Statistica Sinica Preprint No: SS-2021-0350	
Title	Identifiability of Hierarchical Latent Attribute Models
Manuscript ID	SS-2021-0350
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202021.0350
Complete List of Authors	Yuqi Gu and Gongjun Xu
Corresponding Authors	Gongjun Xu
E-mails	gongjun@umich.edu

Identifiability of Hierarchical Latent Attribute Models

Yuqi Gu[1] and Gongjun Xu[2]

[1]*Columbia University* and [2]*University of Michigan*

Abstract: Hierarchical latent attribute models (HLAMs) are a family of discrete latent variable models that are attracting increasing attention in the educational, psychological, and behavioral sciences. An HLAM includes a binary structural matrix and a directed acyclic graph specifying hierarchical constraints on the configurations of the latent attributes. These components encode practitioners' design information and carry important scientific meaning. However, despite the popularity of HLAMs, the fundamental issue of identifiability remains unaddressed. The existence of the attribute hierarchy graph leads to a degenerate parameter space, and the potentially unknown structural matrix further complicates the identifiability problem. Here, we identify the latent structure and model parameters underlying an HLAM, and develop sufficient and necessary identifiability conditions. These results directly and sharply characterize the effects on identifiability of different attribute types in the graph. The proposed conditions provide insights into diagnostic test designs under the attribute hierarchy, and serve as tools that we can use to assess the validity of an estimated HLAM.

Key words and phrases: Identifiability, Attribute hierarchy graph, **Q**-matrix, Cognitive diagnosis.

1. Introduction

Latent attribute models are a family of discrete latent variable models popular in multiple scientific disciplines, including cognitive diagnosis in educational assessments (Junker and Sijtsma, 2001; von Davier, 2008; Henson et al., 2009; Rupp et al., 2010; de la Torre, 2011; Wang et al., 2018), psychiatric diagnosis of mental disorders (Templin and Henson, 2006;

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

de la Torre et al., 2018), and epidemiological and medical measurement studies (Wu et al., 2017; O'Brien et al., 2019). Based on subjects' responses (often binary) to a set of items, a latent attribute model enables a fine-grained inference on the status of an underlying set of the subjects' latent traits. This further allows us to cluster the population into interpretable subgroups based on the inferred attribute patterns. In a latent attribute model, each attribute is often assumed to be binary and carries a specific scientific meaning. For example, in an educational assessment, the observed responses are students' correct or wrong answers to a set of test items. Here the latent attributes indicate students' binary states of mastery or deficiency of certain skills, as measured by the assessment (Junker and Sijtsma, 2001; von Davier, 2008; Rupp et al., 2010). The dependence between the latent attributes can be further modeled to incorporate practitioners' prior knowledge. A particularly popular and powerful way of modeling attribute dependence in educational and psychological studies is to enforce hard constraints on the hierarchical configurations of the attributes. Specifically, educational experts often postulate that some prerequisite relations exist among the binary skill attributes, such that mastering some skills serves as a prerequisite for mastering others (Leighton et al., 2004). Such a family of *Hierarchical latent attribute models* (HLAMs) are attracting increasing attention in cognitive diagnostic applications; see Leighton et al. (2004), Gierl et al. (2007), Templin and Bradshaw (2014), and Wang and Lu (2020). However, despite their popularity, the fundamental identifiability issue of HLAMs remains unaddressed. This study fills this gap by providing the identifiability theory for HLAMs.

HLAMs have close connections with many other popular statistical and machine learning models. Because each possible configuration of the discrete attributes represents a pattern defining a latent subpopulation, an HLAM can be viewed as a structured mixture model

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

(McLachlan and Peel, 2004) and gives rises to model-based clustering (Fraley and Raftery, 2002) of multivariate categorical data. HLAMs are related to several multivariate discrete latent variable models in the machine learning literature, including latent tree graphical models (Choi et al., 2011), restricted Boltzmann machines (Hinton, 2002), and latent feature models (Ghahramani and Griffiths, 2006), but with two key differences. First, the observed variables are assumed to have a certain structured dependence on the latent attributes. This dependence is summarized by a structural matrix, the so-called $\mathbf{Q}$-matrix (Tatsuoka, 1990), to encode scientific interpretations. The second key feature is that HLAMs incorporate the hierarchical structure among the latent attributes. For instance, in educational cognitive diagnosis, possessing certain skill attributes is often assumed to be prerequisite for possessing some other skills (Leighton et al., 2004; Templin and Bradshaw, 2014).

