Non-identifiability, equivalence classes, and attribute-specific classification in Q-matrix based Cognitive Diagnosis Models

Stephanie S. Zhang, Lawrence T. DeCarlo, and Zhiliang Ying
Columbia University

March 5, 2013

Abstract

There has been growing interest in recent years in Q-matrix based cognitive diagnosis models. Parameter estimation and respondent classification under these models may suffer due to identifiability issues. Non-identifiability can be described by a partition separating attribute profiles into groups of those with identical likelihoods. Marginal identifiability concerns the identifiability of individual attributes. Maximum likelihood estimation of the proportion of respondents within each equivalence class is consistent, making possible a new measure of assessment quality reporting the proportion of respondents for whom each individual attribute is marginally identifiable. Arising from this is a new posterior-based classification method adjusting for non-identifiability.

Keywords: CDM, diagnostic classification, DINA, DINO, NIAD-DINA, Q-matrix, consistency, identifiability

1 Introduction

Diagnostic assessments are created with the goal of making classification-based decisions about respondents' possession of multiple latent traits, also known as attributes. Researchers have brought a number of tools to bear on the problem of diagnostic classification, including multidimensional IRT, factor analysis, the rule-space method, the attribute hierarchy method, clustering methods, and cognitive diagnosis models (CDMs); for a recent review, see Rupp, Templin, and Henson (2010). CDMs are multidimensional latent variable models with a vector of binary latent variables representing mastery of a finite set of skills whose analysis results in a probabilistic attribute profile; this makes them well-suited to diagnostic classification. Well known models include the Deterministic Input, Noisy "And" Gate (DINA) model, the Deterministic Input, Noisy "Or" Gate (DINO) model, the Noisy Inputs, Deterministic "And" Gate (NIDA) model,

1

the Noisy Inputs, Deterministic "Or" Gate (NIDO) model, and the Conjunctive Reparameterized Unified Model (C-RUM), among others (Rupp et al., 2010; Haertel, 1989; Junker & Sijtsma, 2001; de la Torre & Douglas, 2004; Maris, 1999; Templin & Henson, 2006; C. Tatsuoka, 2002; Templin, 2006; de la Torre, 2008).

The DINA model is one of the best known and widely used CDMs. Underlying the model is the assumption that, before slipping and guessing come into play, a respondent must have mastered all necessary (as specified by a loading matrix known as the Q-matrix) attributes required by a particular item in order to answer that item correctly. Thus it is a conjunctive, non-compensatory model, well-suited to educational assessments in areas such as mathematics where correct answers are obtained by correctly employing all of an item's required skills together. The DINA model has been frequently employed in the analysis of assessments, including the widely analyzed fraction subtraction data set of K. K. Tatsuoka (1990) (see de la Torre, 2009; de la Torre & Douglas, 2004, 2008; Templin, Henson, & Douglas, 2006; DeCarlo, 2011; Henson, Templin, & Willse, 2009). However, even after many refinements to the methodology there are still some persistent issues. In the fraction subtraction data, for example, respondents who answer all items incorrectly are often classified as having most of the skills (DeCarlo, 2011). Classification issues of this type can result model misspecification, but they can also be the unavoidable consequence of the structure of the assessment. Specifically in the DINA model, attributes that appear solely in conjunction with other attributes are problematic (DeCarlo, 2011). This is due to an issue with attribute identifiability, which has long been known (K. K. Tatsuoka, 1991; DiBello, Stout, & Roussos, 1995; DeCarlo, 2011), but tends to be ignored in practice.

This paper gives a formal treatment of the identifiability issue of the DINA model and related CDMs. Since an assessment with fully identifiable attributes is often unavailable, we include guidelines for classification under non-identifiability and a consistent measure for the extent of non-identifiability. This allows classification error control and assessment evaluation in terms of identifiability.

The paper begins by reviewing some basic concepts, including the DINA model and its variants, in Section 2. We introduce the issue of identifiability for Q-matrix based assessments in Section 3. In Section 4, we explain the use of equivalence classes and partitions to fully describe the structure of the attribute profile space, in terms of identifiability; an algorithm to generate the partition, given any Q-matrix, is included. Partitioning allows consistent estimation of the proportion of individuals in each group of equivalent attribute profiles, as explained in Section 5. These results are extended to individual attributes via marginal identifiability in Section 6. In fact, the consistent estimation of the proportion of the population for which each attribute is marginally identifiable leads to a reliable measure of exam quality (with respect to identifiability). We also create a decision rule for respondent classification which controls misclassification probabilities in Section 7. Section 8 examines the implications of these methods to several variants of the DINA, in addition to another Q-matrix based CDM, the DINO model. Finally, results derived from both simulation

2

and K. K. Tatsuoka's fractions data set are reported in Section 9.

2 Background

Throughout the paper, we will be using some standard concepts from the study of CDMs. Some specific terminology and notations are listed below.

Attributes are the respondent's (unobserved) mastery of certain skills. If we suppose that there are N respondents and K attributes, let the matrix of attributes be $A = (\alpha_{ik})$, where where $\alpha_{ik} \in \{0, 1\}$ indicates the presence or absence of the k-th attribute in the i-th respondent. An *attribute profile* $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K)^\top$ is the vector of all attributes; an individual respondent i will have attribute profile $\boldsymbol{\alpha}^i$ such that $\alpha_k^i = \alpha_{ik}$.

Responses are the respondent's binary responses to items. Given N respondents and J items, the responses can be written as a $N \times J$ matrix $X = (X_{ij})$, where $X_{ij} \in \{0, 1\}$ is the response of the i-th respondent to the j-th item. The i-th respondent's responses will be denoted by the vector $\mathbf{X}^i$, where the j-th element $X_j^i = X_{ij}$ for all i, j.

The Q-matrix is the link between the items and their attribute requirements. It is a $J \times K$ matrix $Q = (q_{jk})$, where for each j, k, $q_{jk} \in \{0, 1\}$ indicates whether the j-th item requires the k-th attribute. From the Q-matrix we can extract the attribute requirements of an item j as the vector $\mathbf{q}^j$, where the k-th element $q_k^j = q_{jk}$ for all j, k.

2.1 The DINA Model

This paper focuses on the DINA model, one of the most widely used CDMs. Under the DINA model, given an attribute profile $\boldsymbol{\alpha}$ and a Q-matrix Q, we can further define the quantity

$$\xi_j(Q, \boldsymbol{\alpha}) = \prod_{k=1}^{K} (\alpha_k)^{q_{jk}} = \mathbf{1}(\alpha_k \geq q_{jk} : k = 1, \ldots, K),$$

which indicates whether a respondent with attribute profile $\boldsymbol{\alpha}$ possesses all the attributes required for item j. If we suppose no uncertainty in the response, then a respondent i with attribute profile $\boldsymbol{\alpha}$ will have responses $X_{ij} = \xi_j(Q, \boldsymbol{\alpha})$ for $j = 1, \ldots, J$. Thus, the vector $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_J)^\top$ is known as the *ideal response vector*.

In the DINA model, uncertainty is incorporated at the item level. With each item $j = 1, \ldots, J$, we associate a slipping parameter $s_j = P(X_j = 0 | \xi_j = 1)$ and a guessing parameter $g_j = P(X_j = 1 | \xi_j = 0)$. Each X_j is Bernoulli with success probability $(1 - s_j)^{\xi_j} g_j^{1 - \xi_j}$. Thus, the probability of a particular response vector

$\mathbf{x}$ given the ideal response vector $\boldsymbol{\xi}$ is

$$P(\mathbf{x}|\boldsymbol{\xi},\mathbf{s},\mathbf{g}) = \prod_{j=1}^{J}[(1-s_j)^{\xi_j}g_j^{1-\xi_j}]^{x_j}[1-(1-s_j)^{\xi_j}g_j^{1-\xi_j}]^{1-x_j}$$

$$= \prod_{j=1}^{J}(1-s_j)^{\xi_j x_j}g_j^{(1-\xi_j)x_j}s_j^{\xi_j(1-x_j)}(1-g_j)^{(1-\xi_j)(1-x_j)} \qquad (1)$$

In addition to $\mathbf{s}$ and $\mathbf{g}$, the response distribution also depends on $\boldsymbol{\nu} = (\nu_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}\in\{0,1\}^K}$, the proportion of individuals possessing each attribute profile. Generally, diagnostic classification is based on the posterior $p(\boldsymbol{\alpha}|\mathbf{x})$, which is calculated using and can be very sensitive to the prior, $\boldsymbol{\nu}$. When $\mathbf{s}$, $\mathbf{g}$, and $\boldsymbol{\nu}$ are unknown, these parameters must be simultaneously estimated.

