Cognitive Diagnosis Modeling: An Introduction and Its Implementation in R

Wenchao Ma

The University of Alabama

May 27, 2023

浙江大学心理与行为科学系MAP开放课堂

Outline

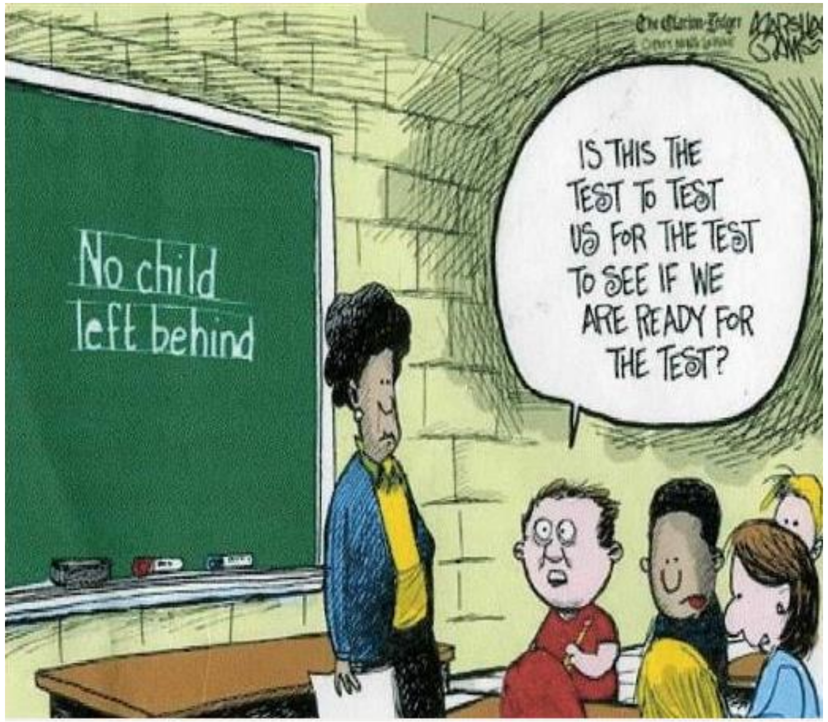
- Introduction to Diagnostic Assessments
- Cognitive Diagnosis: Terminology, Inputs and Outputs
- Cognitive Diagnosis Basics Again: Attributes
- Cognitive Diagnosis: Models
- Cognitive Diagnosis: Applications and New Developments
- Cognitive Diagnosis in R

Educational Assessments

- Educational assessment is a process designed to systematically measure or evaluate the characteristics or performance of individuals, programs, or other entities, for purposes of drawing inferences; sometimes used synonymously with test. (NCME)
- Assessments can be grouped into two categories (Bloom, 1969):
  - Assessments **OF** learning
  - Assessments **FOR** learning

Assessments OF Learning

- Goal
  - To determine what students know and can do after completing a particular phase of education
- Characteristics
  - Often presented as standardized tests
  - Measuring student’s overall proficiency in a particular subject
  - High-stakes
  - Sometimes used for accountability
- Examples
  - End of course exam
  - College admission exams



By Dr. Helen Teague

Assessments FOR Learning

- Goal
  - To provide feedback to adjust ongoing teaching and learning to improve students' achievement
- Characteristics
  - An integration of process and purposefully designed methodology or instrumentation (Bennett, 2011)
  - Low-stakes
  - Immediate feedback
- Examples
  - Diagnostic tests
  - 'interim' assessments

Assessments FOR Learning

- What type of feedback should we give to students and teachers?

> “*If I have to reduce all educational psychology to just one principle, I would say this: the most important single factor influencing learning is what the learner already knows. Ascertain this and teach him [them] accordingly.*”

Ausubel (1968)

Assessments FOR Learning

- To understand what students know and what they do not know, we consider cognitively diagnostic assessments:
  - Standards- or skills-based
  - Conceptually multidimensional
  - Statistically reliable
  - Didactically teachable

Reasoning From Evidence

- What all educational assessments have in common is the desire to reason from particular things students say and do, and make inferences about what they know or can do more broadly
- An assessment is a tool designed to observe students’ behavior and produce data that can be used to draw reasonable inferences about what students know
- The process of collecting evidence to support the type of inferences one wants to draw is referred to as *reasoning from evidence*

Reasoning From Evidence

Developing assessments for gathering evidence

Cognitively diagnostic tests

Collecting evidence

Task performance data

Interpreting evidence

Cognitive diagnosis modeling
Reliability and validity consideration

Acting on interpretations

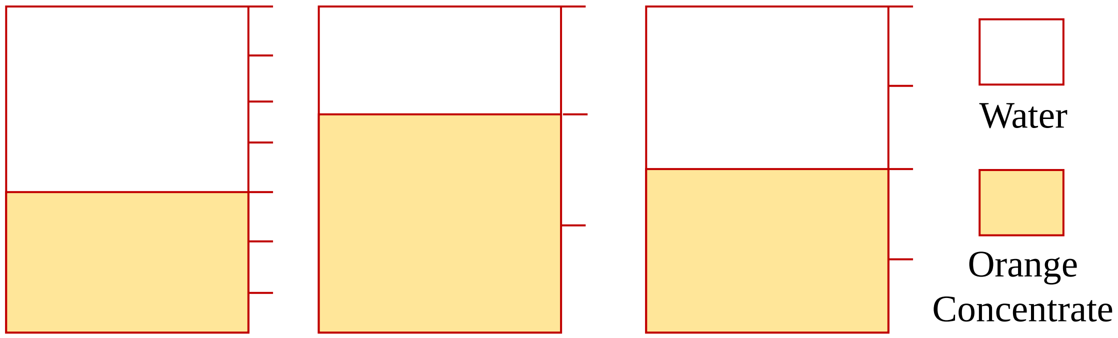
Instruction adjustments
Personalized learning recommendations

An example

A Proportional Reasoning Test

A Proportional Reasoning Test

1) Nate and Dale are making s'mores. Nate has 4 marshmallows and 3 crackers. Dale has 7 marshmallows and twice as many crackers as Nate. Whose s'mores have a stronger marshmallow taste (greater marshmallows-to-crackers ratio)?
2) Solve for x in the equation $\frac{5}{3} = \frac{8}{x}$?
3) Three recipes for orange juice are shown below. Put the recipes in order from the one with the smallest fraction of orange concentrate to the one with the greatest fraction of orange concentrate.



(See Tjoe & de la Torre, 2014)

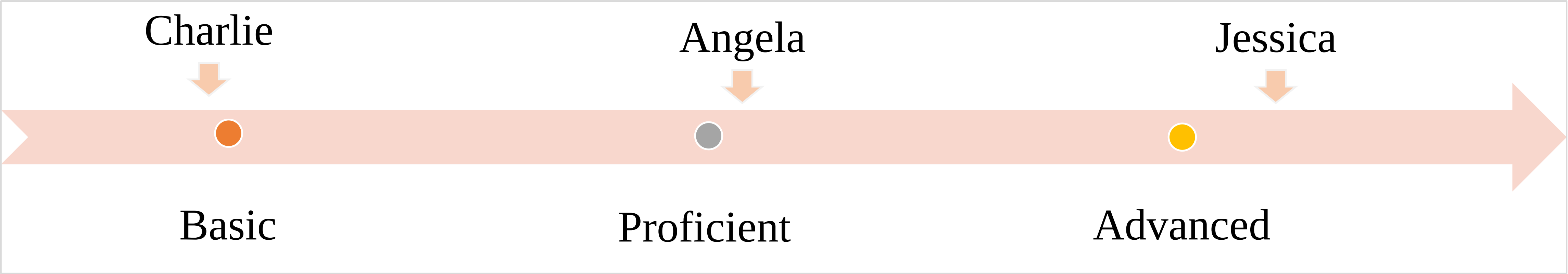
A Proportional Reasoning Test

Students' responses

	Item 1	Item 2	Item 3	…
Angela	√	√	×	
Charlie	×	×	√	
Jessica	√	√	√	
…				

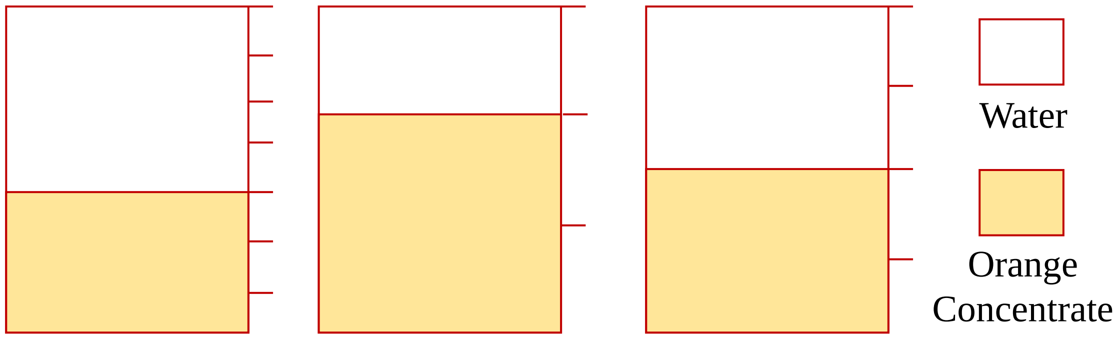
A Proportional Reasoning Test

- Based on CTT or IRT analyses, we could
  - Put all students on the same ability scale
  - Compare their proficiency
  - Make other decisions accordingly



A Proportional Reasoning Test

1) Nate and Dale are making s'mores. Nate has 4 marshmallows and 3 crackers. Dale has 7 marshmallows and twice as many crackers as Nate. Whose s'mores have a stronger marshmallow taste (greater marshmallows-to-crackers ratio)?
2) Solve for x in the equation $\frac{5}{3} = \frac{8}{x}$?
3) Three recipes for orange juice are shown below. Put the recipes in order from the one with the smallest fraction of orange concentrate to the one with the greatest fraction of orange concentrate.