The real-world applications of HLAMs are challenged by the identifiability of the attribute hierarchy, structural $\mathbf{Q}$-matrix, and other model parameters. First, in many applications, the attribute hierarchy and the structural $\mathbf{Q}$-matrix are specified by domain experts, based on their understanding of the diagnostic tests. Such specifications can be subjective, and may not reflect the underlying truth. Second, the attribute hierarchy and the $\mathbf{Q}$-matrix may even be entirely unknown in an exploratory data analysis, where researchers hope to identify and estimate these quantities directly from the observed data. In both of these situations, a fundamental, yet open question is whether and when the attribute hierarchy and the structural $\mathbf{Q}$-matrix are identifiable. The identifiability of HLAMs is closely connected to the uniqueness of tensor decompositions, because the probability distribution of an HLAM can be written as a mixture of highly constrained higher-order tensors. In particular, HLAMs can be viewed as a special family of restricted latent class models, with the $\mathbf{Q}$-matrix

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

imposing constraints on the model parameters. However, related works on the identifiability of latent class models and the uniqueness of tensor decompositions (e.g. Allman et al., 2009; Anandkumar et al., 2014) cannot be applied directly to HLAMs because of the constraints induced by the **Q**-matrix.

To tackle identifiability under such structural constraints, several recent works (Xu, 2017; Xu and Shang, 2018; Gu and Xu, 2019b; Fang et al., 2019; Gu and Xu, 2020, 2019a; Chen et al., 2020) have proposed identifiability conditions for latent attribute models. However, most of these studies consider scenarios without any attribute hierarchy; Gu and Xu (2020) assumed that both the true **Q**-matrix and true configurations of the attribute patterns are known and fixed; and Gu and Xu (2019a) considered the problem of learning the set of truly existing attribute patterns, but assumed the **Q**-matrix is correctly specified beforehand. Nevertheless, none of these works directly consider the hierarchical graphical structure of the attribute hierarchy. Therefore, their results cannot provide explicit and sharp identifiability conditions for an HLAM. On the other hand, in the cognitive diagnostic modeling literature, researchers (Köhn and Chiu, 2019; Cai et al., 2018) have recently studied the "completeness" of the **Q**-matrix, a relevant concept revisited in Section 3, under an attribute hierarchy. However, these results cannot ensure that we can uniquely identify the model parameters that determine the probabilistic HLAM. In summary, establishing identifiability without assuming any knowledge of the **Q**-matrix and the attribute hierarchy remains unaddressed in the literature, and is indeed a technically challenging task.

We address this identifiability question for popular HLAMs under an arbitrary attribute hierarchy. We develop explicit sufficient conditions for identifying the attribute hierarchy, **Q**-matrix, and all model parameters in an HLAM. These sufficient conditions become necessary

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

when the latent pattern space is saturated with no hierarchy. For nonempty hierarchies, we discuss the necessity of these individual conditions, and relax them in several nontrivial and interesting ways. We then establish the fully general necessary and sufficient identifiability conditions for the attribute hierarchy and model parameters under a fixed $\mathbf{Q}$-matrix. Our results sharply characterize the roles played by different types of attributes in an attribute hierarchy graph, and can be used to assess the validity of an estimated HLAM obtained from any estimation method. They also provide insights into designing useful diagnostic tests under an attribute hierarchy with minimal restrictions.

The rest of the paper is organized as follows. In Section 2, we introduce the HLAM model setup. In Section 3, we present sufficient conditions for the identifiability of $\mathbf{Q}$, the attribute hierarchy, and the model parameters. In Section 4, to close the gap between the necessity and sufficiency of the identifiability conditions, we focus on the case where $\mathbf{Q}$ is fixed, and derive the fully general necessary and sufficient conditions for identifying the attribute hierarchy and model parameters. In Section 5, we extend the identifiability result to other types of HLAMs that have potentially more parameters than those studied in Sections 3 and 4. Section 6 concludes the paper. All technical proofs are presented in the Supplementary Material.