2.2   Variants of the DINA Model

Several variants of the DINA can be constructed by restricting $\boldsymbol{\nu}$ to some lower-dimensional subspace. For example, assuming independence among the attributes so that

$$\nu_{\boldsymbol{\alpha}} = p(\boldsymbol{\alpha}) = \prod_{k=1}^{K}p(\alpha_k)$$

reduces the 2^K-1-dimensional parameter space to a K-dimensional one. This restriction is referred to as the independent DINA (ind-DINA) from hereon. It is convenient to model each α_k with a logistic link, so that

$$p(\alpha_k) = \exp(\alpha_k b_k)/[1+\exp(b_k)],$$

where b_k denotes the attribute's 'difficulty.'

Another alternative is the higher-order DINA (HO-DINA) model (de la Torre & Douglas, 2004; Templin, Henson, Templin, & Roussos, 2008). This model assumes that the probability of possessing a skill is dependent on a continuous skill factor θ following the standard normal distribution, so that

$$\nu_{\boldsymbol{\alpha}} = p(\boldsymbol{\alpha}) = \int_{\theta}p(\boldsymbol{\alpha}|\theta)p(\theta)d\theta.$$

Each individual attribute is assumed to be conditionally independent given θ, so that

$$p(\boldsymbol{\alpha}|\theta) = \prod_{k=1}^{K}p(\alpha_k|\theta).$$

Finally the individual probabilities $p(\alpha_k|\theta)$ can be modeled with a logistic link,

$$p(\alpha_k|\theta) = \exp(\alpha_k(b_k+a_k\theta))/[1+\exp(b_k+a_k\theta)],$$

where b_k denotes the attribute's 'difficulty,' and a_k is the attribute discrimina-

tion parameter. It is also possible to fit a restricted version of this model, for which all the a_k must be equal, as in de la Torre and Douglas (2004). This is referred to as the restricted higher order DINA (RHO-DINA) model (DeCarlo, 2011).

3 Identifiability Issues in the DINA

Diagnostic assessments are meant to provide detailed information about respondents' possession of a variety of traits. Preferably, a well-designed exam will be able to provide information about each trait for every respondent. However, recovering information about the latent variables from a '0' response may be difficult; in comparison to a '1' response, which suggests that a respondent is more likely to possess each attribute associated with that item, a '0' response may indicate the failure to master any one or several of the required attributes. Consider the following two simple Q-matrices for the DINA model:

$$Q_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ Q_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}. \tag{2}$$

In assessments based on the Q-matrix Q_1, a correct response to each item generally indicates a higher probability that the respondent possesses the corresponding attribute, while an incorrect response indicates a lower probability of the same. However, with Q_2, an incorrect response to the second item only implies that at least one of the attributes is probably missing. In fact, given that a student does not possess Attribute 1, Item 2 provides no information about his or her mastery of Attribute 2, and so respondents with attribute profiles $(0, 0)$ and $(0, 1)$ have statistically identical responses. Thus, the assessment as a whole is incapable of differentiating between the two profiles, and any classification decision between them will solely be a reflection of the prior information.

A slightly more complicated situation appears if we add a third attribute to the example above. Consider an assessment following the DINA model with Q-matrix Q_3, where

$$Q_3 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}. \tag{3}$$

The attribute requirements of the first two items match those of the items corresponding to Q_2. Now, however, the proportion of individuals for whom Attribute 2 is not identifiable is smaller. Of those who do not possess Attribute 1, some will possess Attribute 3. Then Attribute 2 is identifiable because of differing response distributions on Item 3. However, response distributions for those with attribute profiles $(0, 1, 0)$ and $(0, 0, 0)$ are still indistinguishable. Thus, although the assessment provides no information about Attribute 2 for a smaller part of the population, the issue has not been completely resolved.

4 Partitioning the Attribute Profile Space

We begin with an intuitive criterion for deciding whether an assessment has the ability to differentiate between two attribute profiles.

Definition 1. *Two attribute profiles are* separable *if they lead to different response distributions.*

The differing response distributions of separable attribute profiles imply that the data will favor one profile or the other; there is some differential effect on the likelihood and thus the posterior. Profiles that are not separable are statistically identical, with equivalent likelihood functions, making any differences in their posteriors simply artifacts of the prior.

Determining whether attributes are separable can be done without the full response distribution;in fact, only the ideal responses $\boldsymbol{\xi}(Q, \boldsymbol{\alpha})$ are necessary.

Proposition 1. *Given a Q-matrix Q and slipping and guessing parameters* **s** *and* **g***, two attribute profiles $\boldsymbol{\alpha}^1$ and $\boldsymbol{\alpha}^2$ can be separated if and only if they produce ideal response vectors $\boldsymbol{\xi}^1 = \boldsymbol{\xi}(Q, \boldsymbol{\alpha}^1)$ and $\boldsymbol{\xi}^2 = (Q, \boldsymbol{\alpha}^2)$ such that for some $j \in \{1, \ldots, J\}$, $\xi_j^1 \neq \xi_j^2$ and $1 - s_j \neq g_j$.*

Throughout the rest of this paper we assume that $1 - s_j \neq g_j$ for each $j = 1, \ldots, J$, which simplifies Proposition 1 into Corollary 1. Should such an item indeed be present, then it has no discriminating power and may be omitted.

Corollary 1. *If every item j has different success probabilities given $\xi_j = 1$ or given $\xi_j = 0$, i.e. $1 - s_j \neq g_j$ for $j = 1, \ldots, J$, then two attribute profiles can be separated if and only if they produce different ideal response vectors.*

Lastly, it is also of interest whether an attribute profile can be separated from all other attribute profiles, and is thus identifiable. This definition of identifiability will be tied to the general statistical concept in Section 5.

Definition 2. *An attribute profile $\boldsymbol{\alpha}$ is* identifiable *when it can be separated from any other attribute profile $\boldsymbol{\alpha}' \neq \boldsymbol{\alpha}$.*

4.1 Complete Separation of Attribute Profiles

The first step in understanding the identifiability issue is determining under what circumstances all attribute profiles are identifiable. This depends on the Q-matrix, which is called complete when it leads to full identifiability(Chiu, Douglas, & Li, 2009). Formally, we have the following definition:

Definition 3. *Under a* complete *Q-matrix, all attribute profiles are identifiable, i.e. $\boldsymbol{\xi}(Q, \boldsymbol{\alpha}) \neq \boldsymbol{\xi}(Q, \boldsymbol{\alpha}')$ iff $\boldsymbol{\alpha} \neq \boldsymbol{\alpha}'$.*

The requirements for completeness have long been known (K. K. Tatsuoka, 1991; DiBello et al., 1995; Chiu et al., 2009). In essence, the assessment must contain at least one item devoted solely to each attribute. In terms of the Q-matrix, this means that for each $k \in \{1, \ldots, K\}$, there should be at least one row with an entry of '1' solely in the k-th position.

Proposition 2. *Let R_Q be the set of row vectors of Q-matrix Q. Then Q is complete iff $\{e_k : k = 1, \ldots, K\} \subset R_Q$, where e_k is a vector such that the k-th element is one and all other elements are zero.*

4.2 Partial Separation of Attribute Profiles

While a complete Q-matrix is necessary for full identifiability, many of the Q-matrices used in practice are unfortunately incomplete. In fact, creating assessments with complete Q-matrices is oftentimes infeasible, and requiring a complete matrix for analysis would make models like the DINA model highly impractical. This makes the partition, a standard mathematical construct, an essential tool in accurately and systematically describe the structure of nonidentifiability in the DINA. Partitions are formed from equivalence relations, which have the following requirements:

Definition 4. *The relation 'a $\sim$ b' is an* equivalence relation *if it is*

- *reflexive: $a \sim a$*

- *symmetric: $a \sim b$ iff $b \sim a$*

- *transitive: If $a \sim b$ and $b \sim c$, then $a \sim c$.*

Proposition 3. *Let '$\sim$' denote the binary relation 'cannot be separated,' where $\boldsymbol{\alpha}^1 \sim \boldsymbol{\alpha}^2$ if and only if $\boldsymbol{\xi}(Q, \boldsymbol{\alpha}^1) = \boldsymbol{\xi}(Q, \boldsymbol{\alpha}^2)$. Then '$\sim$' is an equivalence relation.*

Putting profiles into groups, known as equivalence classes, based on an equivalence relation results in a partition; in this case, any two attribute profiles in the same equivalence class cannot be separated, while any two in different classes can be. We denote a particular equivalence class by $[\boldsymbol{\alpha}]$, where $\boldsymbol{\alpha}$ may be any attribute profile in the class; literally, $[\boldsymbol{\alpha}]$ can be read as "the set of attribute profiles equivalent to $\boldsymbol{\alpha}$."