(See Tjoe & de la Torre, 2014)

A Proportional Reasoning Test

1) Nate and Dale are making s'mores. Nate has 4 marshmallows and 3 crackers. Dale has 7 marshmallows and twice as many crackers as Nate. Whose s'mores have a stronger marshmallow taste (greater marshmallows-to-crackers ratio)?

[Skills: Prerequisite skills; Constructing ratios; Comparing fractions]

2) Solve for x in the equation $\frac{5}{3} = \frac{8}{x}$?

[Skills: Prerequisite skills; Applying algorithms]

3) Three recipes for orange juice are shown below. Put the recipes in order from the one with the smallest fraction of orange concentrate to the one with the greatest fraction of orange concentrate.

[Skills: Ordering fractions]

A Proportional Reasoning Test

Item and attribute association matrix (Q-matrix; Tatsuoka, 1983)

	Prerequisite skills	Comparing fractions	Ordering fractions	Constructing ratios	Applying algorithms
Item 1	1	1	0	1	0
Item 2	1	0	0	0	1
Item 3	0	0	1	0	0
⋮					

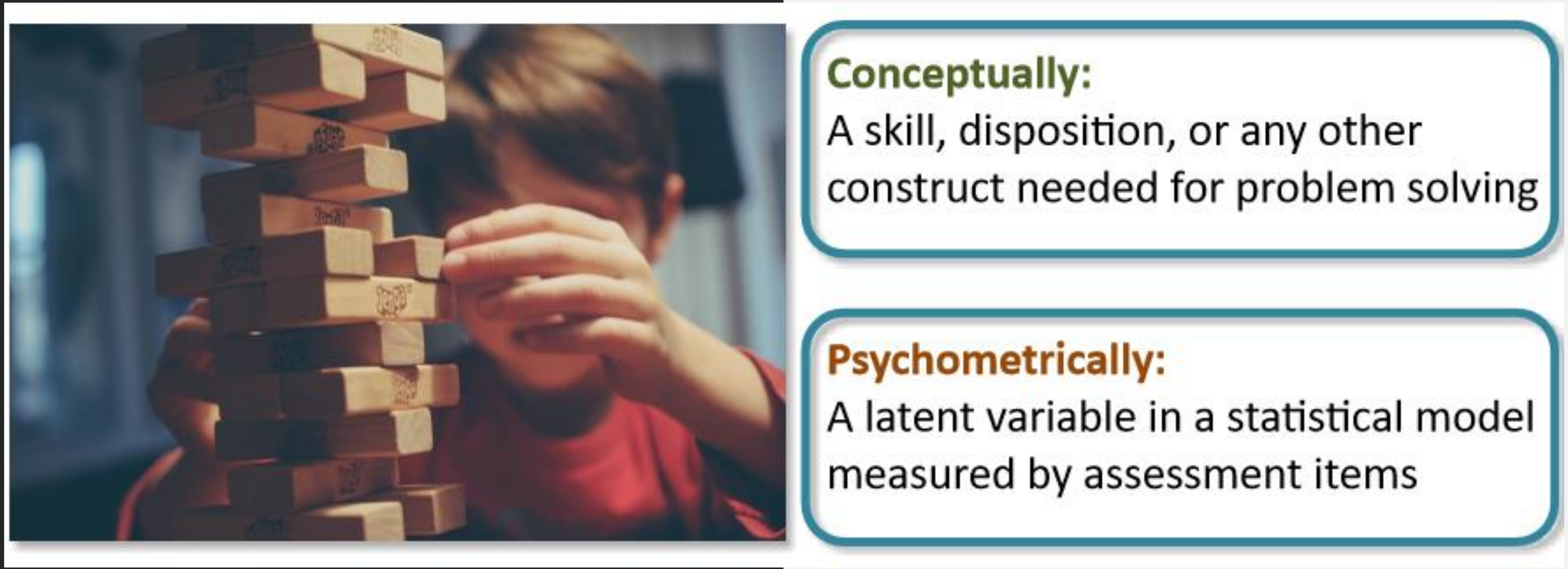
A Proportional Reasoning Test

Score report from CDM analyses

	Prerequisite skills	Comparing fractions	Ordering fractions	Constructing ratios	Applying algorithms
Angela	0	1	0	1	0
Charlie					
Jessica					
⋮					

Outline

- Introduction to Diagnostic Assessments
- Cognitive Diagnosis: Terminology, Inputs and Outputs
- Cognitive Diagnosis Basics Again: Attributes
- Cognitive Diagnosis: Models
- Cognitive Diagnosis: Applications and New Developments
- Cognitive Diagnosis in R



Attribute:

any procedure, skill, or process that involves in the problem-solving process

- Attributes are discrete
- Attributes can have hierarchical structures
- Attributes are latent variables

	Prerequisite skills	Comparing fractions	Ordering fractions	Constructing ratios	Applying algorithms
Angela	0	0	0	0	0
Charlie	1	0	0	0	0
Jessica	0	1	0	0	0
⋮					

If the test measures 5 attributes, 5 mastery statuses need to be estimated for each student.

Attribute:

any procedure, skill, or process that involves in the problem-solving process

Attributes: Why latent variables?

- Students' responses do not always a reflection of what they know
  - Guessing
  - Carelessness
- Tasks often involve multiple attributes
  - Hard to know why a student fail
  - Hard to know why a student succeed
- Latent variables allow us to make principled inferences from evidence

Conceptually:
An unobserved grouping of learners that share similar characteristics

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

Latent class

	Prerequisite skills	Comparing fractions	Ordering fractions	Constructing ratios	Applying algorithms
α_1	0	0	0	0	0
α_2	1	0	0	0	0
α_3	0	1	0	0	0
$\vdots$					
α_{32}	1	1	1	1	1

If a test measures 5 attributes, there are $2^5=32$ possible attribute profiles, each labelling a latent class.

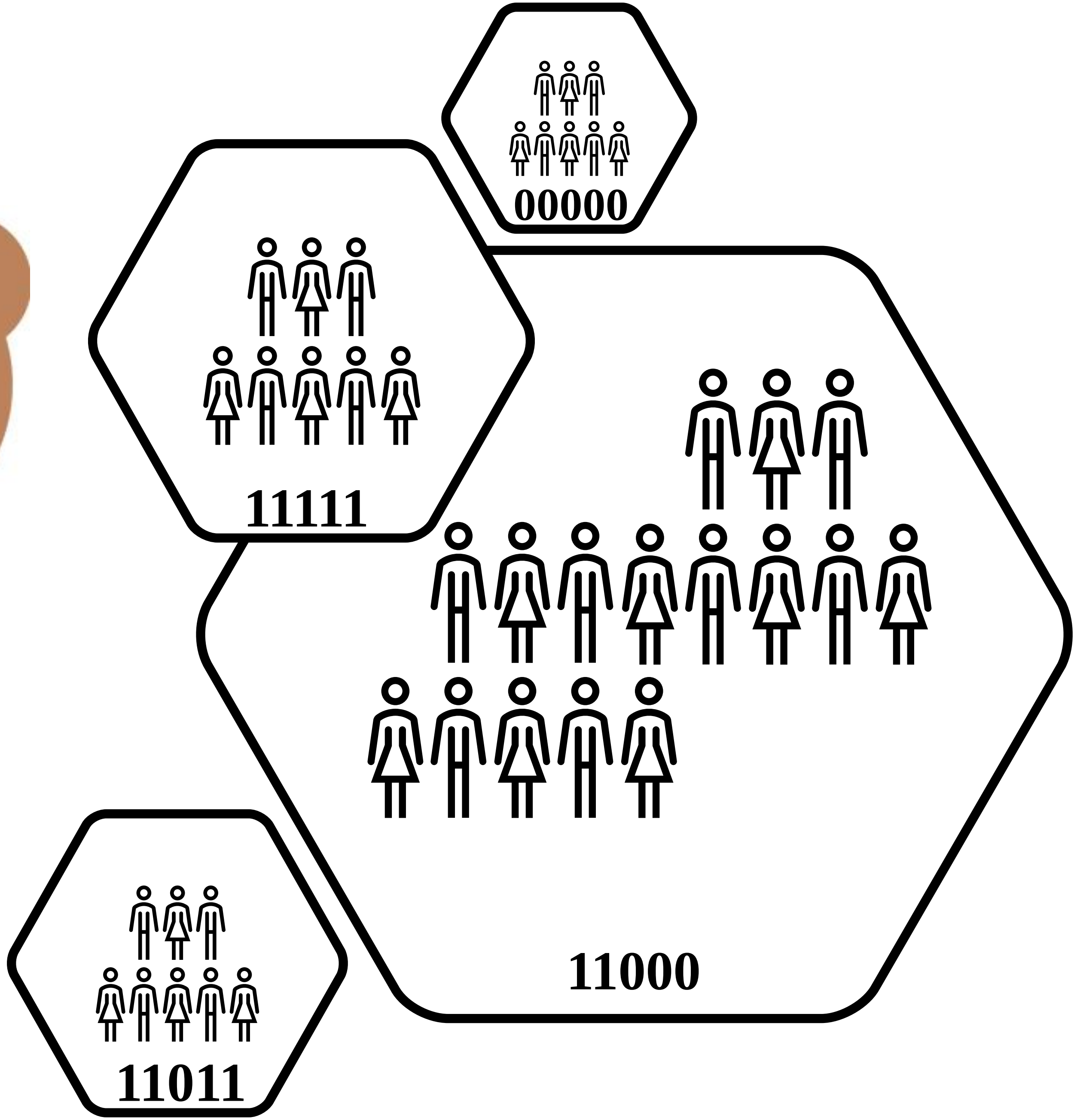
Latent class

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

Diagnosis



CDM inputs: Item response data

Columns = Items

Rows = Learners

	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

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

CDM inputs: Q-matrix

Columns = Attributes

Rows = Items

	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

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

CDM outputs

- Item parameters

- Different CDMs have **different parameterizations**

Deterministic input noisy "and" gate (DINA) model:

