

A Brief Introduction to Longitudinal Diagnostic Classification Models

Matthew Madison, University of Georgia

Workshop Overview

Access session materials at
www.matthewmadison.com/workshop

Workshop Overview

- Conceptual and statistical foundations (3 – 3:25 pm)
- Empirical demonstrations in R Studio (3:25 – 3:50 pm)
- Discussion and Q&A (3:50 – 4 pm)

A Little About Me

- From Columbia, South Carolina
- Research interests: DCMs, measurement of growth
- Hobbies: sports (basketball, golf), true crime



DCM Preliminaries

Conceptual Example

- Suppose a basketball coach wants to test basic basketball ability
- Using traditional psychometric approaches, the coach could obtain an ability estimate for each player
 - » Classical test theory: obtain a test score (# items correct)
 - » Item response theory: obtain a latent ability estimate
- Both these approaches are normative in nature, only provide information on students' location on a continuum

Conceptual Example

- Traditional testing procedures measure an **overall** ability in an area with a **continuous** latent variable:

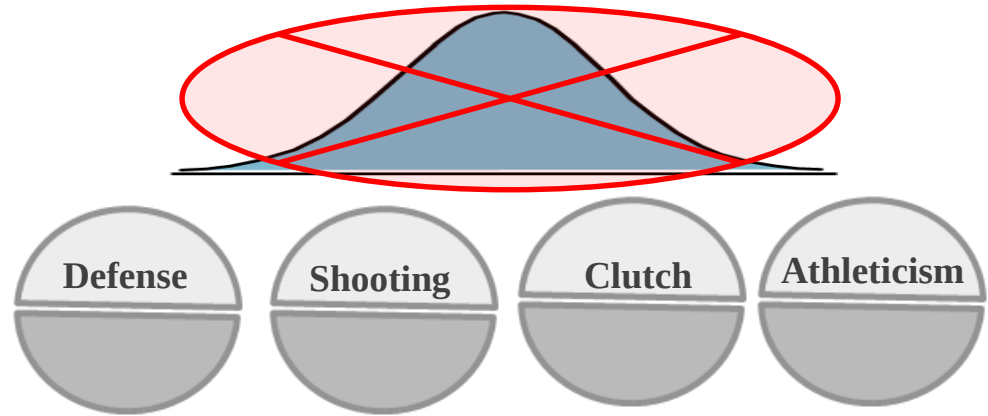


Basketball Ability

Basketball Ability Example

- Instead of measuring an overall basketball ability, let's break basketball down into a set of skills or *attributes*:

- Defense
- Shooting
- Clutch
- Athleticism



- These attributes are categorical (not continuous).
- Often, they are dichotomous:
 - Model places examinees into two categories (e.g., Proficiency vs Non-proficiency)

Diagnostic Classification Models

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

Player	Defense	Shooting	Clutch	Athleticism
	✓	✓	✗	✗
	✓	✓	✗	✓
	✓	✓	✓	✓

Conceptual Foundation Summary

- DCMs are psychometric models designed to classify
 - Diagnostic and criterion-referenced interpretations
- DCMs provide valuable information with more feasible data demands than other psychometric models
 - Higher reliability than IRT/MIRT models
 - Naturally accommodate multidimensionality
 - Complex items structures possible
 - Alignment of assessment goals and psychometric model

Questions? Comments?



Statistical Foundation

- Latent class models use responses to probabilistically put respondents into latent classes
 - » Latent class models are exploratory
- DCMs are confirmatory latent class models
 - » Latent classes specified a priori as attribute profiles
 - » Q-matrix specifies item-attribute structure (confirmatory)
 - » Person parameters are attribute proficiency probabilities

DCMs as Latent Class Models

- DCMs are constrained latent class models

Observed Data: Probability of observing examinee r 's vector of item responses to all I items

Measurement Component:
Product of Conditional Item Response Probabilities (Item Responses are Independent)

$$P(X_r = x_r) = \sum_{c=1}^C v_c \prod_{i=1}^I \pi_{ic}^{x_{ir}} (1 - \pi_{ic})^{1-x_{ir}}$$

Structural component:
Proportion of examinees in each class

Statistical Foundation

- Latent class models (LCMs) use responses to probabilistically put respondents into latent classes
- DCMs are confirmatory latent class models
 - » Latent classes specified a priori as attribute profiles
 - » Q-matrix specifies item-attribute structure (confirmatory)
 - » Person parameters are attribute proficiency probabilities
- DCMs are the categorical latent trait analogue of IRT models
 - » Item parameters model relationships between attribute proficiency and responses