The simplest way of determining the partition would be to calculate the ideal response vector of each of the 2^K attribute profiles and sort them lexicographically. This runs quickly in $\mathcal{O}(JK \cdot 2^K)$ time (refer to Table 1 for the step-by-step algorithm). For an alternative algorithm using Boolean algebra, see K. K. Tatsuoka (1991). Note that our algorithm results in equivalence classes labeled by their smallest member, which shall be called the *minimal representative*. The minimal representative has additional meaning as the attribute requirements of the corresponding ideal response vector and is therefore the most preferable member for labeling.

As seen in Table 2, performing the algorithm on the 3×3 Q-matrix Q_3 from (3) results in five different equivalence classes, each of which is labeled with by its minimal representative: $[000] = \{000, 010, 001\}$, $[011] = \{011\}$, $[100] = \{100, 101\}$, $[110] = \{110\}$, and $[111] = \{111\}$. Note that since the bracket notation may be read as 'the equivalence class containing,' it is possible to change the labeling of each equivalence class by choosing any other member as the titular profile: $[000]$, $[010]$, and $[001]$ all refer to the same equivalence class, for example.

Table 1: Algorithm for partitioning an attribute profile space

Step	Procedure
Input:	A $J \times K$ Q-matrix Q.
(0)	(optional) Remove items with duplicate attribute requirements
(1)	List all 2^K attribute profiles $\boldsymbol{\alpha}$.
(2)	Find the ideal response vector $\boldsymbol{\xi}(Q, \boldsymbol{\alpha})$ for each $\boldsymbol{\alpha}$.
(3)	Do a lexicographic (alphabetic) sort of the ideal response vectors.
(4)	Check whether each successive profile has the same ideal response vector as the previous profile. If so, $\boldsymbol{\alpha}$ is the first member of a new equivalence class $[\boldsymbol{\alpha}]$. Else, $\boldsymbol{\alpha}$ is part of the current equivalence class.
Output:	A list of equivalence classes $[\boldsymbol{\alpha}]$ and their members.

Table 2: Generating the partition associated with the Q-matrix Q_3.

Q_3		$\boldsymbol{\alpha}$	$\boldsymbol{\xi}(Q_3, \boldsymbol{\alpha})$		$\boldsymbol{\alpha}$	$\boldsymbol{\xi}(Q_3, \boldsymbol{\alpha})$		$[\boldsymbol{\alpha}]$
		000	000		000	000		[000]
		100	100		010	000		
		010	000		001	000		
$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$	$\xrightarrow{(1),(2)}$	001	000	$\xrightarrow{(3)}$	011	001	$\xrightarrow{(4)}$	[011]
		110	110		100	100		[100]
		101	100		101	100		
		011	001		110	110		[110]
		111	111		111	111		[111]

Steps from Table 1 labeled (1), (2), (3), and (4).

5 Consistent Estimation

We now consider the problem of parameter estimation, specifically that of $\nu_{\boldsymbol{\alpha}}$, the proportion of the population possessing each attribute profile $\boldsymbol{\alpha}$. Unless $\boldsymbol{\nu}$ is assumed known, its consistent estimation has important consequences for respondent classification and exam validity. Unfortunately, when an assessment's Q-matrix is incomplete, it is impossible to consistently estimate $\boldsymbol{\nu}$. For each equivalence class $[\boldsymbol{\alpha}]$, let $\nu_{[\boldsymbol{\alpha}]}$ be the proportion of the population possessing an attribute profile within that equivalence class. Then,

$$\nu_{[\boldsymbol{\alpha}]} = \sum_{\boldsymbol{\alpha}' \in [\boldsymbol{\alpha}]} \nu_{\boldsymbol{\alpha}'}. \tag{4}$$

The probability of observing any particular set of data depends only on $\nu_{[\boldsymbol{\alpha}]}$, since the probability of any response depends only on equivalence class membership, not on the respondent's possession of a specific profile. With an incomplete Q-matrix it is possible to observe populations with different distributions $\boldsymbol{\nu}^1 \not\equiv \boldsymbol{\nu}^2$ over the attribute profile space that have identical distributions over the equivalence classes $[\boldsymbol{\alpha}]$, i.e., $\nu_{[\boldsymbol{\alpha}]}^1 = \nu_{[\boldsymbol{\alpha}]}^2$ for all $\boldsymbol{\alpha} \in \{0,1\}^K$, and thus identical response distributions. The phenomenon where different parameter values lead to identical response distributions is generally known as non-identifiability, and it destroys the ability of likelihood-based estimation methods to achieve consistency.

While consistent estimation of $\nu_{\boldsymbol{\alpha}}$ cannot be achieved, it is possible to consistently estimate the proportion of individuals within each equivalence class $[\boldsymbol{\alpha}]$.

Theorem 1. *Suppose an assessment follows the DINA model, with known Q-matrix Q and item parameters $\mathbf{s}$ and $\mathbf{g}$. Let $\nu_{[\boldsymbol{\alpha}]}$, representing the proportion of the population possessing an attribute profile $\boldsymbol{\alpha}' \in [\boldsymbol{\alpha}]$, be defined as in (4), and let the population parameter $\boldsymbol{\nu}$ be the vector of all $\nu_{[\boldsymbol{\alpha}]}$. We may write its likelihood as*

$$L(\boldsymbol{\nu}) = p(X|\boldsymbol{\nu}) = \prod_{i=1}^{N} p(\mathbf{x}^i|\boldsymbol{\nu}) = \prod_{i=1}^{N} \sum_{[\boldsymbol{\alpha}]} p(\mathbf{x}^i|[\boldsymbol{\alpha}])\nu_{[\boldsymbol{\alpha}]}.$$

Then the maximum likelihood estimate $\hat{\boldsymbol{\nu}}$ of $\boldsymbol{\nu}$ is consistent as $N \to \infty$.

Consistent estimation of the $\nu_{[\boldsymbol{\alpha}]}$ is an important result, justifying the results of the following sections. To emphasize the differences in parameter space and procedure, work based on equivalence classes $[\boldsymbol{\alpha}]$ rather than profiles $\boldsymbol{\alpha}$ will from hereon be referred to under the name of the Non-Identifiability ADjusted DINA (NIAD-DINA) model.

6 Marginal Identifiability

We now wish to extend the concept of identifiability to individual attributes. This is motivated by the fact that, though the presence of multiple profiles in the same equivalence class signals non-identifiability, some individual attributes may still be identifiable within the class. To illustrate, consider the the Q-matrix Q_3 from (3) and one of its equivalence classes, $[000] = \{000, 010, 001\}$. If a profile $\boldsymbol{\alpha} \in [000]$, then its first component $\alpha_1 = 0$, but the values of α_2 and α_3 are uncertain. Thus, posterior weight $p([000]|x)$ on this class counts as positive evidence that $\alpha_1 = 0$, but does not help in deciding α_2 or α_3. This observation motivates the following definition:

Definition 5. *An attribute is* marginally identifiable *within an equivalence class when either all members of that class possess that attribute or none of them do.*

Define the marginal identifiability indicator $\delta_{[\alpha],k}$ as follows:

$$\delta_{[\alpha],k} = \prod_{\alpha' \in [\alpha]} \alpha'_k + \prod_{\alpha' \in [\alpha]} (1 - \alpha'_k). \tag{5}$$

Then, $\delta_{[\alpha],k} = 1$ when Attribute k is marginally identifiable within equivalence class $[\alpha]$. Posterior weight on a class $[\alpha]$ only provides information about the k-th attribute when $\delta_{[\alpha],k} = 1$; otherwise, there is no information beyond the prior.

6.1 Exam Quality: the Marginal Identifiability Rate

Since non-identifiability is frequently unavoidable with Q-matrix based CDMs, it is important to measure its extent. For a more nuanced view, this is done on a marginal, basis.

Given the proportion ν_α of each attribute profile α, the proportion of the population for which the k-th attribute is marginally identifiable can be quantified by ζ_k, as follows:

$$\zeta_k = \sum_{\{\alpha : \delta_{[\alpha],k}=1\}} \nu_\alpha. \tag{6}$$

Let ζ be the vector of all ζ_k. Then ζ is the proportion of the population for which each attribute is marginally identifiable, i.e., the marginal identifiability rate.

Oftentimes ν_α, and thus ζ, is unknown. Under the conditions of Theorem 1, ζ can be consistently estimated by its maximum likelihood estimator $\hat{\zeta}$.