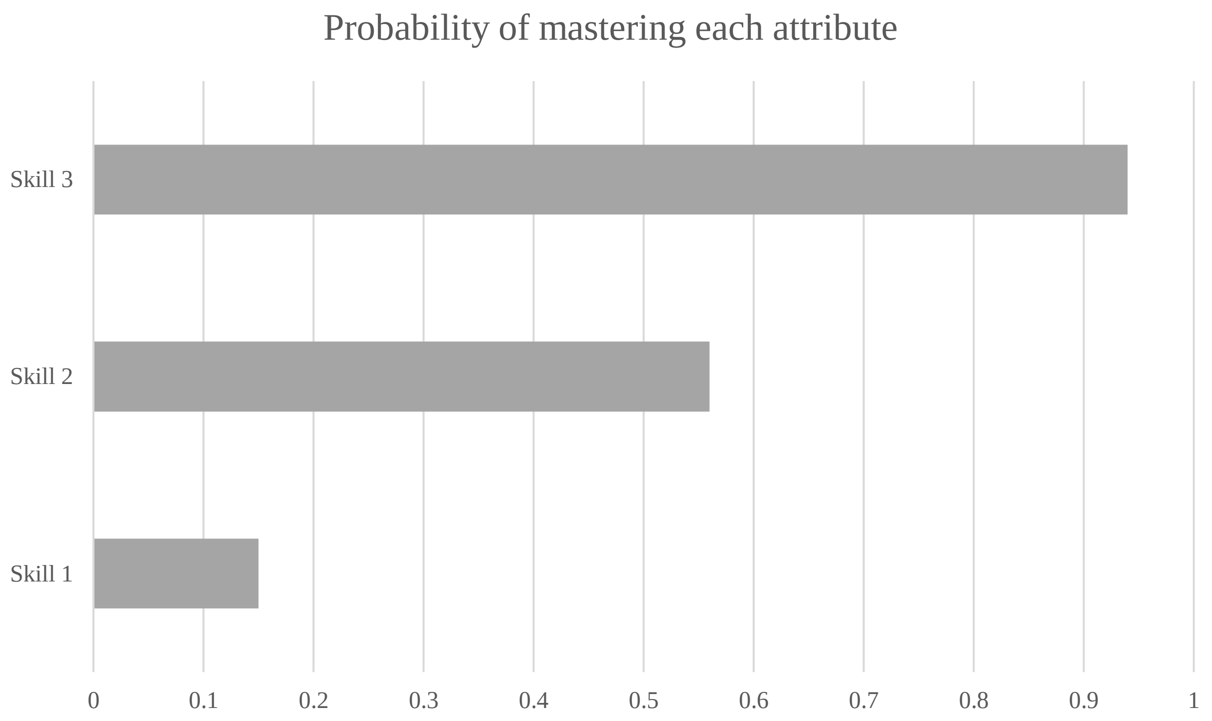
- guessing parameter
- slipping parameter

- **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

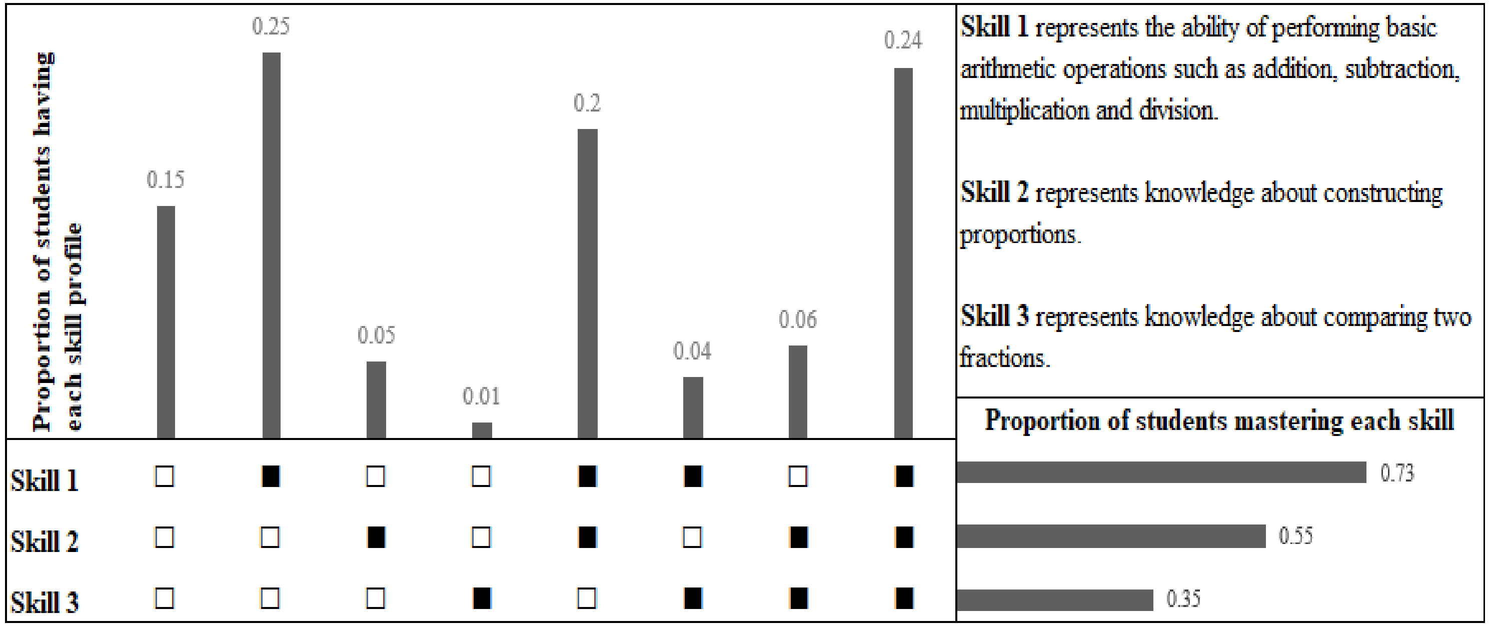
CDM outputs

- Item parameters
- Individual-level profiles



CDM outputs

- Item parameters
- Individual-level profiles
- Population-level profiles



Outline

- Introduction to Diagnostic Assessments
- Cognitive Diagnosis: Terminology, Inputs and Outputs
- Cognitive Diagnosis Basics Again: Attributes
- Cognitive Diagnosis: Models
- Cognitive Diagnosis: Applications and New Developments
- Cognitive Diagnosis in R

Attributes

- (Psychometric) attributes characterize test items, and may be interpreted as cognitive processes, procedures, or skills that are required to perform correctly on a particular test item
- In more recent usage, the term *attributes* has been redefined *to* refer to any procedures, skills, or processes that an examinee must possess to solve a test item
- That is, the term attributes has been used to refer to person-specific, not just item-specific characteristics

Attributes

- As in conventional IRT, attributes in cognitive diagnosis modeling are construed as latent constructs and are represented by latent variables in the CDMs
- In cognitive diagnosis modeling, the goal is to provide detailed information about the examinees' (cognitive) attributes

Attributes

- For diagnostic purpose, attributes need to be closely tied to classroom instruction:
  - Conceptually multidimensional
  - Fine-grained in nature
  - Curriculum- and skill-based
  - Statistically reliable
  - Didactically actionable

An example: attributes in a proportional reasoning test

Attribute	Description
A1	Prerequisite skills and concepts required in proportional reasoning
A2a	Comparing (two) fractions
A2b	Ordering (three or more) fractions
A3a	Constructing ratios
A3b	Constructing proportions
A4	Identifying a multiplicative relationship between sets of values
A5	Differentiating a proportional from a non-proportional relationship
A6	Applying algorithms in solving proportional reasoning problems

Attributes in Clinical Psychology

- A diagnostic classification test for internet addiction (Tu et al., 2017) based on Diagnostic and Statistical Manual of Mental Disorders (DSM-5).
- Based on the, DSM-5, to be classified as a internet gaming disorder, an individual must meet 5 or more criteria

ID	Symptom criteria
S1	Preoccupation with Internet (games)
S2	Withdrawal symptoms when Internet (gaming) is taken away
S3	Tolerance—the need to spend increasing amounts of time engaged in Internet (games)
S4	Unsuccessful attempts to control the participation in Internet (games)
S5	Loss of interests in previous hobbies and entertainment as a result of, and with the exception of, Internet (games)
S6	Continued excessive use of Internet (games) despite knowledge of psychosocial problems
S7	Has deceived family members, therapists, or others regarding the amount of Internet gaming
S8	Use of Internet (games) to escape or relieve a negative mood
S9	Has jeopardized or lost a significant relationship, job, or educational or career opportunity because of participation in Internet (games)

Other examples of Attributes in Clinical Psychology

- Alcohol-related problems, anxiety, hostility, and depression (Tan et al., 2022)
- Anxety, somatoform, thought disorder, major depression (de la Torre, van der Ark, & Rossi, 2018)

CDM Application in Personnel Selection

- Situational judgment tests (SJTs) have become popular for personnel selection
- These tests are designed to evaluate candidate’s judgments regarding situations encountered in the workplace
- Test takers are usually asked to choose an option from a set of possible course of actions

Attributes in Personnel Selection

- When studying for an exam, do you find that you reach best results when:
  - you start planning and setting aside time in advance
  - work in a clean environment, even if it means taking time away from studying
  - wait for inspirations before becoming involved in most important study tasks
  - wait until the last day or so to study, knowing that you have to get it done now