Item Response Function Basics

- 2-PL item response function

- » Discrimination (a)
- » Difficulty (b)
- » Ability (θ)

$$\begin{aligned}\text{logit}(X_i = 1) &= a_i(\theta - b_i) \\ &= a_i\theta - a_ib_i \\ &= \boxed{\psi_i + \delta_i\theta}\end{aligned}$$

- DCM

- » Intercept (λ_0)
- » Main effects (λ_1)
- » Attribute proficiency status (α)

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

- α is either 0 or 1
- λ_0 is log-odds of a correct response for non-proficient
- λ_1 is the increase in log-odds of a correct response for proficient

Statistical Foundations Summary

- DCMs are confirmatory latent class models
- DCMs are categorical latent trait IRT models
- Person parameters (α): attribute proficiency probabilities
- Item parameters (λ): proficiency $\sim$ responses

DCM Inputs and Outputs

- Inputs:
 - Item response data ($N \times I$ matrix)
 - Q-matrix ($I \times A$ matrix)
 - Type of DCM you want to estimate, plus options
- Outputs
 - Item parameter estimates
 - Person parameter estimates (proficiency probabilities)
 - Attribute parameter estimates (proficiency proportions, correlations)
 - Model fit statistics

Estimation in R

- Several packages estimate DCMs
 - **CDM**
 - GDINA
 - mirt
 - blatant
 - hmcdm
 - npcd
 - PACER
- Other software programs as well
 - Mplus, flexMIRT, Latent Gold, mdltm

Questions? Comments?



Longitudinal DCMs

Assessing Growth in Education

- Recent shift in educational policy
 - » Meeting state standards → measuring student progress
- Need to quantify learning
 - » Examinees take pre-test
 - » Intervention (e.g., instruction)
 - » Examinees take post-test
- Goal: assess growth in students



Assessing Growth in Education

- Recent shift in educational policy
 - » Meeting state standards → measuring student progress
- Need to evaluate instructional methods and materials
- Is Method A >> Method B?
 - » Compare methods
 - » Examine differential growth

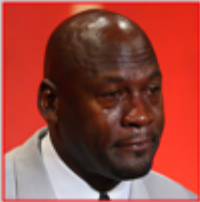


Growth and Item Response Models

- Classical Test Theory (CTT)
 - » Difference in pre-test and post-test score = Gain Score
 - » Gain score is interpreted as quantification of learning
 - » Can suffer from low reliability
- Item Response Theory (IRT)
 - » Difference in pre-test and post-test ability = growth
 - » Longitudinal IRT models
 - » Lack diagnostic feedback, reliability

Growth and DCMs

- Diagnostic classification models (DCMs)
 - » Classify on categorical latent traits: *attributes*
 - » Attribute mastery or non-mastery

Time Point	Student	Addition	Subtraction	Multiplication	Division
Pre-test		✗	✗	✗	✗
Post-test		✓	✓	✓	✗

Pre-test Post-test

Growth and DCMs

- Diagnostic classification models (DCMs)
 - » Classify on categorical latent traits: *attributes*
 - » Attribute mastery or non-mastery
- Growth in a DCM framework
 - » Individual: non-mastery $\leftrightarrow$ mastery
 - » Group: 30% at pre-test $\rightarrow$ 80% at post-test
- Longitudinal DCMs can be the psychometric foundation of a categorical growth model

Latent Class Models

- DCMs are constrained latent class models

Observed Data: Probability of the observed item responses

Measurement Component:
Models how latent classes respond to items

The diagram shows the equation for the probability of observed data in a Latent Class Model. The equation is $\mathbb{P}(\mathbf{X}_r = \mathbf{x}_r) = \sum_{c=1}^C \nu_c \prod_{i=1}^I \pi_{ic}^{X_{ir}} (1 - \pi_{ic})^{1-X_{ir}}$. The terms are grouped into three colored ovals: a green oval around $\mathbb{P}(\mathbf{X}_r = \mathbf{x}_r)$, a blue oval around $\sum_{c=1}^C \nu_c$, and a red oval around $\prod_{i=1}^I \pi_{ic}^{X_{ir}} (1 - \pi_{ic})^{1-X_{ir}}$. A green arrow points from the green oval to the 'Observed Data' box. A blue arrow points from the blue oval to the 'Structural component' box. A red arrow points from the red oval to the 'Measurement Component' box.