Proposition 4. *Suppose an assessment follows the DINA model, with known Q-matrix Q and item parameters $\mathbf{s}$ and $\mathbf{g}$. Let $\hat{\nu}_{[\alpha]}$ be the MLE estimate of $\nu_{[\alpha]}$. Then*

$$\hat{\zeta}_k = \sum_{\{[\alpha]:\delta_{[\alpha],k}=1\}} \hat{\nu}_{[\alpha]}, \ \ k = 1, \ldots, K. \tag{7}$$

is consistent as $N \to \infty$.

The consistency of $\hat{\zeta}$ is a direct consequence of the consistency of $\hat{\nu}_{[\alpha]}$ in Theorem 1. We thus obtain a very reasonable measure of exam quality, in terms of the proportion of the population for which each attribute is marginally identifiable.

7 Classification

Non-identifiability has potentially serious effects on respondent classification. Classification is generally conducted based on the posterior distribution $p(\alpha|\mathbf{x}) \propto p(\mathbf{x}|\alpha)p(\alpha)$. Recall that profiles in the same equivalence class have the same likelihood. Thus, the posterior will simply be a reflection of the prior, without

any added information from the data (DeCarlo, 2011). In fact, within an equivalence class the posteriors are proportional to the priors. Given a prior $p(\boldsymbol{\alpha})$, for any $\boldsymbol{\alpha}^1, \boldsymbol{\alpha}^2 \in [\boldsymbol{\alpha}]$,

$$\frac{p(\boldsymbol{\alpha}^1|\mathbf{x})}{p(\boldsymbol{\alpha}^2|\mathbf{x})} = \frac{p(\mathbf{x}|\boldsymbol{\alpha}^1|)p(\boldsymbol{\alpha}^1)}{p(\mathbf{x}|\boldsymbol{\alpha}^2)p(\boldsymbol{\alpha}^2)} = \frac{p(\mathbf{x}|[\boldsymbol{\alpha}])p(\boldsymbol{\alpha}^1)}{p(\mathbf{x}|[\boldsymbol{\alpha}])p(\boldsymbol{\alpha}^2)} = \frac{p(\boldsymbol{\alpha}^1)}{p(\boldsymbol{\alpha}^2)}.$$

Posteriors are often calculated by maximizing the marginal maximum likelihood $L(\boldsymbol{\nu}, \mathbf{s}, \mathbf{g})$ via the E-M algorithm (Haertel, 1989; de la Torre, 2009; Rupp et al., 2010). Then, since all vectors $\boldsymbol{\nu}$ with identical weights on each class $\nu_{[\boldsymbol{\alpha}]}$ have identical likelihoods, any $\boldsymbol{\nu}$ achieving the maximizing $\nu_{[\boldsymbol{\alpha}]}$ may result. The values chosen are determined by the starting values, which have little validity for classification.

When the posterior is sensitive to the prior, it is important to work with $p([\boldsymbol{\alpha}])$, which can be estimated consistently, rather than $p(\boldsymbol{\alpha})$. Thus classification here will be conducted based on $p([\boldsymbol{\alpha}]|\mathbf{x} \propto p(\mathbf{x}|[\boldsymbol{\alpha}])p([\boldsymbol{\alpha}])$ instead of the usual posterior. This calculation does not require a separate fitting of the model, since

$$p([\boldsymbol{\alpha}]|\mathbf{x}) = \frac{p(\mathbf{x}|[\boldsymbol{\alpha}])\nu_{[\boldsymbol{\alpha}]}}{p(\mathbf{x})} = \frac{p(\mathbf{x}|[\boldsymbol{\alpha}])\sum_{\boldsymbol{\alpha}'\in[\boldsymbol{\alpha}]}\nu_{\boldsymbol{\alpha}'}}{p(\mathbf{x})} = \sum_{\boldsymbol{\alpha}'\in[\boldsymbol{\alpha}]} p(\boldsymbol{\alpha}'|\mathbf{x}) \qquad (8)$$

From this posterior, we then define

$$p_k^{\min}(\mathbf{x}) = \sum_{[\boldsymbol{\alpha}]:\alpha_k=1, d_{[\boldsymbol{\alpha}],k}=1} p([\boldsymbol{\alpha}]|\mathbf{x}), \qquad (9)$$

$$p_k^{\max}(\mathbf{x}) = p_k^{\min}(\mathbf{x}) + \sum_{[\boldsymbol{\alpha}]:d_{[\boldsymbol{\alpha}],k}=0} p([\boldsymbol{\alpha}]|\mathbf{x}), \qquad (10)$$

where $\delta_{[\boldsymbol{\alpha}],k}$ is the marginal identifiability indicator defined in (5). Classification follows from the fact that, depending upon the specific hyperprior on $\boldsymbol{\nu}$ or starting point of the E-M algorithm, the DINA model may produce marginal posterior probabilities of mastery $p(\alpha_k = 1|\mathbf{x})$ anywhere in the range $[p_k^{\min}(\mathbf{x}), p_k^{\max}(\mathbf{x})]$. Thus, it is only appropriate to conclude that $\alpha_k = 1$ when $p_k^{\min}(\mathbf{x})$ is high, or that $\alpha_k = 0$ when $p_k^{\max}(\mathbf{x})$ is low. A natural cutoff for both is 0.5, but it may be adjusted as necessary. This classification method, from hereon referred to as the NIAD-DINA classification algorithm, accounts for both uncertainty in the prior and uncertainty caused by slipping and guessing. It is summarized in Table 3.

Table 3: NIAD-DINA classification algorithm

Step	Procedure	q.v.
Input:	Q-matrix $Q = (q_{jk})_{J \times K}$, data $X = (x_{ik})_{N \times J}$.	
(1)	Fit the model to produce $p(\boldsymbol{\alpha}|\mathbf{x})$.	
(2)	Partition the attribute profile space.	Table 1
(3)	Calculate the marginal identifiability vector $\delta_{[\boldsymbol{\alpha}]}$.	(5)
(4)	Sum posteriors $p(\boldsymbol{\alpha}|x)$ for $p([\boldsymbol{\alpha}]|\mathbf{x})$.	(8)
(5)	Calculate $p_k^{\min}(\mathbf{x})$ and $p_k^{\max}(\mathbf{x})$ for every $k, \mathbf{x}$.	(9), (10)
(6)	Classify:	
	$\quad$ If $p_k^{\min} > 0.5$, then $\hat{\alpha}_k = 1$.	
	$\quad$ If $p_k^{\max} < 0.5$, then $\hat{\alpha}_k = 0$.	
	$\quad$ Else, $\hat{\alpha}_k = *$ (unclassified).	
Output:	Classifications $\hat{\alpha}_k^i \in \{0, 1, *\}$ for all i, k.	

8   Extensions

8.1   Variants of the DINA

The methodology of partitioning in Section 4 applies to any model where the presence of a difference in the ideal response pattern fully determines the presence of a difference in the likelihood function. Included among these models are all variants of the DINA model listed in Section 2.2. Since these variants are, in essence, a restriction on the space of ν, the consistency result for $\nu_{[\boldsymbol{\alpha}]}$ in Section 5, along with all the following results, holds when the model is in fact correct. However, if the true $\nu_{[\boldsymbol{\alpha}]}$ do not fall under the set of values consistent with the restriction on the parameter space, then even estimates of $\nu_{[\boldsymbol{\alpha}]}$ will no longer be consistent. Thus, large differences in the $\hat{\nu}_{[\boldsymbol{\alpha}]}$ calculated under restricted models from those calculated under the NIAD-DINA model are symptomatic of model misspecification, and may imply that the DINA variant chosen is overly restrictive on the prior. Goodness-of-fit measures such as the AIC and BIC will reflect lack of fit appropriately if the saturated model has the correct number of parameters $2J + L$, rather than $2J + 2^K - 1$.

8.2   The DINO Model

The DINO model also specifies item and attribute relationships using a Q-matrix, but the ideal responses are calculated as

$$\xi_j(Q, \boldsymbol{\alpha}) = 1 - \prod_{k=1}^{K}(1 - \alpha_k)^{q_{jk}} = \mathbf{1}(\alpha_k = q_{jk} = 1 \text{ for some } k).$$

12

As in the DINA model, the response probabilities are functions of item parameters $s_j = P(X_j = 0 | \xi_j = 1)$ and $g_j = P(X_j = 1 | \xi_j = 0)$. Under the DINA model, ideal responses are *correct* when the respondent possesses all required attributes; under the DINO model, ideal responses are *incorrect* when the respondent *does not* possess all required attributes. Thus, for responses X following the DINO model, the reverse responses $1 - X$ follow the DINA model (with a reversed interpretation of the attribute profile vectors). This dualism implies that all the results of this paper apply to the DINO model.