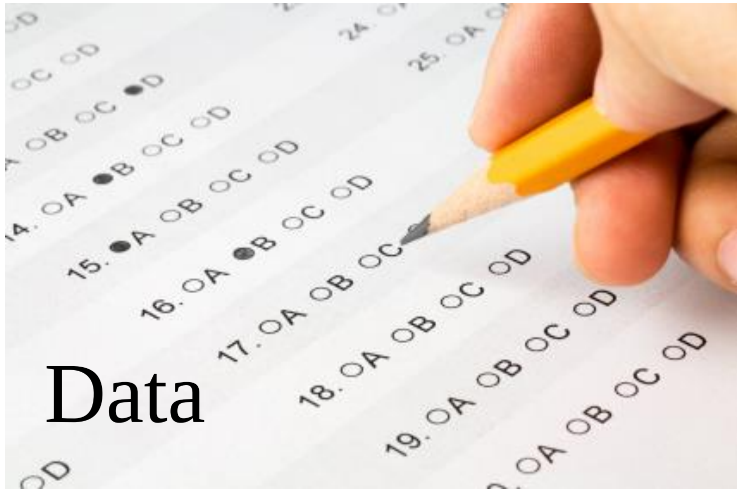
CDM Application in Personnel Selection

- To identify whether applicants have a set of desired characteristics (Sorrel, 2016):
  - A1: Study habits
  - A2: Study attitudes
  - A3: Helping others
  - A4: Generalized compliance

Outline

- Introduction to Diagnostic Assessments
- Cognitive Diagnosis: Terminology, Inputs and Outputs
- Cognitive Diagnosis Basics Again: Attributes
- Cognitive Diagnosis: Models
- Cognitive Diagnosis: Applications and New Developments
- Cognitive Diagnosis in R

Cognitive Diagnosis Models



Attributes

Q-matrix

Items

Cognitive Diagnosis Models → Attributes

Parametric Approaches for Diagnosis

- Cognitive Diagnosis models
  - Attributes are assumed to be categorical (often binary)
  - Also called Diagnostic classification models, Cognitive psychometric models, Multiple classification (latent class) models, Latent response models, Restricted latent class models, Structured located latent class models, Structured IRT models

Parametric Approaches for Diagnosis

- Defining characteristics of CDMs
  - Multidimensional nature
  - Confirmatory nature
  - Complexity of their loading structure
  - Types of observed response variables
  - Types of latent predictor variables
  - Interactions of the latent predictor variables
  - Criterion-referenced interpretations
  - Diagnostic nature of the interpretations

Data

	3X4-2	6-1	5+3X6
Student 1	0	0	0
Student 2	0	1	1
Student 3	1	1	1

Attributes

Items	+	-	X
3X4-2	0	1	1
6-1	0	1	0
5+3X6	1	0	1

Q-matrix

CDMs

Attribute Profiles

	+	-	X
Student 1	0	0	0
Student 2	1	1	0
Student 3	1	1	1

Questions

- Suppose N = sample size, J = test length, K= number of attributes
  - What is the dimension of the item response matrix?
  - What is the dimension of the Q-matrix?
  - What is the dimension of the attribute profile matrix?
  - What is the total number of possible attribute profiles?

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

Cognitive Diagnosis Models

- Models at item response level

$$P(\text{CorrectResponse}) = f\,(\text{AttributeProfile, ItemParameters})$$

- How students use attributes to solve each item
  - Conjunctive rule
  - Disjunctive rule

CDMs

- CDMs are item response models
- For item j

$$P(\text{CorrectResponse})=f\,(\text{AttributeProfile, ItemParameters})$$

- $\boldsymbol{\alpha}_l$ is the lth attribute vector for item j

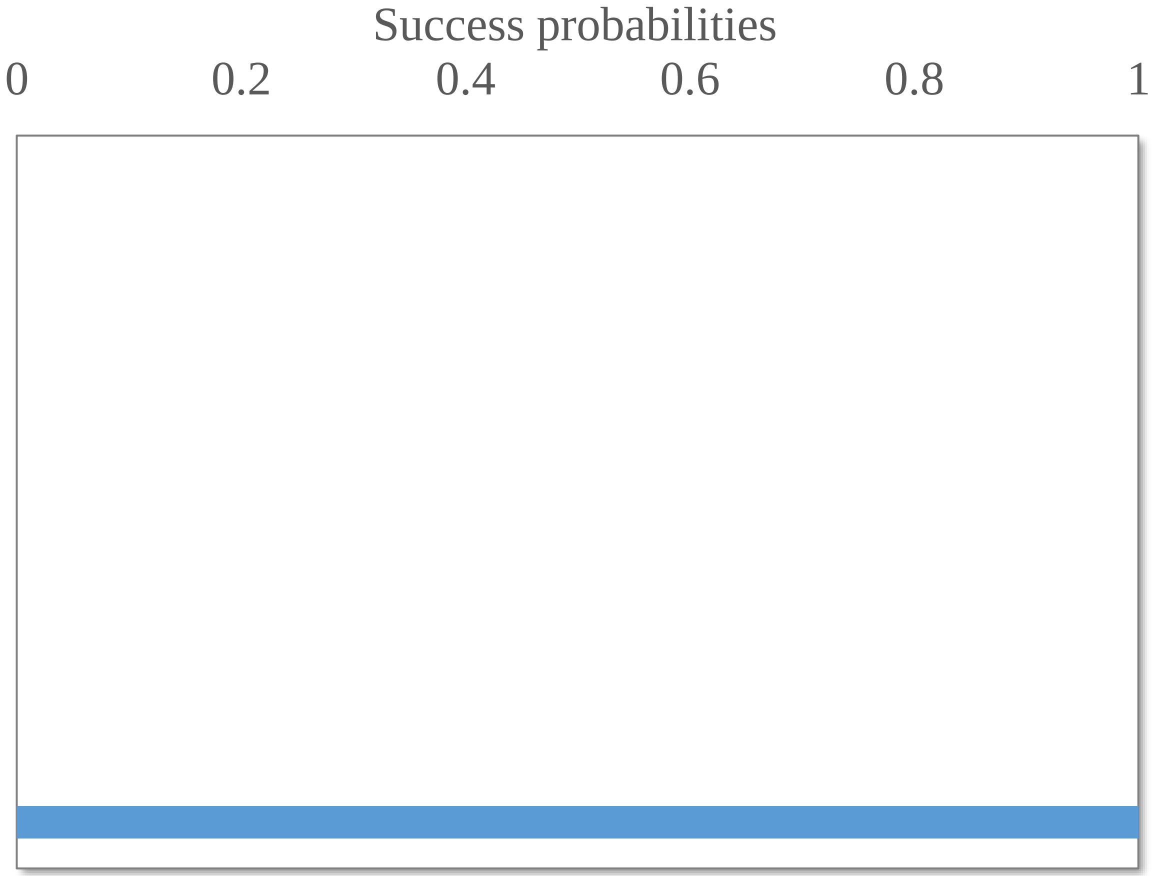
Nate and Dale are making s’mores. Nate has 4 marshmallows and 3 crackers. Dale has 7 marshmallows and twice as many crackers as Nate. Whose s’mores have a stronger marshmallow taste (greater marshmallows-to-crackers ratio)?

Nate and Dale are making s’mores. Nate has 4 marshmallows and 3 crackers. Dale has 7 marshmallows and twice as many crackers as Nate. Whose s’mores have a stronger marshmallow taste (greater marshmallows-to-crackers ratio)?

Prerequisite skills	Constructing ratios	Comparing fractions
0	0	0
1	0	0
0	1	0
0	0	1
1	1	0
0	1	1
1	0	1
1	1	1

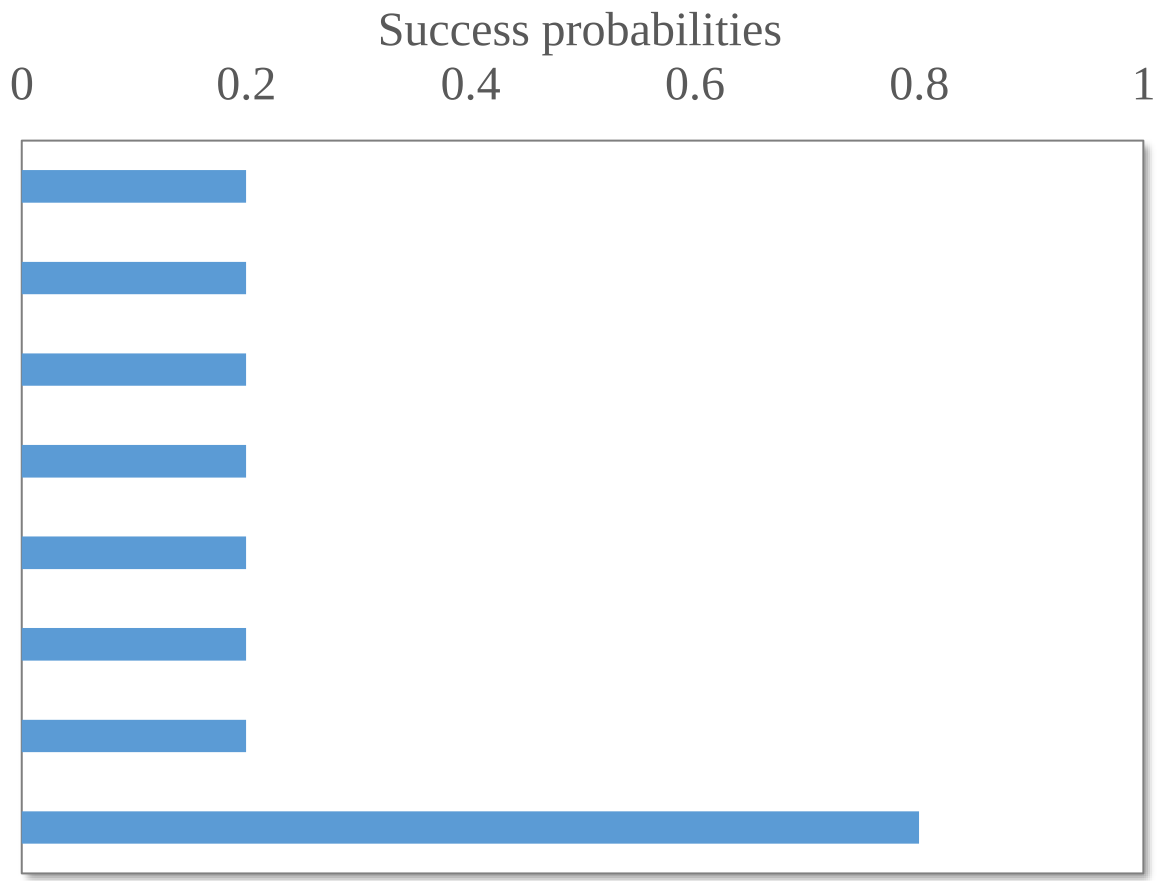
Nate and Dale are making s'mores. Nate has 4 marshmallows and 3 crackers. Dale has 7 marshmallows and twice as many crackers as Nate. Whose s'mores have a stronger marshmallow taste (greater marshmallows-to-crackers ratio)?

