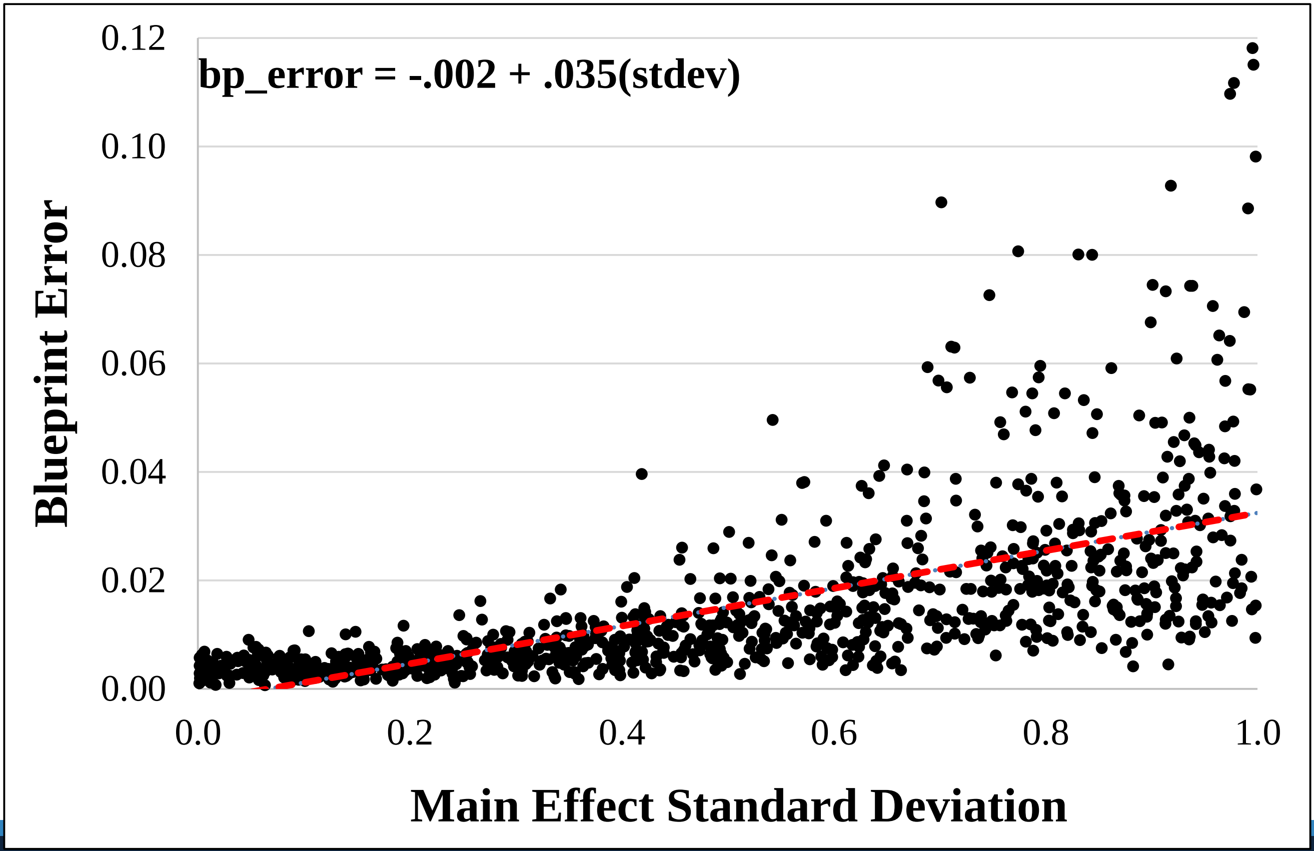
Result #1: Attribute Correlation

- Simulation #1: attribute correlations ~ $U(0,.95)$
- Slope = -.000, p = .269, $R^2 = .001$