9 Results

9.1 Simulation Results

We first demonstrate the procedures on simulated data. Responses are generated for $N = 5000$ resondents taking an assessment with $J = 6$ items measuring $K = 3$ distinct attributes. The respondents' mastery or nonmastery of the measured attributes is randomly generated according to the probability $p_{sim}(\boldsymbol{\alpha})$ of each profile $\boldsymbol{\alpha} \in \{0, 1\}^3$, as listed in Table 4. The responses themselves follow

Table 4: Population proportions of each attribute profile

	$\boldsymbol{\alpha}$							
	000	001	010	011	100	101	110	111
p_{sim}	0.27	0.00	0.01	0.04	0.10	0.16	0.20	0.21

low the DINA model according to the Q-matrix Q_{sim} with slipping s_{sim} and guessing g_{sim} as shown in Table 5.

Table 5: Q-matrix, slipping, and guessing for simulated data.

Item (j)	Attribute vector $(\mathbf{q}^j)$	Slipping (s_j)	Guessing (g_j)
1.	100	0.14	0.10
2.	110	0.12	0.15
3.	011	0.18	0.18
4.	100	0.17	0.18
5.	110	0.08	0.06
6.	011	0.05	0.06

The Q-matrix Q_{sim} is incomplete, and the resulting instability in the posterior becomes clear once the data is fitted multiple times. As an example, the posterior probabilities of each attribute profile given the zero response vector $\mathbf{0} = (0, 0, \ldots, 0)$ are summarized in Table 6. Here, the DINA, HO-DINA, and RHO-DINA are overparameterized and produce a wide range of results for Profiles [000], [001], and [010]. The slight variability in the ind-DINA estimates is a numerical artifact. While the ind-DINA does not suffer from nonidentifiability,

Table 6: Posterior probabilities given zero correct responses, $p(\boldsymbol{\alpha}|\mathbf{x} = \mathbf{0})$

| | $\boldsymbol{\alpha}$ | | | | | | | |
	000	001	010	011	100	101	110	111
truth								
	0.91	0.02	0.05	0.00	0.01	0.01	0.00	0.00
minimums								
DINA	0.01	0.02	0.03	0.00	0.00	0.00	0.00	0.00
HO-DINA	0.13	0.09	0.03	0.00	0.01	0.01	0.00	0.00
RHO-DINA	0.55	0.11	0.29	0.00	0.02	0.01	0.00	0.00
ind-DINA	0.29	0.24	0.37	0.00	0.02	0.02	0.00	0.00
maximums								
DINA	0.71	0.86	0.56	0.00	0.02	0.03	0.00	0.00
HO-DINA	0.62	0.81	0.71	0.00	0.02	0.02	0.00	0.00
RHO-DINA	0.58	0.13	0.30	0.00	0.02	0.01	0.00	0.00
ind-DINA	0.31	0.28	0.40	0.01	0.03	0.02	0.00	0.00

Note: Minimum and maximum values of the posterior $p(\boldsymbol{\alpha}|\mathbf{x} = 0)$, as generated over ten runs of the (random start) E-M algorithm.

it still does not give accurate estimates in this case since the model assumptions are incorrect.

Partitioning the attribute profile space as directed by Table 1 produces the five equivalence classes listed in Table 7, two of which have multiple members. The table also reports the marginal identifiability vector $\delta_{[\boldsymbol{\alpha}]}$ for each class. Note that since Items 1 and 4 are devoted to Attribute 1, it is always marginally identifiable and $\delta_{[\boldsymbol{\alpha}],1} \equiv 1$. It is also clear that nonidentifiability most seriously affects Attribute 3, which is marginally non-identifiable for members of both [000] and [100]. Finally, Table 7 also reports E-M estimates of the proportion of respondents in each class under the DINA and several variants, along with the true proportion. Note the accuracy of the DINA estimates, which are consistent, and the inaccuracy of the ind-DINA estimates due to model misfit.

Table 7: Equivalence classes, along with their class sizes, true and maximum likelihood probabilities, and marginal identifiability vectors.

| | | | | $\nu_{[\boldsymbol{\alpha}]}$ | | | |
$[\boldsymbol{\alpha}]$	Size	True	DINA	HO-DINA	RHO-DINA	ind-DINA	$\delta_{[\boldsymbol{\alpha}]}$
[000]	3	0.29	0.30	0.29	0.29	0.22	100
[100]	2	0.26	0.26	0.27	0.26	0.31	110
[011]	1	0.04	0.04	0.04	0.04	0.08	111
[110]	1	0.20	0.20	0.19	0.20	0.20	111
[111]	1	0.21	0.21	0.21	0.21	0.18	111

We now consider variability in the marginal posterior probabilities $p(\alpha_k = 1|\mathbf{x})$. Table 8 gives the sample range of $p(\alpha_k = 1|\mathbf{x})$ after ten runs of the E-M algorithm, in addition to the theoretical range. Note the large theoretical ranges for $p(\alpha_k = 1|\mathbf{x} = \mathbf{0}), k = 2, 3$.

Table 8: Variability in $p(\alpha_k = 1|\mathbf{x} = \mathbf{0})$, the marginal posterior given the zero response vector.

	k		
	1	2	3
sample min	0.03	0.03	0.04
$p_k^{min}(\mathbf{0})$	0.03	0.00	0.00
sample max	0.03	0.56	0.89
$p_k^{max}(\mathbf{0})$	0.03	0.97	1.00

Note: Probabilities calculated by fitting the DINA model over ten runs of E-M algorithm with random starts.

Classification was conducted on a marginal basis, based on $p(\alpha_k = 1|\mathbf{x})$, under each of the models. In addition, NIAD-DINA classification was performed (see Table 3). Marginal misclassification rates $p(\hat{\alpha}_k \neq \alpha_k)$ are compared in Table 9. Note that NIAD-DINA classification results in unclassified individuals; for example, $\hat{\boldsymbol{\alpha}} = (0 * *)$ for those with the zero response vector. This rate is also listed in Table 9. The DINA and HO-DINA are overparameterized and the misclassification rate for Attribute 3 may reach over 40% in both models. The ind-DINA also performs poorly, but due to an overly restricted parameter space rather than nonidentifiability. Adjusting classification under the DINA to account for nonidentifiability according to the method described in Section 7 solves both these issues. It may leave a large proportion of individuals unclassified, but this is a necessary consequence of the assessment design. Classifying these individuals would require further assumptions beyond the model.

Table 9: Marginal misclassification rates under a variety of models.

	Model				
k	DINA	HO-DINA	RHO-DINA	ind-DINA	NIAD-DINA
1	0.07	0.07	0.07	0.09	0.07 (0.00)
2	0.07-0.32	0.07-0.32	0.07	0.26	0.04 (0.32)
3	0.19-0.44	0.19-0.43	0.20	0.21	0.04 (0.56)

Note: Range over 10 runs reported for overparameterized models. All cut-offs equal to 0.5. The proportion of respondents left unclassified under the NIAD-DINA is displayed within parentheses.

In addition to controlling misclassification errors, we may also evaluate the

quality of the assessment by measuring the marginal identifiability rate ζ defined in (6). Table 10 shows both true and estimated values for ζ. Note once again that nonidentifiability affects Attribute 3 more severely than it does Attribute 2. In addition, estimates are generally accurate, except in the case of the ind-DINA, which suffers from lack of fit.

Table 10: True and estimated values for ζ, marginal identifiability rate.

		Model			
k	true	DINA	HO-DINA	RHO-DINA	ind-DINA
1	1.00	1.00	1.00	1.00	1.00
2	0.71	0.70	0.71	0.71	0.78
3	0.44	0.45	0.45	0.45	0.47

In terms of model selection, reducing the number of parameters for the DINA model to $2M + L$ from the original $2M + 2^K$ reduces the comparative advantage of the restricted models. In Table 11, the AIC value of the RHO-DINA barely edges out that of the DINA with identifiability adjustment.

Table 11: AIC and BIC for the DINA, RHO-DINA, and ind-DINA.

	parameters	AIC	BIC
NIAD-DINA	17	32862.8	32973.6
RHO-DINA	16	32861.2	32965.5
ind-DINA	15	32995.9	33093.6

9.2 A Fractions Data Example

We now turn to the widely analyzed fraction subtraction data set of K. K. Tatsuoka (1990). It is composed of the twenty items listed in Table 12. The

Table 12: Items from the fraction subtraction data set (K. K. Tatsuoka, 1990).