Prerequisite skills	Constructing ratios	Comparing fractions
0	0	0
1	0	0
0	1	0
0	0	1
1	1	0
0	1	1
1	0	1
1	1	1



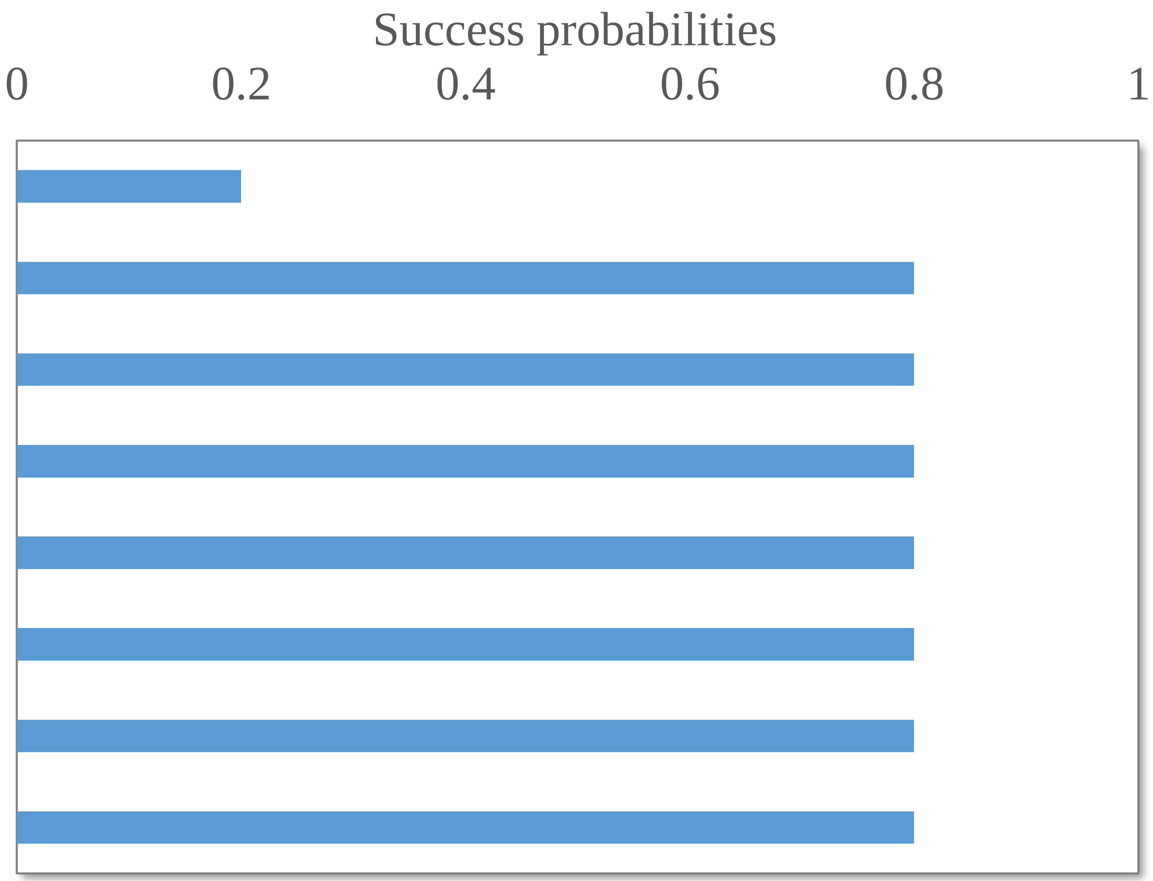
Nate and Dale are making s’mores. Nate has 4 marshmallows and 3 crackers. Dale has 7 marshmallows and twice as many crackers as Nate. Whose s’mores have a stronger marshmallow taste (greater marshmallows-to-crackers ratio)?

Prerequisite skills	Constructing ratios	Comparing fractions
0	0	0
1	0	0
0	1	0
0	0	1
1	1	0
0	1	1
1	0	1
1	1	1



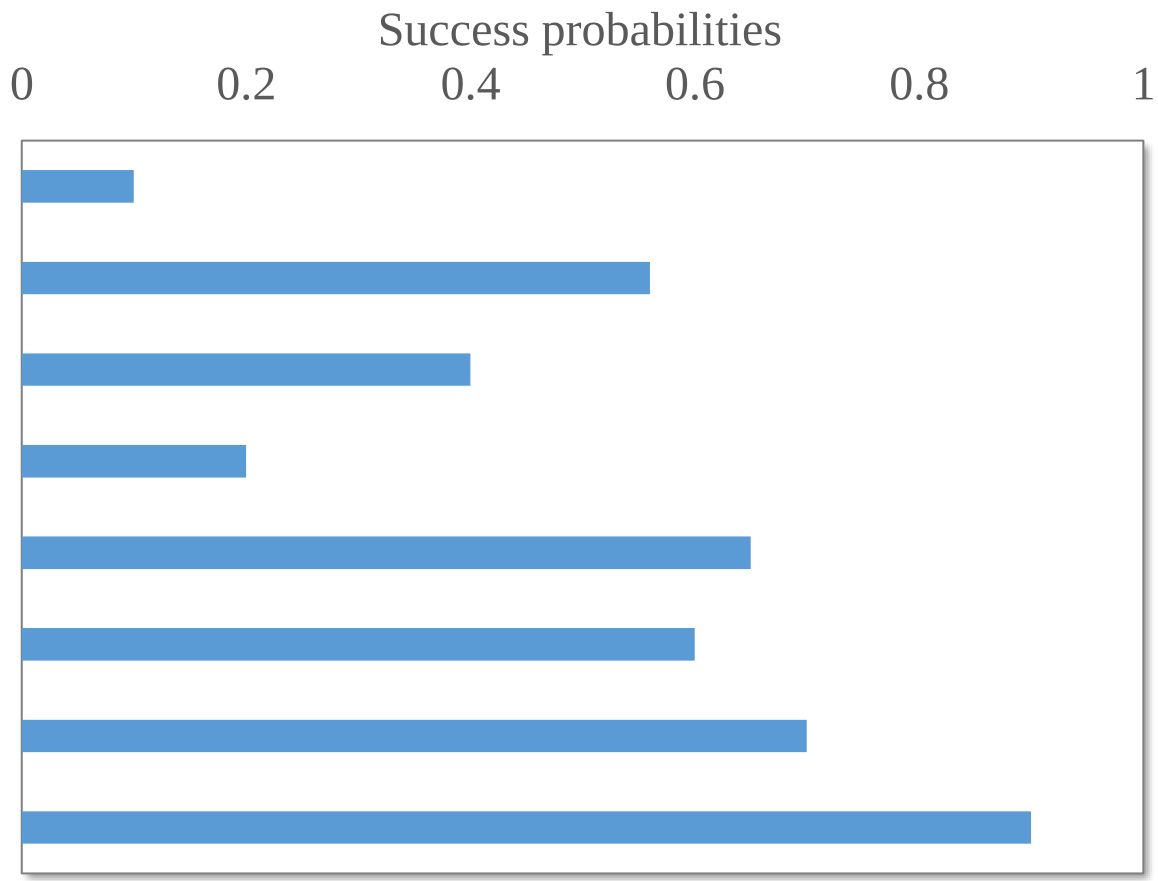
Nate and Dale are making s'mores. Nate has 4 marshmallows and 3 crackers. Dale has 7 marshmallows and twice as many crackers as Nate. Whose s'mores have a stronger marshmallow taste (greater marshmallows-to-crackers ratio)?