2. Model Setup and Examples

This section introduces the HLAM model setup. We first introduce some notation. For an integer m, denote $[m] = \{1, 2, \ldots, m\}$. For a set $\mathcal{A}$, denote its cardinality by $|\mathcal{A}|$. Denote $K \times K$ identity matrix by I_K, and K-dimensional all-one and all-zero vectors by $\mathbf{1}_K$ and $\mathbf{0}_K$, respectively.

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

An HLAM consists of two types of subject-specific binary variables, namely, the observed responses $\boldsymbol{r} = (r_1, \ldots, r_J) \in \{0,1\}^J$ to J *items*, and the latent attribute pattern $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K) \in \{0,1\}^K$, with α_k indicating the mastery or deficiency of the kth *attribute*. In this work, K is assumed to be known and fixed. This assumption is well suited for the motivating applications in cognitive diagnosis, where the number and the real-world meanings of the latent attributes are usually known in the context of the application, and it is of interest to identify and learn other quantities from the data. Next, we describe the distribution of the latent attributes. Attribute k is said to be a prerequisite of attribute ℓ, denoted by $k \to \ell$, if any pattern $\boldsymbol{\alpha}$ with $\alpha_k = 0$ and $\alpha_\ell = 1$ is "forbidden" to exist. This is a common assumption in applications such as cognitive diagnosis that model subjects' learning processes (Leighton et al., 2004; Templin and Bradshaw, 2014). A subject's latent pattern $\boldsymbol{a}$ is assumed to follow a categorical distribution of population proportion parameters $\boldsymbol{p} = (p_{\boldsymbol{\alpha}}, \boldsymbol{\alpha} \in \{0,1\}^K)$, with $p_{\boldsymbol{\alpha}} \geq 0$ and $\sum_{\boldsymbol{\alpha}} p_{\boldsymbol{\alpha}} = 1$. In particular, any pattern $\boldsymbol{\alpha}$ not respecting the hierarchy is deemed impossible to exist with population proportion $p_{\boldsymbol{\alpha}} = 0$. An attribute hierarchy is a set of prerequisite relations among the K attributes:

$$\mathcal{E} = \{k \to \ell : \text{attribute } k \text{ is a prerequisite for } \ell\}.$$

In general, an attribute hierarchy $\mathcal{E}$ implies a directed acyclic graph (DAG) among the K attributes, with no directed cycles. This graph constrains which attribute patterns are permissible or forbidden. Specifically, any $\mathcal{E}$ induces a set of allowable configurations of attribute patterns out of $\{0,1\}^K$, which we denote by $\mathcal{A}(\mathcal{E})$, or simply $\mathcal{A}$ when it causes no confusion. For an arbitrary $\mathcal{E}$, the all-zero and all-one attribute patterns $\mathbf{0}_K$ and $\mathbf{1}_K$ always belong to the induced $\mathcal{A}$. This is because no prerequisite relation among the attributes can rule out the existence of a pattern possessing all or none of the attributes. When there is no

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

attribute hierarchy among the K attributes, $\mathcal{E} = \varnothing$ and $\mathcal{A} = \{0,1\}^K$. The set $\mathcal{A}$ is a proper subset of $\{0,1\}^K$ if $\mathcal{E} \neq \varnothing$. An attribute hierarchy determines the sparsity pattern of the vector of proportion parameters $\boldsymbol{p}$, because $p_{\boldsymbol{\alpha}} > 0$ if and only if $\boldsymbol{\alpha} \in \mathcal{A}(\mathcal{E})$, that is, if and only if $\boldsymbol{\alpha}$ is permissible under $\mathcal{E}$. In this sense, a nonempty attribute hierarchy necessarily leads to degenerate parameter space for $\boldsymbol{p}$, because certain entries of $\boldsymbol{p}$ will be constrained to be zero.

In an attribute hierarchy in cognitive diagnosis, the case of $k \to \ell$ and $\ell \to k$ indicates that the skill attributes α_k and α_ℓ are prerequisites for each other, which is not interpretable, and hence is not used in modeling. Similarly, having any cycle in the attribute hierarchy graph in the form of $k_1 \to k_2 \to \cdots \to k_m \to k_1$ is also not interpretable. Therefore, a DAG structure among the latent attributes is well suited to describe the hierarchical nature of the attributes that carry these substantive meanings. Note that the DAG of an attribute hierarchy in an HLAM differs from that in a Bayesian network (Pearl, 1986), because the former encodes hard constraints on which variable patterns are permissible/forbidden, whereas the latter encodes the conditional independence relations among the variables.