$$\mathbb{P}(\mathbf{X}_r = \mathbf{x}_r) = \sum_{c=1}^C \nu_c \prod_{i=1}^I \pi_{ic}^{X_{ir}} (1 - \pi_{ic})^{1-X_{ir}}$$

Structural component:
Models how classes are related, class proportions

Latent Class Models

- There are longitudinal latent class models
 - » Latent transition analysis (LTA)

$$\mathbb{P}(\mathbf{X}_r = \mathbf{x}_r) = \sum_{c_1=1}^C \sum_{c_2=1}^C \cdots \sum_{c_T=1}^C v_{c_1} \tau_{c_2|c_1} \tau_{c_3|c_2} \cdots \tau_{c_T|c_{T-1}} \prod_{t=1}^T \prod_{i=1}^I \pi_{ic_t}^{x_{rit}} (1 - \pi_{ic_t})^{1-x_{rit}}$$

The diagram illustrates the components of the Latent Transition Analysis (LTA) equation. The equation is shown with three parts highlighted by colored ovals and arrows pointing to explanatory boxes:

- Structural component:** A blue oval highlights the initial class proportions v_{c_1} . An arrow points from this oval to a blue-bordered box containing the text: "Structural component: Proportion of examinees in each class at Time Point 1".
- Transition Probabilities:** A purple oval highlights the transition probabilities $\tau_{c_2|c_1} \tau_{c_3|c_2} \cdots \tau_{c_T|c_{T-1}}$. An arrow points from this oval to a purple-bordered box containing the text: "Transition Probabilities: Probabilities of transition between latent classes from time to time".
- Measurement Component:** A red oval highlights the product of conditional item response probabilities $\prod_{t=1}^T \prod_{i=1}^I \pi_{ic_t}^{x_{rit}} (1 - \pi_{ic_t})^{1-x_{rit}}$. An arrow points from this oval to a red-bordered box containing the text: "Measurement Component: Product of Conditional Item Response Probabilities (Across time points)".

Transition DCM

- Transition Diagnostic Classification Model (TDCM)
 - » Combines LTA with the LCDM as measurement model
 - » General form (T time points):

$$\mathbb{P}(\mathbf{X}_r = \mathbf{x}_r) = \sum_{c_1=1}^C \sum_{c_2=1}^C \cdots \sum_{c_T=1}^C \nu_{c_1} \tau_{c_2|c_1} \tau_{c_3|c_2} \cdots \tau_{c_T|c_{T-1}} \prod_{t=1}^T \prod_{i=1}^I \pi_{ic_t}^{x_{rit}} (1 - \pi_{ic_t})^{1-x_{rit}}$$

- » Pre-test/post-test form (2 times points w/ invariance):

$$\mathbb{P}(\mathbf{X}_r = \mathbf{x}_r) = \sum_{c_1=1}^C \sum_{c_2=1}^C \nu_{c_1} \tau_{c_2|c_1} \prod_{t=1}^2 \prod_{i=1}^I \pi_{ic_t}^{x_{rit}} (1 - \pi_{ic_t})^{1-x_{rit}}$$

Measurement Invariance

- Reflects the notion that measurement is happening the same way across groups or over time
- In a longitudinal psychometric analysis, person estimates should be determined the exact same way
- For the TDCM, we want that proficiency is being determined the same way over time
 - » Otherwise, change in proficiency status is meaningless
- Can be assessed by statistically comparing a constrained model to a free model

Madison & Bradshaw (2018)

- National Assessment of Educational Progress, 2011
 - » 23% of students w/o disabilities scored below *basic*
 - » 65% of students with disabilities scored below *basic*
- Basic: operations with rational numbers and real world problem solving
- Research suggests that students with disabilities struggle with problems embedded in text
- University of Kentucky (Bottge, et. al) focus on improving mathematics instruction and assessment for students with disabilities