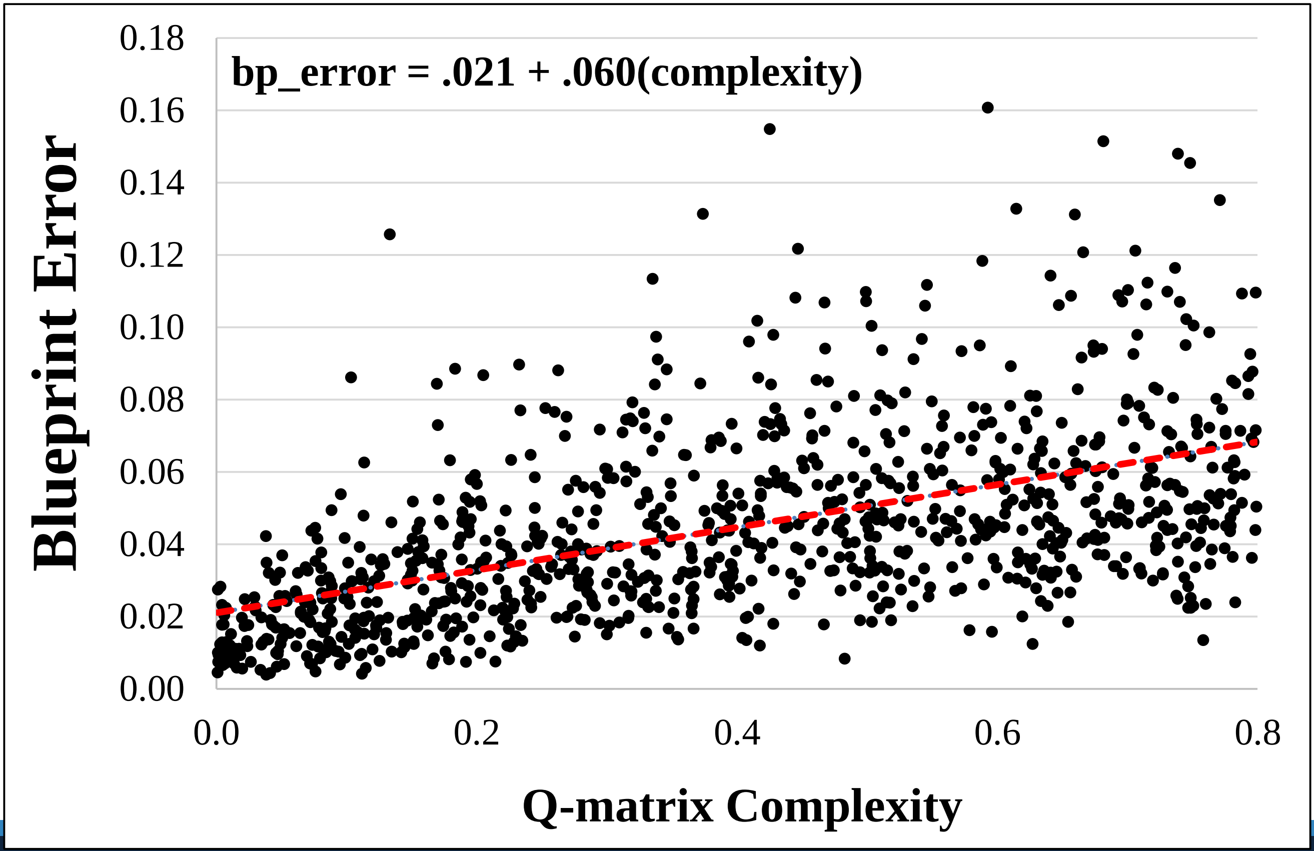
Result #2: Main Effect Equality

- Simulation #2: main effects ~ $N(2, u)$, $u \sim U(0,1)$
- Slope = .035, p<.001, $R^2 = .416$



Result #3: Q-matrix Complexity

- Simulation #3: Q-matrix complexity ~ $U(0, .80)$
- Slope = .060, $p < .001, R^2 = .250$



Conclusions

- The 1-PLCDM is the DCM version of the 1-PL IRT model
  - » Similar properties: sufficiency, invariance
  - » Blueprint matching
  - » Strong assumptions
- We examined its robustness to assumption violations
  - » Highly robust to attribute independence violations
  - » Slight to moderate negative effects of main effect variance
  - » Moderate negative effects of high Q-matrix complexity
- Overall, 1-PLCDM is highly robust
  - » Only in extreme violations, did we observe notable decreases
  - » Even in worst cases, overall blueprint error was within acceptable ranges

Conclusions

- Despite these results, we do not recommend ignoring moderate to extreme violations
- While mean error was low, towards high end of the violations, variance in performance increases (heteroscedastic)
- Finite simulation conditions can't cover all situations
  - » Combinations of violations with other types of misfit
- Acceptable error depends on context
- We recommend comparing models and evaluating assumptions

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

mjmadison@uga.edu

Access slides and paper at www.matthewmadison.com/research