Prerequisite skills	Constructing ratios	Comparing fractions
0	0	0
1	0	0
0	1	0
0	0	1
1	1	0
0	1	1
1	0	1
1	1	1



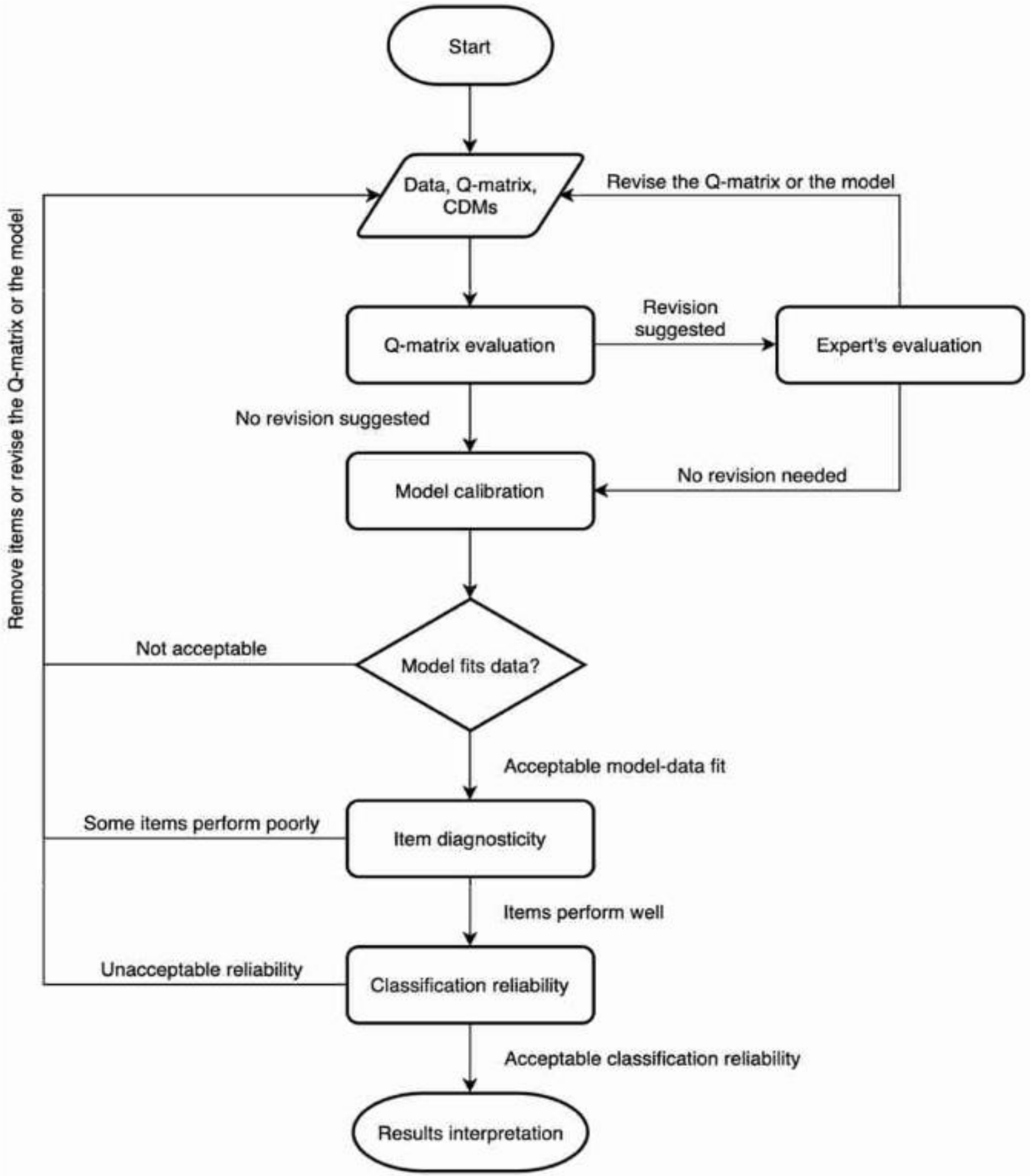
Nate and Dale are making s’mores. Nate has 4 marshmallows and 3 crackers. Dale has 7 marshmallows and twice as many crackers as Nate. Whose s’mores have a stronger marshmallow taste (greater marshmallows-to-crackers ratio)?

Prerequisite skills	Constructing ratios	Comparing fractions
0	0	0
1	0	0
0	1	0
0	0	1
1	1	0
0	1	1
1	0	1
1	1	1



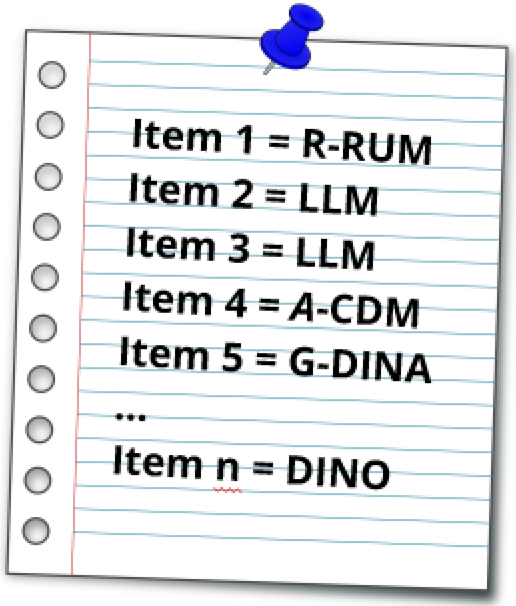
CDM Analysis Procedure

Shi, Q., Ma, W., Robitzsch, A., Sorrel, M. A., & Man, K. (2021). Cognitively Diagnostic Analysis Using the G-DINA Model in R. *Psych*, *3*(4), 812–835. https://doi.org/10.3390/psych3040052



Item-level model comparison

- **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

Model-data fit

- CDM or Q-matrix misspecifications can happen
- For CDM inferences to be valid, it is important to evaluate how well the model fits the data → absolute fit evaluation
- With the availability of various CDMs, it is also important to choose the most appropriate model → relative fit evaluation

Q-matrix validation

- Recall: A Q-matrix specifies which attributes are necessary for each item
- Most CDM analyses assume that the Q-matrix is correctly specified
- Thus, model misfit attributable to the Q-matrix are not addressed and remedied
- Q-matrix estimation vs. Q-matrix validation

Outline

- Introduction to Diagnostic Assessments
- Cognitive Diagnosis: Terminology, Inputs and Outputs
- Cognitive Diagnosis Basics Again: Attributes
- Cognitive Diagnosis: Models
- Cognitive Diagnosis: Applications and New Developments
- Cognitive Diagnosis in R

Methodological Developments

- The DINA, DINO and G-DINA models are most basic CDMs
- They have been used as building blocks for more complex models

Methodological Developments

- Dichotomous data → polytomous data
  - Constructed-response tasks
  - Likert-scale items
  - Multiple-choice questions with coded options

British Journal of Mathematical and Statistical Psychology (2016)

The British Psychological Society
www.wileyonlinelibrary.com

A sequential cognitive diagnosis model for polytomous responses

Wenchao Ma* and Jimmy de la Torre
Rutgers, The State University of New Jersey, New Brunswick, New Jersey, USA

This paper proposes a general polytomous cognitive diagnosis model for a special type of graded responses, where item categories are attained in a sequential manner, and associated with some attributes explicitly. To relate categories to attributes, a category-level Q-matrix is used. When the attribute and category association is specified *a priori*, the proposed model has the flexibility to allow different cognitive processes (e.g., conjunctive, disjunctive) to be modelled at different categories within a single item. This model can be extended for items where categories cannot be explicitly linked to attributes, and for items with unordered categories. The feasibility of the proposed model is examined using simulated data. The proposed model is illustrated using the data from the Trends in International Mathematics and Science Study 2007 assessment.

A general diagnostic classification model for rating scales

Ren Liu & Zhehan Jiang

Behavior Research Methods **52**, 422–439 (2020) | Cite this article

Abstract

This study proposes and evaluates a general diagnostic classification model (DCM) for rating scales. We applied the proposed model to a dataset to compare its performance with traditional DCMs for polytomous items. We also conducted a simulation study based on the applied study condition in order to evaluate the parameter recovery of the proposed model. The findings suggest that the proposed model shows promise for (1) accommodating much smaller sample sizes by reducing a large number of parameters for estimation; (2) obtaining item category response probabilities and individual scores very similar to those from a traditional saturated model; and (3) providing general item information that is not available in traditional DCMs for polytomous items.

Applied Psychological Measurement
Volume 33 Number 3
May 2009 163-183
© 2009 SAGE Publications
10.1177/0146621608320523
http://apm.sagepub.com
hosted at
http://online.sagepub.com

A Cognitive Diagnosis Model for Cognitively Based Multiple-Choice Options

Jimmy de la Torre
Rutgers, The State University of New Jersey

Cognitive or skills diagnosis models are discrete latent variable models developed specifically for the purpose of identifying the presence or absence of multiple fine-grained skills. However, applications of these models typically involve dichotomous or dichotomized data, including data from multiple-choice (MC) assessments that are scored as right or wrong. The dichotomization approach to the analysis of MC data ignores the potential diagnostic information that can be found in the distractors and is therefore deemed diagnostically suboptimal. To maximize the diagnostic value of MC assessments, this article prescribes how MC options should be constructed to make them more cognitively diagnostic and proposes a cognitive diagnosis model for analyzing such data. The article discusses the specification of the proposed model and estimation of its parameters. Moreover, results of a simulation study evaluating the viability of the model and an estimation algorithm are presented. Finally, practical considerations concerning the proposed framework are discussed.

Methodological Developments

- Dichotomous attributes → polytomous attributes → continuous attributes
  - Accommodate coarser-grained attributes
  - Improve model-data fit

Applied Psychological Measurement
37(6) 419–437

DOI: 10.1177/0146621613479818
apm.sagepub.com
SAGE

A General Cognitive Diagnosis Model for Expert-Defined Polytomous Attributes

Jinsong Chen[1] and Jimmy de la Torre[2]

Abstract

Polytomous attributes, particularly those defined as part of the test development process, can provide additional diagnostic information. The present research proposes the polytomous generalized deterministic inputs, noisy, "and" gate (pG-DINA) model to accommodate such attributes. The pG-DINA model allows input from substantive experts to specify attribute levels and is a general model that subsumes various reduced models. In addition to model formulation, the authors evaluate the viability of the proposed model by examining how well the model parameters can be estimated under various conditions, and compare its classification accuracy against that of the conventional G-DINA model with a modified classification rule. A real-data example is used to illustrate the application of the model in practice.

> Multivariate Behav Res. 2022 Mar-May;57(2-3):408-421. doi: 10.1080/00273171.2020.1860731. Epub 2021 Jan 12.

A Higher-Order Cognitive Diagnosis Model with Ordinal Attributes for Dichotomous Response Data

Wenchao Ma [1]

Affiliations + expand

PMID: 33434081 DOI: 10.1080/00273171.2020.1860731

Abstract

Most existing cognitive diagnosis models (CDMs) assume attributes are binary latent variables, which may be oversimplified in practice. This article introduces a higher-order CDM with ordinal attributes for dichotomous response data. The proposed model can either incorporate domain experts' knowledge or learn from the data empirically by regularizing model parameters. A sequential item response model was employed for joint attribute distribution to accommodate the sequential mastery mechanism. The expectation-maximization algorithm was employed for model estimation, and a simulation study was conducted to assess the recovery of model parameters. A set of real data was also analyzed to assess the viability of the proposed model in practice.