Enhanced Anchored Instruction

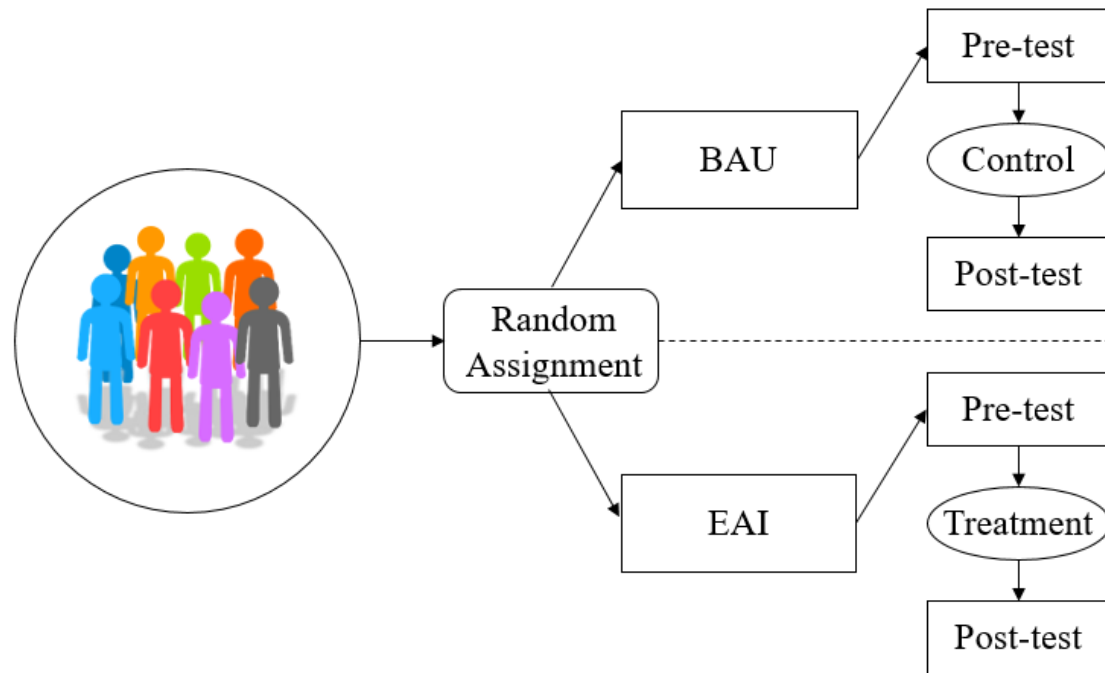
- Acronym (EAI)
- Instructional method developed to improve problem solving
- Realistic problems are embedded in interesting contexts
 - » 8-15 minute videos
 - » Hands-on applications
- Include modules and interactive media
 - » Help students visualize abstract concepts

EAI Study Design

- Overall research questions:
 - » To what extent does EAI improve students' problem solving abilities?
 - » How well does it work compared to traditional instruction?
- Team developed an assessment targeting multiple areas
 - » 21-item problem-solving test
 - » Measured four Common Core *strands*
 - Ratios and proportional relationships
 - Measurement and data
 - Numbers systems (fractions)
 - Geometry (graphing)

EAI Study Design

- Business as Usual (BAU, $N = 456$) vs. EAI ($N = 423$)
- Sample of 6-8 graders



- Is EAI better than traditional instruction?

Analysis

- Evaluate intervention effect
 - » Use multiple-group TDCM
 - » Analyze growth for BAU and EAI
 - » Compare growth
- How does growth for EAI students compare to growth for BAU students?
 - » For which attributes does EAI appear to be better than BAU?
 - » Where can EAI be improved?