Remark 1. *Our attribute hierarchy constraints that "$k \to \ell$ implies $\alpha_k = 0$ and $\alpha_\ell = 1$ is impossible" have interesting connections to some other constraints in the statistics literature. In variable selection, where the main effects of the variables and their interaction effects may be present, the effect heredity principle (Hamada and Wu, 1992) posits that a corresponding interaction effect only exists if the main effects of the variables exist. In particular, with θ_i and θ_j denoting the continuous regression coefficients associated with two heredity terms, Yuan et al. (2009) used a linear inequality $\theta_i \leq \theta_j$ (a continuous relaxation of the hard constraints on the binary indicators of the variable inclusions) to cleverly enforce the*

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

heredity constraint and facilitate computation. In causal inferences, the monotonicity constraint in instrumental variable analyses (Hernán and Robins, 2006; Swanson et al., 2015) posits that if the instrumental variable satisfies $z_1 < z_2$, then the counterfactual treatment is a nondecreasing function of the instrument, that is, $X_i^{z_1} \leq X_i^{z_2}$, for all subjects i. A key difference between the attribute hierarchy constraints and the aforementioned constraints is the involvement of many latent variables in HLAMs; indeed, all the $\alpha_1, \ldots, \alpha_K$ among which the hierarchical constraints exist are latent. The binary patterns $\boldsymbol{\alpha}$ that respect the attribute hierarchy $\mathcal{E}$ follow an unknown categorical distribution with parameters $\boldsymbol{p} = (p_{\boldsymbol{\alpha}})$ with $\sum_{\boldsymbol{\alpha}} p_{\boldsymbol{\alpha}} = 1$, and the observed data distribution is obtained after marginalizing out the latent structure, which is quite complicated.

Example 1. Fig 1 presents several hierarchies with the size of the associated $\mathcal{A}$, where a dotted arrow from α_k to α_ℓ indicates $k \to \ell$ and k is a direct prerequisite for ℓ. Note that under the hierarchy in Fig 1(a), the prerequisite $1 \to 3$ is an indirect prerequisite implied by $1 \to 2$ (or 4) and 2 (or 4) $\to 3$.

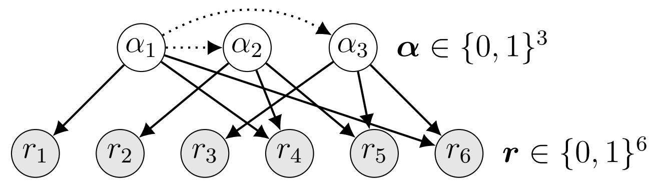
Figure 1: Different attribute hierarchies among binary attributes for $K = 4$, where $|\{0,1\}^4| = 16$. For example, the set of allowed attribute patterns under hierarchy (a) is $\mathcal{A}_1 = \{\mathbf{0}_4$, (1000), (1100), (1001), (1101), $\mathbf{1}_4\}$.

On top of the model of the latent attributes, an HLAM uses a $J \times K$ binary matrix $\mathbf{Q} = (q_{j,k})$ to encode the structural relationship between the J observed response variables and the K latent attributes. In cognitive diagnostic assessments, the matrix $\mathbf{Q}$ is often specified by domain experts to summarize which abilities each test item targets (Tatsuoka,

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

$$\mathbf{Q}_{6\times 3} := \begin{pmatrix} \boldsymbol{q}_1 \\ \boldsymbol{q}_2 \\ \boldsymbol{q}_3 \\ \boldsymbol{q}_4 \\ \boldsymbol{q}_5 \\ \boldsymbol{q}_6 \end{pmatrix} := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix};$$

(a) **Q**-matrix



(b) graphical model & attribute hierarchy

Figure 2: A binary structural matrix and the corresponding graphical model, with (solid) directed edges from the latent to the observed variables representing dependencies. Below the observed variables in (b) are the row vectors of $\mathbf{Q}_{6\times 3}$, that is, the item loading vectors. The dotted arrows indicate the attribute hierarchy with $\mathcal{E} = \{1 \to 2, 1 \to 3\}$ and $\mathcal{A} = \{\mathbf{0}_3, (100), (110), (101), \mathbf{1}_3\}$.