No.	Item	No.	Item	No.	Item
1.	$5/3 - 3/4$	8.	$2/3 - 2/3$	15.	$2 - 1/3$
2.	$3/4 - 3/8$	9.	$3\,7/8 - 2$	16.	$4\,5/7 - 1\,4/7$
3.	$5/6 - 1/9$	10.	$4\,4/12 - 2\,7/12$	17.	$7\,3/5 - 4/5$
4.	$3\,1/2 - 2\,3/2$	11.	$4\,1/3 - 2\,4/3$	18.	$4\,1/10 - 2\,8/10$
5.	$4\,3/5 - 3\,4/10$	12.	$11/8 - 1/8$	19.	$4 - 1\,4/3$
6.	$6/7 - 4/7$	13.	$3\,3/8 - 2\,5/6$	20.	$4\,1/3 - 1\,5/3$
7.	$3 - 2\,1/5$	14.	$3\,4/5 - 3\,2/5$		

Q-matrix in Table 13 comes from de la Torre and Douglas (2004), and specifies

the following eight attributes: α_1 = convert a whole number to a fraction; α_2 = separate a whole number from a fraction; α_3 = simplify before subtracting; α_4 = find a common denominator; α_5 = borrow from whole number part; α_6 = column borrow to subtract the second numerator from the first; α_7 = subtract numerators; and α_8 = reduce answers to simplest form.

Table 13: Q-matrix from de la Torre and Douglas (2004).

Item	Attributes ($\mathbf{q}^j$)	Item	Attributes ($\mathbf{q}^j$)	Item	Attributes ($\mathbf{q}^j$)
1.	00010110	8.	00000010	15.	10000010
2.	00010010	9.	01000000	16.	01000010
3.	00010010	10.	01001011	17.	01001010
4.	01101010	11.	01001010	18.	01001110
5.	01010011	12.	00000011	19.	11101010
6.	00000010	13.	01011010	20.	01101010
7.	11000010	14.	01000010		

As pointed out by DeCarlo (2011), this assessment exemplifies the identifiability issues of the DINA model. While Attributes 2 and 7 have items dedicated solely to them, all other attributes appear only in combination. In fact, Attribute 3 only appears in Item 4, in conjunction with Attributes 2, 5, and 7. Attribute 7 is required for all items except one, making it difficult to draw conclusions about other attributes when it has not been mastered. Table 14 displays the marginal posterior probabilities of mastery for each attribute, given the zero response vector. The posterior displayed for the DINA is just one possible output of the E-M algorithm for this data; meanwhile, note the high probabilities of mastery under the ind-DINA model. Common sense dictates that something is out of place when the analysis states that students with a score of zero cannot subtract numerators, but can do everything else, from finding a common denominator to borrowing to reducing to simplest form.

Table 14: Marginal posterior probabilities of mastery given the zero response vector, $p(\alpha_k = 1 | \mathbf{x} = \mathbf{0})$

	k							
	1	2	3	4	5	6	7	8
DINA	0.50	0.08	0.50	0.52	0.53	0.41	0.00	0.59
HO-DINA	0.00	0.13	0.31	0.05	0.02	0.30	0.00	0.25
RHO-DINA	0.02	0.13	0.12	0.05	0.02	0.25	0.00	0.18
ind-DINA	0.74	0.86	0.96	0.86	0.75	0.98	0.00	0.94

With eight attributes in the Q-matrix, there are a total of 256 possible attribute profiles. They can be divided into just 58 different equivalence classes by the partitioning algorithm, 32 of them containing a single identifiable element. The 26 multiple-profile equivalence classes are listed in Table 15, which also

Table 15: Multiple-member equivalence classes, along with their class sizes, maximum likelihood probabilities, and marginal identifiability vectors.

| $[\boldsymbol{\alpha}]$ | Size | $\nu_{[\boldsymbol{\alpha}]}$ | | | | $\delta_{[\boldsymbol{\alpha}]}$ |
		DINA	HO-DINA	RHO-DINA	ind-DINA	
[00000000]	64	0.15	0.12	0.12	0.02	01000010
[01000000]	64	0.04	0.06	0.06	0.31	01000010
[00000010]	8	0.01	0.02	0.03	0.00	11010011
[10000010]	8	0.00	0.00	0.00	0.00	11010011
[00000011]	8	0.02	0.03	0.02	0.00	11010011
[10000011]	8	0.00	0.00	0.00	0.00	11010011
[01000010]	4	0.03	0.04	0.04	0.00	11011011
[01000011]	4	0.11	0.09	0.08	0.00	11011011
[11000010]	4	0.00	0.00	0.00	0.00	11011011
[11000011]	4	0.01	0.01	0.02	0.01	11011011
[00010010]	4	0.00	0.00	0.00	0.00	11010111
[10010010]	4	0.00	0.00	0.00	0.00	11010111
[00010011]	4	0.00	0.00	0.00	0.00	11010111
[10010011]	4	0.00	0.00	0.00	0.00	11010111
[00010110]	4	0.02	0.00	0.00	0.00	11010111
[10010110]	4	0.00	0.00	0.00	0.00	11010111
[00010111]	4	0.00	0.01	0.01	0.01	11010111
[10010111]	4	0.00	0.00	0.00	0.03	11010111
[01010010]	2	0.00	0.00	0.00	0.00	11011111
[11010010]	2	0.00	0.00	0.00	0.00	11011111
[01010011]	2	0.01	0.01	0.00	0.00	11011111
[11010011]	2	0.00	0.00	0.00	0.00	11011111
[01010110]	2	0.00	0.01	0.01	0.00	11011111
[11010110]	2	0.00	0.00	0.00	0.01	11011111
[01010111]	2	0.05	0.06	0.06	0.03	11011111
[11010111]	2	0.06	0.06	0.06	0.09	11011111

displays their class sizes, maximum likelihood probabilities, and marginal identifiability vectors. Within these multiple-profile equivalence classes, Attributes 2 and 7 are always marginally identifiable, while Attribute 3 is never so; this is natural considering our previous observations about the Q-matrix. Profiles within the largest classes contain many zeroes, since under the DINA model non-identifiability affects a particular attribute only for respondents who do not possess other attributes used in combination with that attribute. Also note that the ind-DINA shows signs of model misspecification, since its estimates $\hat{\nu}_{[\boldsymbol{\alpha}]}$ deviate strongly from the estimates derived from the other models.

Table 16 shows the estimated marginal identifiability rates, $\hat{\boldsymbol{\zeta}}$. At the low end, $\hat{\zeta}_3 = 0.48$, bringing into question the ability of this assessment to measure mastery of Attribute 3. Attribute 6 does only slightly better, with $\hat{\zeta}_6 = 0.64$.

Note that Attribute 6 is only utilized in Items 1 and 18; in both cases it appears in conjunction with at least two other attributes.

Table 16: Estimated marginal identifiability rates ζ_k.

				k				
	1	2	3	4	5	6	7	8
DINA	0.81	1.00	0.48	0.81	0.75	0.64	1.00	0.81
HO-DINA	0.82	1.00	0.47	0.82	0.75	0.64	1.00	0.82
RHO-DINA	0.82	1.00	0.48	0.82	0.75	0.63	1.00	0.82
ind-DINA	0.66	1.00	0.47	0.66	0.62	0.64	1.00	0.66

Finally, it is useful to consider how the reduction of the parameter space for the DINA model, based on identifiability, affects model selection by AIC and BIC. In particular, the AIC no longer prefers the ind-DINA to the full DINA model once the reduced parameter space has been applied. The BIC, which will generally choose sparser models than the AIC, still reports lower values for the ind-DINA, but the comparison is much tighter.

Table 17: AIC and BIC for the DINA, RHO-DINA, and ind-DINA.

	parameters	AIC	BIC
DINA	296	9397.0	10665.2
NIAD-DINA	98	9001.0	9420.9
HO-DINA	56	8959.7	9199.6
RHO-DINA	49	8961.9	9171.9
ind-DINA	48	9208.3	9413.9

10   Discussion

In general, it is difficult to obtain a complete Q-matrix. Oftentimes, due to the demands of practicality, assessments must involve items that require a combination of skills. Using the tools discussed in this paper, it is possible to determine the extent to which nonidentifiability affects classification and estimation under the DINA model. Marginal identifiability rates ζ, which can be estimated consistently, provide an overall measure of the extent of non-identifiability; meanwhile, NIAD-DINA classification takes marginal identifiability into consideration in order to control classification errors that are otherwise quite sensitive to the prior information. The results here suggest that when designing items to test a particular attribute, if using a combination of skills is unavoidable, it is best to combine that attribute with basic attributes mastered by a large proportion of the population. After all, it is only impossible to recover information about a

particular attribute when the respondent does not posess one or more of the other attributes tested by the same item.