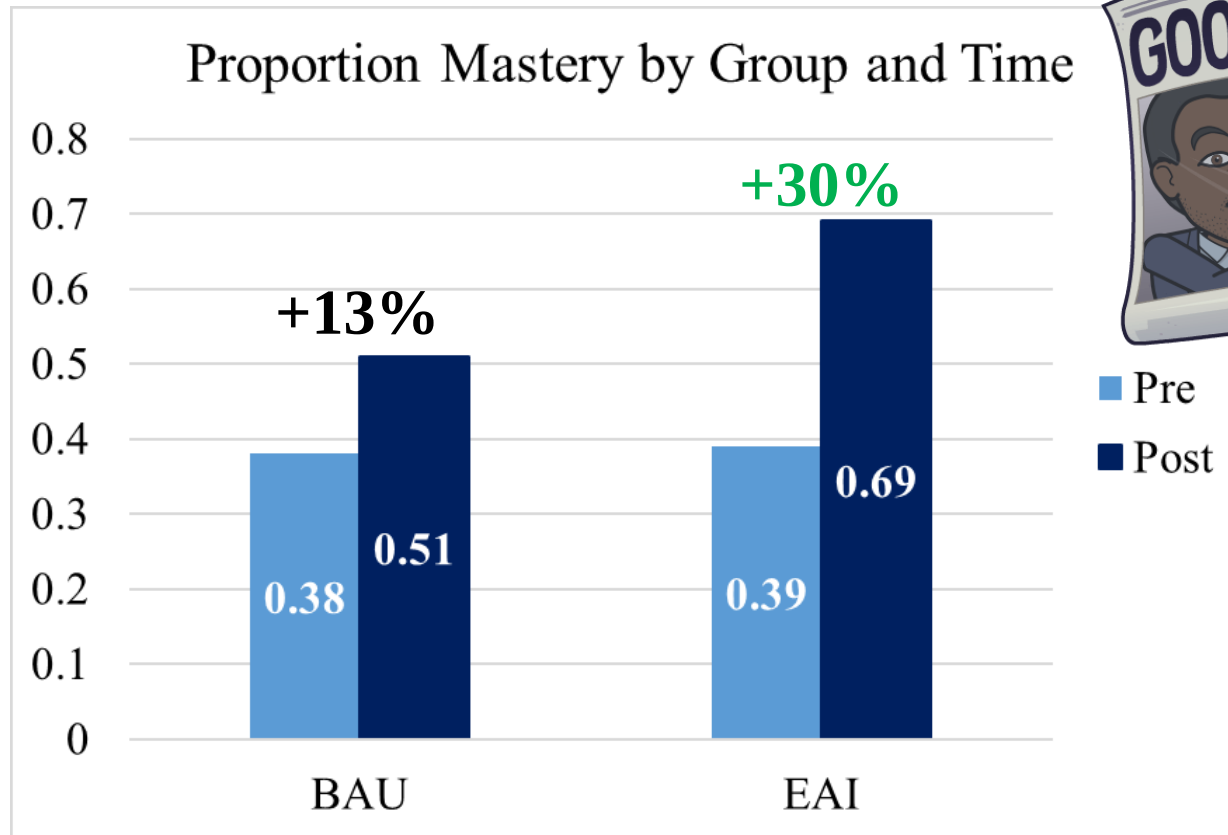
Analysis

- To analyze the intervention effect, extend TDCM to multiple groups
 - » Attribute proficiency proportions and transitions depend on group membership (g)

$$\mathbb{P}(\mathbf{X}_r = \mathbf{x}_r | G = g) = \sum_{c_1=1}^C \sum_{c_2=1}^C \boxed{v_{c_1|g} \tau_{c_2|c_1,g}} \prod_{t=1}^2 \prod_{i=1}^I \pi_{ic}^{x_{rit}} (1 - \pi_{ic})^{1-x_{rit}}$$

Results: Differential Growth

- Attribute 1: Ratios and Proportions



Results: Differential Growth

Attribute	Growth (EAI)	Growth (BAU)	Difference
Ratios and proportional relationships	.31	.13	.18 (.05)
Measurement and data	.07	.07	.00 (.04)
Number Systems (fractions)	.18	.13	.05 (.04)
Geometry (graphing)	.26	.11	.15 (.05)

Results: Conditional Growth

- Proficiency transitional probabilities
 - » Ratios and proportional relationships

BAU

		Post-test	
		0	1
Pre-test	0	.63	.37
	1	.27	.73

EAI

		Post-test	
		0	1
Pre-test	0	.41	.59
	1	.14	.86

Alternative Analysis

- Can use gain scores and repeated measures MANOVA to analyze this data
 - » Interaction effects test differential growth
- Could also use a longitudinal IRT model
- Why use a DCM?
 - » Two main reasons

Interpretive Value

- Effect size measures
 - » TDCM: difference in groups' growth in attribute mastery
 - » Gain scores: average difference in groups' gain scores
 - » Longitudinal IRT: difference in groups' growth in ability
- Effect sizes:

Attribute	TDCM	Gain Scores	Long IRT
Ratios and proportions	.18 (72)	.49	.51



Reliability

- Refers to the consistency of the growth estimates
 - » Are they trustworthy?