Partial-mastery cognitive diagnosis models

Zhuoran Shang, Elena A. Erosheva, Gongjun Xu

Author Affiliations +

Ann. Appl. Stat. 15(3): 1529-1555 (September 2021). DOI: 10.1214/21-AOAS1439

ABOUT | FIRST PAGE | CITED BY | REFERENCES | SUPPLEMENTAL CONTENT

Abstract

Cognitive diagnosis models (CDMs) are a family of discrete latent attribute models that serve as statistical basis in educational and psychological cognitive diagnosis assessments. CDMs aim to achieve fine-grained inference on individuals' latent attributes, based on their observed responses to a set of designed diagnostic items. In the literature CDMs usually assume that items require mastery of specific latent attributes and that each attribute is either fully mastered or not mastered by a given subject. We propose a new class of models, partial mastery CDMs (PM-CDMs), that generalizes CDMs by allowing for partial mastery levels for each attribute of interest. We demonstrate that PM-CDMs can be represented

Methodological Developments

- Cross-sectional data → longitudinal data
  - Monitor students' progress
  - Evaluating intervention effects
  - May provide more accurate estimation

Journal of Educational and Behavioral Statistics
Volume 43, Issue 1, February 2018, Pages 57-87

https://doi-org.libdata.lib.ua.edu/10.3102/1076998617719727

SAGE journals

Article

Tracking Skill Acquisition With Cognitive Diagnosis Models: A Higher-Order, Hidden Markov Model With Covariates

Shiyu Wang[1], Yan Yang, Steven Andrew Culpepper, and Jeffrey A. Douglas[2]

Abstract
A family of learning models that integrates a cognitive diagnostic model and a higher-order, hidden Markov model in one framework is proposed. This new framework includes covariates to model skill transition in the learning environment. A Bayesian formulation is adopted to estimate parameters from a learning model. The developed methods are applied to a computer-based assessment with a learning intervention. The results show the potential application of the proposed model to track the change of students' skills directly and provide immediate remediation as well as to evaluate the efficacy of different interventions by investigating how different types of learning interventions impact the transitions from nonmastery to mastery.

Assessing Growth in a Diagnostic Classification Model Framework

Matthew J. Madison & Laine P. Bradshaw

Psychometrika **83**, 963–990 (2018) | Cite this article

Abstract

A common assessment research design is the single-group pre-test/post-test design in which examinees are administered an assessment before instruction and then another assessment after instruction. In this type of study, the primary objective is to measure growth in examinees, individually and collectively. In an item response theory (IRT) framework, longitudinal IRT models can be used to assess growth in examinee ability over time. In a diagnostic classification model (DCM) framework, assessing growth translates to measuring changes in attribute mastery status over time, thereby providing a categorical, criterion-referenced interpretation of growth. This study introduces the Transition Diagnostic

AERA Open
January-December 2021, Vol. 7, No. 1, pp. 1–15
DOI: 10.1177/23328584211060804

Does Diagnostic Feedback Promote Learning? Evidence From a Longitudinal Cognitive Diagnostic Assessment

Fang Tang
Peida Zhan
Zhejiang Normal University

Assessment for learning emphasizes the importance of feedback to promote learning. To explore whether cognitive diagnostic feedback (CDF) promotes learning and whether it is more effective than traditional feedback in promoting learning, this study conducted a quasi-experiment by utilizing a longitudinal cognitive diagnostic assessment to compare the effect of three feedback modes on promoting learning, including CDF, correct–incorrect response feedback (CIRF), and no feedback. The results provided some evidence for the conclusion that CDF can promote students' learning and is more effective than CIRF in promoting learning, especially in more challenging areas of knowledge.

Methodological Developments

- Other model extensions
  - Single strategy → Multiple strategies
  - Confirmatory CDMs → Exploratory CDMs
  - Skills-based models → models for misconceptions, disengaged behaviors, etc.
  - Models with task response data → Models with process data
  - Parametric models → nonparametric approaches

Methodological Developments

- New estimation methods
  - for small samples
  - large number of attributes
- Conditions for identifiability
- Q-matrix estimation or validation
- Reliability estimation
- CD-CAT

Applications (Some examples)

- Math education
  - Tjoe, H., & de la Torre, J. (2014). The identification and validation process of proportional reasoning attributes: an application of a cognitive diagnosis modeling framework. Mathematics Education Research Journal, 26(2), 237–255. doi: 10.1007/s13394-013- 0090-7
- Science education
  - Zhai, X., Haudek, K. & Ma, W. (Accepted). Assessing argumentation using machine learning and cognitive diagnostic modeling. *Research in Science Education*. https://doi.org/10.1007/s11165-022-10062-w
- Language assessments
  - Lee, Y.-W., & Sawaki, Y. (2009b). Cognitive diagnosis approaches to language assessment: An overview. Language Assessment Quarterly, 6(3), 172–189. doi: 10.1080/15434300902985108
- Game-based assessments
  - Yu, J., Ma, W., Moon, J., & Denham, A. (Accepted). Developing a Stealth Assessment System Using a Continuous Conjunctive Model. *Journal of Learning Analytics*, *9*(3), 11-31. https://doi.org/10.18608/jla.2022.7639
- Mental health
  - Tan, Z.*, de la Torre, J., Ma, W., Huh, D., Larimer, M., & Mun, E-Y. (2022). A tutorial on cognitive diagnosis modeling for characterizing mental health symptom profiles using existing item responses. *Prevention Science*. http://doi.org/10.1007/s11121-022-01346-8
- I/O psychology
  - Sorrel, M. A., Olea, J., Abad, F. J., de La Torre, J., Aguado, D., & Lievens, F. (2016). Validity and reliability of situational judgement test scores. Organizational Research Methods, 19(3), 506–532. doi: 10.1177/1094428116630065
- Cognitive psychology
  - Wang S, Hu Y, Wang Q, Wu B, Shen Y and Carr M (2020) The Development of a Multidimensional Diagnostic Assessment With Learning Tools to Improve 3-D Mental Rotation Skills. *Front. Psychol.* 11:305.

Real-world applications

Real-world applications

NAVVY EDUCATION

About Us | Pandemic Assessment Solution | The Navvy System | Testimonials

about us

pandemic assessment solution

NAVVY

Standard-by-Standard Diagnostic Assessment System

What is Navvy?

Navvy is a student-friendly and technology-savvy formative assessment system that provides short, standard-by-standard assessments that are embedded in classroom practice and available on-demand. Navvy provides immediate, actionable results to inform personalized instruction and successfully navigate each student's learning journey.

Innovation Approved!

Navvy's new way to approach assessment in schools has been approved by the US Department of Education and Georgia's State Board of Education as an innovative assessment system being implemented under USED's Innovative Assessment Demonstration Authority. Learn more below!

Navvy Competency Checks

Students earn a digital microcertification for each standard they show they have learned on a short Navvy Competency Check. Students have multiple opportunities to earn a microcert for each standard throughout the year, and Navvy keeps learning profiles up-to-date on the student's Learning Map as the student re-assesses. Learning Maps allow for standards-level within- and across-grade progress monitoring of student competencies.

Navvy Quick Checks

Navvy provides a complete suite of standard-by-standard practice assessments called Quick Checks. Navvy Quick Checks can be given remotely for both prior-grade and on-grade standards, providing a quick way to identify unfinished learning from prior grades or gauge the progress of an on-grade standard.

Navvy Writing Checks

Students earn microcerts for ELA writing and language standards through formative and/or summative Navvy Writing Checks. Students provide an extended response to a writing task, and Navvy's team of raters provides standard-by-standard competency diagnoses.

Real-world applications

ALEKS K-12 Higher Ed Independent Use Course Products About Search

Overview | Company Info | Leadership | Awards | Major New Releases

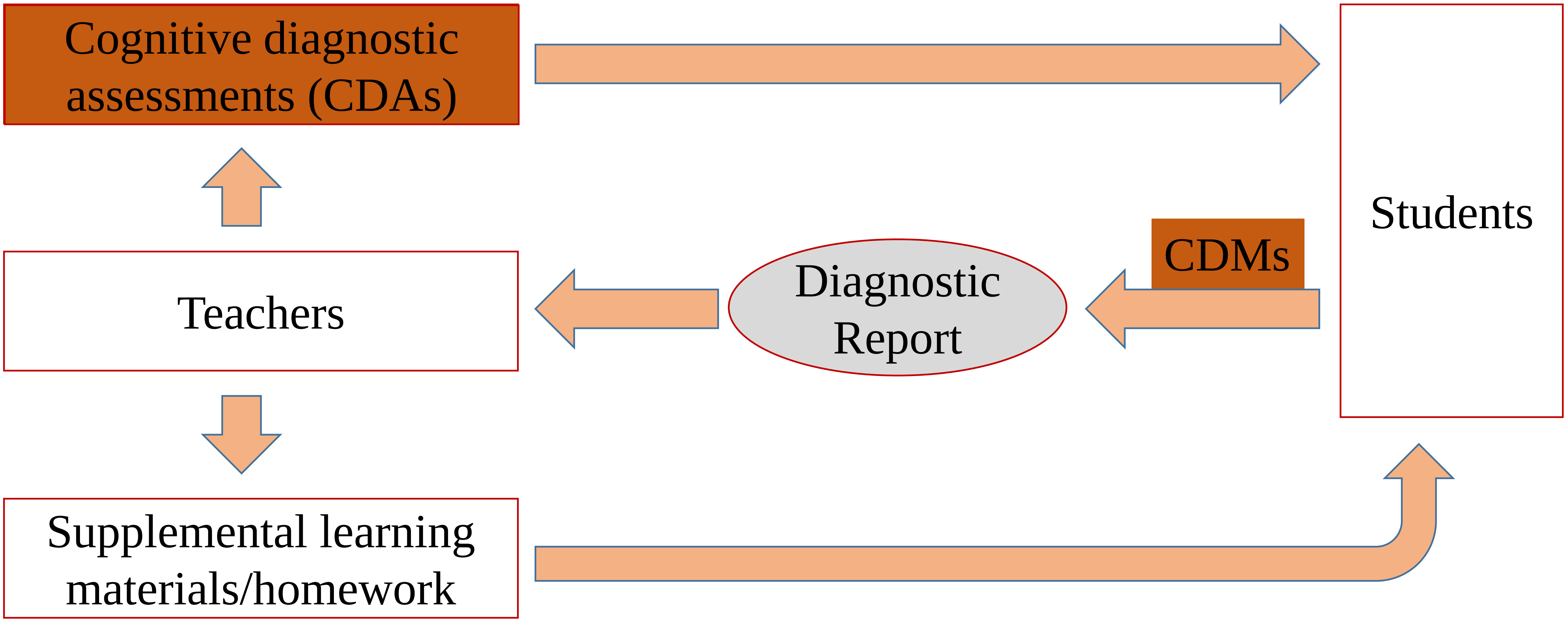
ALEKS is based on Learning Spaces, a type of *Knowledge Space*. A *Knowledge Space* is a representation of a domain of knowledge (such as Algebra 1) as a combinatorial structure that delineates the combinations of elements of knowledge (problem types in Algebra, for example) that comprise all the feasible states of knowledge of individual students. A student's *knowledge state* is the complete set of problems that the individual student is capable of solving in the particular domain of knowledge. For example, Algebra 1 is regarded as a domain of roughly 400-500 core problem types - giving rise to a *knowledge space* of a few trillion empirically feasible states of knowledge. That is, each Algebra 1 student could be in any one of a few trillion feasible knowledge states.