Many of the methods currently in use may resolve idenifiability issues by enforcing restrictions on the attribute profile space. Variants of the DINA such as the ind-DINA, HO-DINA, and RHO-DINA accomplish this by specifying a structure and a prior on the probabilities $p(\boldsymbol{\alpha}) = \nu_{\boldsymbol{\alpha}}$. Although this may eliminate non-identifiability and create a unique global maximum for the likelihood, model misspecification becomes a risk. Thus, careful comparison of these variants to the NIAD-DINA becomes important.

A Proofs

Proof of Proposition 1. Suppose $\xi_j^1 = \xi_j^2$ for all j such that $1 - s_j \neq g_j$. If $1 - s_j = g_j$, then the response distribution for item j does not depend on $\boldsymbol{\xi}$:

$$P(x_j|\boldsymbol{\xi}, s, g) = (1 - s_j)^{\xi_j x_j} g_j^{(1-\xi_j)x_j} s_j^{\xi_j(1-x_j)} (1 - g_j)^{(1-\xi_j)(1-x_j)}$$
$$= (1 - s_j)^{x_j} s_j^{1-x_j} = g_j^{x_j} (1 - g_j)^{1-x_j}.$$

Thus, for every $\mathbf{x} \in \{0, 1\}^J$,

$$P(\mathbf{x}|\boldsymbol{\xi}^1, s, g) = \prod_{j=1}^m P(x_j|\xi_j^1, s_j, g_j)$$

$$= \prod_{\{i:1-s_j=g_j\}} P(x_j|\xi_j^1, s_j, g_j) \prod_{\{i:1-s_j\neq g_j\}} P(x_j|\xi_j^1, s_j, g_j)$$

$$= \prod_{\{i:1-s_j=g_j\}} P(x_j|\xi_j^2, s_j, g_j) \prod_{\{i:1-s_j\neq g_j\}} P(x_j|\xi_j^2, s_j, g_j) = P(\mathbf{x}|\boldsymbol{\xi}^2, s, g)$$

and $\boldsymbol{\alpha}^1$ cannot be separated from $\boldsymbol{\alpha}^2$.

Now suppose that $\xi_j^1 \neq \xi_j^2$ for some j such that $1 - s_j \neq g_j$. Then

$$P(x_j = 1|\boldsymbol{\xi}^1, s, g) = (1 - s_j)^{\xi_j^1} g_j^{1-\xi_j^1} \neq g_j^{\xi_j^1} (1 - s_j)^{1-\xi_j^1}$$
$$= g_j^{1-\boldsymbol{\xi}^2} (1 - s_j)^{\xi_j^2} = P(x_j = 1|\boldsymbol{\xi}^2, s, g),$$

so the response distributions differ. $\qquad\square$

Proof of Proposition 2. Suppose that Q is complete. WLOG, for $j = 1, \ldots, K$ let $\mathbf{q}^j = e_j$. Then, for $j = 1, \ldots, K$, $\xi_j(Q, \boldsymbol{\alpha}) = \alpha_j$ and given any two attribute profiles $\boldsymbol{\alpha}^1 \neq \boldsymbol{\alpha}^2$,

$$\xi_{1:K}(Q, \boldsymbol{\alpha}^1) = \boldsymbol{\alpha}^1 \neq \boldsymbol{\alpha}^2 = \xi_{1:K}(Q, \boldsymbol{\alpha}^2)$$

By Proposition 1, Q separates any $\boldsymbol{\alpha}^1 \neq \boldsymbol{\alpha}^2$.

Now suppose that $\exists k_* \in \{1, \ldots, K\}$ such that $e_{k_*} \notin R_Q$. WLOG, suppose $k_* = 1$. Consider profiles $\boldsymbol{\alpha} = e_1$ and $\boldsymbol{\alpha}' = \mathbf{0}$, the zero column-vector. For each item $j = 1, \ldots, J$, if $q_{j1} = 0$ then

$$\xi_j(Q, e_1) = \left(1^0\right) \prod_{k\neq 1} 0^{q_{jk}} = \left(0^0\right) \prod_{k\neq 1} 0^{q_{jk}} = \xi_j(Q, \mathbf{0}).$$

Else, $q_{j1} = 1$ and there exists some $k_{**} \neq 1$ such that $q_{jk_{**}} = 1$ and

$$\xi_j(Q, e_1) = \left(0^1\right) \prod_{k\neq k_{**}} [\mathbf{1}(k = 1)]^{q_{jk}} = 0 = \left(0^1\right) \prod_{k\neq k_{**}} 0^{q_{jk}} = \xi_j(Q, \mathbf{0}).$$

Thus, $\boldsymbol{\xi}(Q, e_1) = \boldsymbol{\xi}(Q, \mathbf{0})$ and by Proposition 1, attribute profiles e_1 and $\mathbf{0}$ cannot be separated. $\qquad\square$

Proof of Proposition 3. The relation '$\sim$' is (i) reflexive: $\boldsymbol{\xi}(Q, \boldsymbol{\alpha}) = \boldsymbol{\xi}(Q, \boldsymbol{\alpha}) \Rightarrow \boldsymbol{\alpha} \sim \boldsymbol{\alpha}$ for all profiles $\boldsymbol{\alpha}$, (ii) symmetric: if $\boldsymbol{\alpha}^1 \sim \boldsymbol{\alpha}^2$, then $\boldsymbol{\xi}(Q, \boldsymbol{\alpha}^1) = \boldsymbol{\xi}(Q, \boldsymbol{\alpha}^2)$ and $\boldsymbol{\alpha}^2 \sim \boldsymbol{\alpha}^1$ for any profiles $\boldsymbol{\alpha}^1, \boldsymbol{\alpha}^2$; (iii) transitive: if $\boldsymbol{\alpha}^1 \sim \boldsymbol{\alpha}^2$ and $\boldsymbol{\alpha}^2 \sim \boldsymbol{\alpha}^3$, then $\boldsymbol{\xi}(Q, \boldsymbol{\alpha}^1) = \boldsymbol{\xi}(Q, \boldsymbol{\alpha}^2) = \boldsymbol{\xi}(Q, \boldsymbol{\alpha}^3)$ and $\boldsymbol{\alpha}^1 \sim \boldsymbol{\alpha}^3$ for any profiles $\boldsymbol{\alpha}^1, \boldsymbol{\alpha}^2, \boldsymbol{\alpha}^3$. $\qquad\square$

Proof of Theorem 1. We can write the likelihood

$$L(\boldsymbol{\nu}) = p(X|\boldsymbol{\nu}) = \prod_{i=1}^{N} p(\mathbf{x}^i|\boldsymbol{\nu})$$

and the log-likelihood as

$$\ell(\boldsymbol{\nu}) = \sum_{i=1}^{N} \log p(\mathbf{x}^i|\boldsymbol{\nu}) = \sum_{\mathbf{X} \in \{0,1\}^J} N_{\mathbf{X}} \log \left(\sum_{[\boldsymbol{\alpha}]} p\left(\mathbf{x}\,|[\boldsymbol{\alpha}]\right) \nu_{[\boldsymbol{\alpha}]} \right),$$

where $N_{\mathbf{X}} = \#\{i : \mathbf{x}^i = \mathbf{x}\}$.

We first check for identifiability. Suppose there are L distinct equivalence classes partitioning the attribute profile space. The probability of each vector $\mathbf{x}$ given each class $[\boldsymbol{\alpha}]$ can be written as a $2^J \times L$ matrix $P = (p_{\mathbf{X},[\boldsymbol{\alpha}]})$, where $p_{\mathbf{X},[\boldsymbol{\alpha}]} = p(\mathbf{x}|[\boldsymbol{\alpha}])$. The vector of total probabilities for each response vector $\mathbf{x}$, for any $\boldsymbol{\nu}$, can be written as the matrix product $P\boldsymbol{\nu}$. Thus, we need to show that there is no $\boldsymbol{\nu}^1, \boldsymbol{\nu}^2$ s.t. $P\boldsymbol{\nu}^1 = P\boldsymbol{\nu}^2$. Define the vector inequality operation "$\geq$" so that $\mathbf{x} \geq \mathbf{y}$ iff $x_j \geq y_j$ for all j. Let the T-matrix be the $2^J \times L$ matrix $T = (t_{\mathbf{X},[\boldsymbol{\alpha}]})$, indexed over all response vectors $\mathbf{x}$ and equivalence classes $[\boldsymbol{\alpha}]$, such that $t_{\mathbf{X},[\boldsymbol{\alpha}]} = p(\mathbf{X} \geq \mathbf{x}|[\boldsymbol{\alpha}])$. This is a variant of the T-matrix used to examine Q-matrix identifiability in Liu, Xu, and Ying (2012). Then $P\boldsymbol{\nu}^1 = P\boldsymbol{\nu}^2$ iff $T\boldsymbol{\nu}^1 = T\boldsymbol{\nu}^2$, and the identifiability condition is equivalent to T being a rank L matrix.