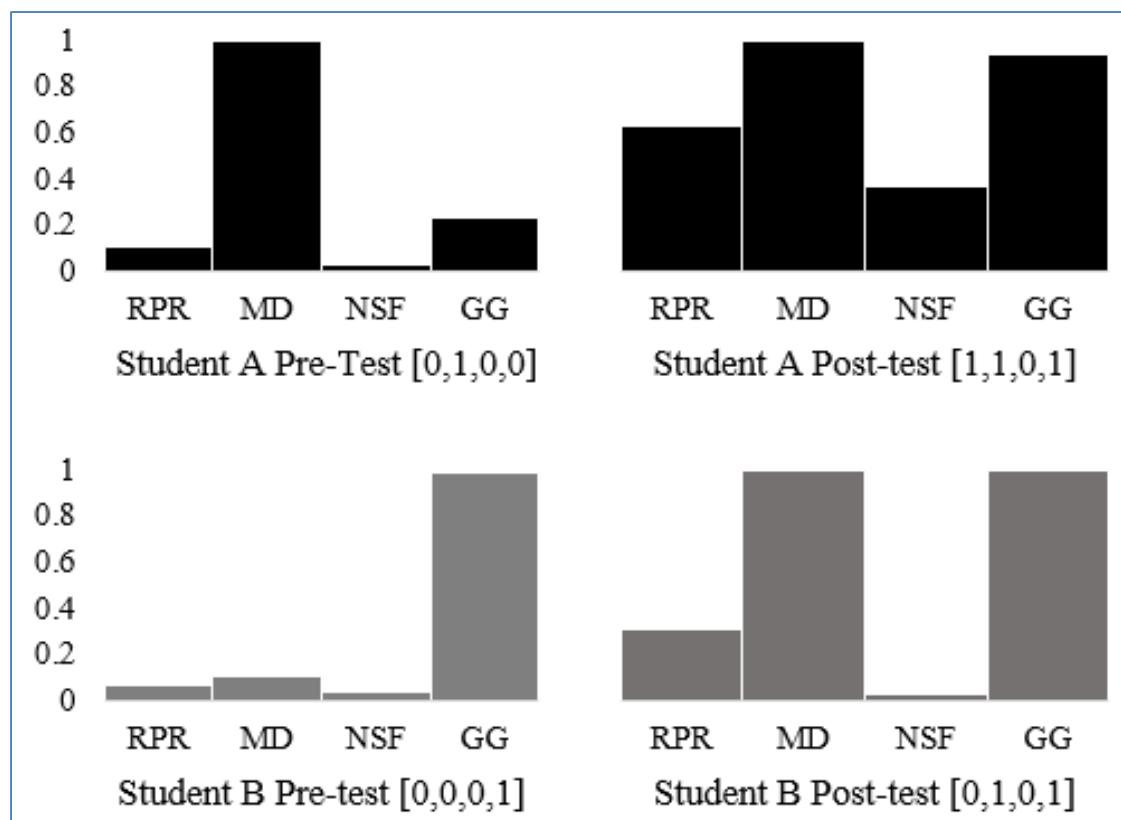
Attribute	TDCM	Gain Scores	Long IRT
Ratios and proportions	.915	.488	.550
Measurement and data	.968	.651	.566
Number Systems	.928	.603	.631
Geometry (graphing)	.911	.619	.638

Beautiful



Interpretive Value

- Individual estimates
- Two examinees (Student A and B):
 - » Pre-test = 9/21
 - » Post-test = 12/21
 - » Gain Score = +3



Conclusions

- Analyzed data from a randomized controlled trial
- Evaluated efficacy of EAI
 - » Used TDCM, which has benefits over traditional analysis
 - » Analysis suggests EAI better than traditional instruction
- Overall, EAI was successful
 - » But...could be improved for number systems and measurement and data
- The TDCM provides a meaningful measure of growth



Estimation in R

- The previous analysis was completed in Mplus
 - » Software is expensive
 - » Estimating DCMs and TDCMs is very tedious
 - » Thousands of lines of code and days to run multigroup TDCM
- New R package (Madison, et al. 2024)
 - » Free and easy
 - » Very fast
 - » Output is optimized for DCM interpretations

Estimation in R

- Open “TDCM Examples.R”
- Three empirical examples
 - » Data included in the package
 - » Demonstrate functionality through these examples