In ALEKS, mathematically rigorous theory facilitates the development of computer algorithms for the construction and mapping of knowledge spaces. This enables ALEKS machine learning software to comprehensively investigate trillions of potential knowledge states to accurately diagnose each individual student's precise knowledge of the subject, and what that individual student is currently ready to learn. Even with such a large number of knowledge states in the knowledge space, the ALEKS adaptive assessment is nevertheless able to rapidly and efficiently assess a particular student's knowledge after the student has answered only 20-30 questions.

The outcome of an ALEKS assessment consists in (i) the precise and comprehensive description of an individual student's competence in a particular subject in the form of a knowledge state that describes the problem types already mastered by that individual student, and (ii) the problem types that the individual is ready to learn next.

Real world applications

Summary



Cognitive diagnostic assessments (CDAs)
Students
CDMs
Diagnostic Report
Teachers
Supplemental learning materials/homework

Outline

- Introduction to Diagnostic Assessments
- Cognitive Diagnosis: Terminology, Inputs and Outputs
- Cognitive Diagnosis Basics Again: Attributes
- Cognitive Diagnosis: Models
- Cognitive Diagnosis: Applications and New Developments
- Cognitive Diagnosis in R

Software Programs for Diagnostic Modeling

- Commercial
  - Mplus (see Templin, & Hoffman, 2013)
  - LatentGold (see DeCarlo, 2010)
  - FlexMIRT (Houts, & Cai, 2015)

- Free R packages
  - CDM (George, et al, 2016)
  - GDINA (Ma, & de la Torre, 2020)

CDM: Cognitive Diagnosis Modeling

- Developed by Alexander Robitzsch, Thomas Kiefer, Ann Cathrice George, and Ali Uenlue
- Functions for cognitive diagnosis modeling and multidimensional item response modeling for dichotomous and polytomous item responses. This package enables the estimation of the DINA and DINO model, the multiple group (polytomous) GDINA model, the multiple choice DINA model, the general diagnostic model, the structured latent class model and regularized latent class analysis.
- Website: https://cran.r-project.org/web/packages/CDM/index.html

GDINA: The Generalized DINA Model Framework

- Developed by Wenchao Ma, Jimmy de la Torre, Miguel Sorrel and Zhehan Jiang
- Package website: https://wenchao-ma.github.io/GDINA/
- Source code: https://github.com/Wenchao-Ma/GDINA/

GDINA: The Generalized DINA Model Framework

- Estimating G-DINA model and a variety of widely-used models subsumed by the G-DINA model, including the DINA model, DINO model, additive-CDM (A-CDM), linear logistic model (LLM), reduced reparametrized unified model (RRUM), multiple-strategy DINA model for dichotomous responses
- Estimating Bugs models for dichotomous responses
- Estimating sequential G-DINA model for ordinal and nominal responses
- Estimating the generalized multiple-strategy cognitive diagnosis models and diagnostic tree model for multiple strategies
- Estimating multiple-choice CDMs
- Accommodating multiple-group model analysis
- Accommodating binary and polytomous attributes
- Validating Q-matrix under the general model framework
- Evaluating absolute and relative item and model fit
- Comparing models at the test and item levels
- Detecting differential item functioning using Wald and likelihood ratio test
- Providing graphical user interface for users less familiar with R

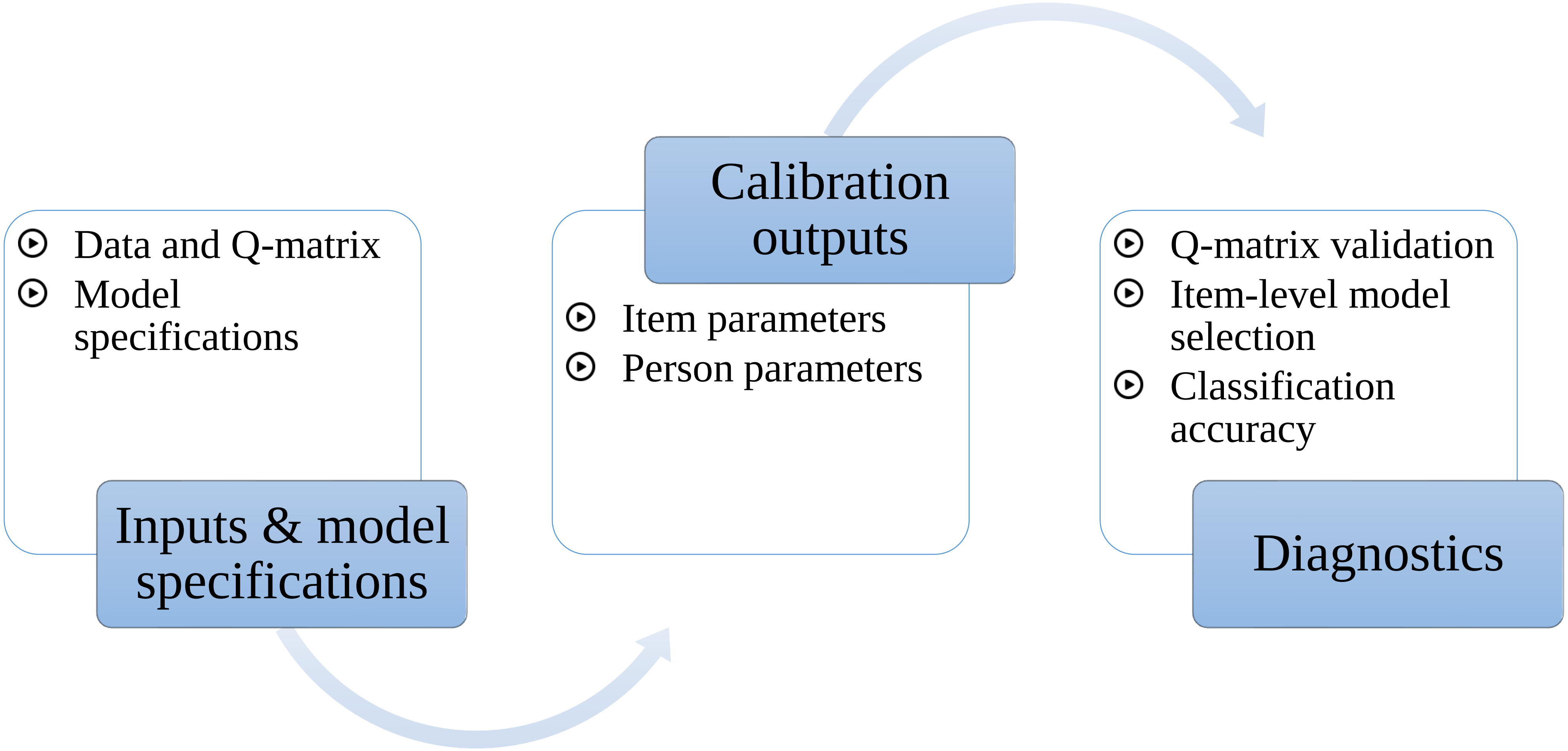
A Demo using GDINA R package

- If you have a computer and would like to try, please download data and Q-matrix from:

wenchaoma.people.ua.edu/downloads

The slides can also be found from the website above.

- The data with responses of 837 students to 15 items were simulated based on a subset of the proportional reasoning test data.
- The test measures 3 attributes.
- We are going to use GDINA R package for illustration.


Data and Q-matrix
Model specifications
Inputs & model specifications
Calibration outputs
Item parameters
Person parameters
Q-matrix validation
Item-level model selection
Classification accuracy
Diagnostics

Exercise

- Please use data2 and Q2 for the following exercises.
- For Item 15, the G-DINA model parameter estimates are P(100) = 0.4___57, P(101) = 0.6___96.
- The proportion of individuals having an attribute pattern of 111 in the population is estimated to be 0.3__12
- The proportion of individuals who master ***a***1 in the population is estimated to be 0.80___7.
- Plot item success probabilities for Items 5, 8 and 15. Which item(s) appear to follow the DINO models?
- Find the EAP estimate of attribute pattern for the third individual.


Thank you!
Questions?