First, suppose that $g \equiv 0$. WLOG, assume that the L equivalence classes $[\boldsymbol{\alpha}^1], \ldots, [\boldsymbol{\alpha}^L]$ are ordered lexicographically by their minimal representatives, $\boldsymbol{\alpha}^{1*}, \ldots, \boldsymbol{\alpha}^{L*}$. Thus, if $\boldsymbol{\alpha}^{k*} \geq \boldsymbol{\alpha}^{\ell*}$, then $k \geq \ell$. Also, let $\mathbf{x}^\ell = \boldsymbol{\xi}(Q, [\boldsymbol{\alpha}^\ell])$ for $\ell = 1, \ldots, L$. Define $T^* = (t_{k\ell}^*)$, where $t_{k\ell}^* = t_{\mathbf{X}^k,[\boldsymbol{\alpha}^\ell]}$. Then T^* is an $L \times L$ submatrix of T, containing the specified rows $\mathbf{x}^1, \ldots, \mathbf{x}^L$. Moreover, T^* is an upper triangular matrix. This is a consequence of the fact that for any $k > \ell$, $\boldsymbol{\alpha}^{\ell*} \not\geq \boldsymbol{\alpha}^{k*}$. Thus, there must be some item $j \in \{1, \ldots, J\}$ for which individuals with profiles $\boldsymbol{\alpha} \in [\boldsymbol{\alpha}^k]$ possess the necessary attributes, but individuals with profiles $\boldsymbol{\alpha} \in [\boldsymbol{\alpha}^\ell]$ do not. Then $p(X_j = 1|[\boldsymbol{\alpha}^\ell]) = g_j = 0 \Rightarrow t_{\mathbf{X}^k,[\boldsymbol{\alpha}^\ell]} = t_{k,\ell}^* = 0$. In addition, on the diagonal, $t_{\ell,\ell}^* = \prod_{\{i:x_j^\ell=1\}} (1 - s_j) \neq 0$. Thus, T^* is a rank L matrix, as is T.

Next suppose that $g \not\equiv 0$. Consider the T matrix as a function of $\mathbf{c} = 1 - \mathbf{s}$ and $\mathbf{g}$. Then the T-matrix $T(\mathbf{c}, \mathbf{g})$ can be written as a linear transformation of

another T-matrix $T(\mathbf{c} - \mathbf{g}, \mathbf{0})$. For any subset of the items S and any constants b_j,

$$\prod_{j \in S}(b_j - g_j) = \prod_{j \in S} b_j - \sum_{j \in S} g_j \prod_{k \neq j} b_k + \sum_{k \neq j \in S} g_j g_k \prod_{\ell \neq j,k} b_\ell - \cdots + (-1)^{\#S} \prod_{j \in S} g_j$$

For each entry of the T-matrix, the b_j will correspond to either c_j or g_j, depending on the value of $\xi_j(Q, [\boldsymbol{\alpha}])$. Then, $T(\mathbf{c} - \mathbf{g}, \mathbf{0}) = D(\mathbf{g}) \cdot T(\mathbf{c}, \mathbf{g})$, where the transformation matrix $D(g)$ is a $2^J \times 2^J$ matrix depending solely on $\mathbf{g}$. Since the rows of T are ordered lexicographically by $\mathbf{x}$, $D(g)$ is lower triangular with diagonal $diag(D) \equiv 1$. Thus, D is full-rank and $\text{rank}(T(\mathbf{c}, \mathbf{g})) = \text{rank}(T(\mathbf{c} - \mathbf{g}, \mathbf{0})) = L$. The model is identifiable, and all other conditions for the consistency of the maximum likelihood estimator are clearly evident. $\square$

Proof of Proposition 4. Since

$$\sum_{\{\boldsymbol{\alpha}:\delta_{[\boldsymbol{\alpha}],k}=1\}} \nu_{\boldsymbol{\alpha}} = \sum_{\{[\boldsymbol{\alpha}]:\delta_{[\boldsymbol{\alpha}],k}=1\}} \sum_{\boldsymbol{\alpha}' \in [\boldsymbol{\alpha}]} \nu_{\boldsymbol{\alpha}} = \sum_{\{[\boldsymbol{\alpha}]:\delta_{[\boldsymbol{\alpha}],k}=1\}} \nu_{[\boldsymbol{\alpha}]},$$

ζ_k can be written in terms of $\nu_{[\boldsymbol{\alpha}]}$ as

$$\zeta_k = \sum_{\{[\boldsymbol{\alpha}]:\delta_{[\boldsymbol{\alpha}],k}=1\}} \nu_{[\boldsymbol{\alpha}]}.$$

By Theorem 1, the MLE $\hat{\nu}_{[\boldsymbol{\alpha}]}$ is consistent as $N \to \infty$ under the conditions of the proposition. Thus, $\hat{\zeta}_k$ is consistent as $N \to \infty$. $\square$

References

Chiu, C.-Y., Douglas, J. A., & Li, X. (2009). Cluster analysis for cognitive diagnosis: Theory and applications. *Psychometrika*, *74*, 633-665.

de la Torre, J. (2008). The generalized DINA model. *Paper presented at the annual International Meeting of the Psychometric Society, Durham, NH*.

de la Torre, J. (2009). DINA model and parameter estimation: a didactic. *Journal of Educational and Behavioral Statistics*, *34*, 115-130.

de la Torre, J., & Douglas, J. A. (2004). Higher order latent trait models for cognitive diagnosis. *Psychometrika*, *69*, 333-353.

de la Torre, J., & Douglas, J. A. (2008). Model evaluation and multiple strategies in cognitive diagnosis: An analysis of fraction subtraction data. *Psychometrika*, *73*, 595-624.

DeCarlo, L. T. (2011). On the analysis of fraction subtraction data: the DINA model, classification, class sizes, and the Q-matrix. *Applied Psychological Measurement*, *35*, 8-26.

DiBello, L. V., Stout, W. F., & Roussos, L. A. (1995). Unified cognitive psychometric diagnostic assessment likelihood-based classification techniques. In P. D. Nichols, S. F. Chipman, & R. L. Brennan (Eds.), *Cognitively diagnostic assessment* (p. 361-390). Hillsdale, NJ: Erlbaum Associates.

Haertel, E. H. (1989). Using restricted latent class models to map the skill structure of achievement items. *Journal of Educational Measurement*, *26*, 301-323.

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

Junker, B. W., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement*, *25*, 258-272.

Liu, J., Xu, G., & Ying, Z. (2012). Data-driven learning of Q-matrix. *Applied Psychological Measurement*, *36*, 548-564.

Maris, E. (1999). Estimating multiple classification latent class models. *Psychometrika*, *64*, 187-212.

Rupp, A. A., Templin, J. L., & Henson, R. A. (2010). *Diagnostic measurement: Theory, methods, and applications.* New York: Guilford Press.

Tatsuoka, C. (2002). Data-analytic methods for latent partially ordered classification models. *Applied Statistics (JRSS-C)*, *51*, 337-350.

Tatsuoka, K. K. (1990). Toward an integration of item-response theory and cognitive error diagnosis. In N. Frederiksen, R. Glaser, A. Lesgold, & M. Shafto (Eds.), *Diagnostic monitoring of skill and knowledge acquisition* (p. 453-488). Hillsdale, NJ: Erlbaum.

Tatsuoka, K. K. (1991). *Boolean algebra applied to determination of universal set of knowledge states* (Research report No. RR-91-44-ONR). Princeton, NJ: Educational Testing Service.

Templin, J. L. (2006). CDM: cognitive diagnosis modeling

with Mplus [Computer software manual]. Available from `http://jtemplin.myweb.uga.edu/cdm/cdm.html`

Templin, J. L., & Henson, R. A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological Methods*, *11*, 287-305.

Templin, J. L., Henson, R. A., & Douglas, J. A. (2006, April). *General theory and estimation of cognitive diagnosis models: Using Mplus to derive model estimates.* Paper presented at the 2007 National Council on Measurement in Education training session. Chicago, IL.

Templin, J. L., Henson, R. A., Templin, S. E., & Roussos, L. A. (2008). Robustness of hierarchical modeling of skill association in cognitive diagnosis models. *Applied Psychological Measurement*, *32*, 559-574.