1990; von Davier, 2008; Rupp et al., 2010; de la Torre, 2011). Specifically, $q_{j,k} = 1$ if and only if the response r_j to the jth item is statistically dependent on the latent variable α_k. The distribution of r_j, that is, $\theta_{j,\boldsymbol{\alpha}} := \mathbb{P}(r_j = 1 \mid \boldsymbol{\alpha})$, depends only on its "parent" latent attributes α_k that are connected to r_j, that is, $\{\alpha_k : q_{j,k} = 1\}$. The structural matrix $\mathbf{Q}$ naturally induces a bipartite graph connecting the latent and the observed variables, with edges corresponding to entries of "1" in $\mathbf{Q} = (q_{j,k})$. Fig 2 presents an example of a structural matrix $\mathbf{Q}$ and its corresponding directed graphical model between the $K = 3$ latent attributes and $J = 6$ observed variables. The solid edges from the latent attributes to the observed variables are specified by $\mathbf{Q}_{6\times 3}$. Furthermore, the observed responses to the J items are conditionally independent, given the latent attribute pattern $\boldsymbol{\alpha}$.

In the psychometrics literature, various HLAMs adopting the **Q**-matrix concept have been proposed with the goal of diagnosing targeted attributes (Junker and Sijtsma, 2001; Templin and Henson, 2006; von Davier, 2008; Henson et al., 2009; de la Torre, 2011). These are often called cognitive diagnostic models. The general family of latent attribute models is also widely used in other scientific areas, including psychiatric evaluations (Templin and

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

Henson, 2006; de la Torre et al., 2018) used to diagnose patients' mental disorders, and epidemiological diagnose of disease etiologies (Wu et al., 2016, 2017; O'Brien et al., 2019). These applications share the common key interest of identifying multivariate discrete latent attributes.

In this work, we mainly focus on a popular and fundamental type of modeling assumption under such a framework. As to be revealed soon, this modeling assumption also has close connections to Boolean matrix factorizations (Ravanbakhsh et al., 2016; Rukat et al., 2017). Specifically, we mainly consider HLAMs that assume a logical *ideal response* $\Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}}$, given an attribute pattern $\boldsymbol{\alpha}$ and an item loading vector $\boldsymbol{q}_j$, in the noiseless case. Then, item-level noise parameters are introduced to account for the uncertainty of the observations. The following are two popular ways of defining the ideal response.

The first is the deterministic input noisy output "And" gate (DINA) model (Junker and Sijtsma, 2001; de la Torre and Douglas, 2004; von Davier, 2014). The DINA model assumes a conjunctive relationship between the attributes. The ideal response of attribute pattern $\boldsymbol{\alpha}$ to item j is

$$\text{(DINA ideal response)} \qquad \Gamma^{\text{AND}}_{\boldsymbol{q}_j,\boldsymbol{\alpha}} = \prod_{k=1}^{K} \alpha_k^{q_{j,k}}, \tag{2.1}$$

where the convention $0^0 \equiv 1$ is adopted. It is not hard to check that the above definition is equivalent to

$$\Gamma^{\text{AND}}_{\boldsymbol{q}_j,\boldsymbol{\alpha}} = \mathbb{1}(\alpha_k \geq q_{j,k} \text{ for all the } k \in [K]). \tag{2.2}$$

This definition intuitively and explicitly explains that the DINA adopts a *conjunctive* modeling assumption, because only if a subject with attribute pattern $\boldsymbol{\alpha}$ possesses all of the attributes required by the loading vector $\boldsymbol{q}_j$ is he/she considered capable of this item j and to have $\Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}} = 1$. Such a conjunctive relationship is often assumed for diagnosis of students' mastery or deficiency of skill attributes in educational assessments, and $\Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}}$ naturally indicates whether a student with $\boldsymbol{\alpha}$ has mastered all the attributes required by the test item j. With $\Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}}$ in (2.1), the uncertainty of the responses is further modeled by the item-specific Bernoulli parameters

$$\theta_j^+ = \mathbb{P}(r_j = 1 \mid \Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}} = 1), \quad \theta_j^- = \mathbb{P}(r_j = 1 \mid \Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}} = 0), \tag{2.3}$$

where $\theta_j^+ > \theta_j^-$ is assumed for identifiability. For each item j, the ideal response $\Gamma_{\boldsymbol{q}_j,\cdot}$, if viewed as a function of attribute patterns, divides the patterns into two latent classes, $\{\boldsymbol{\alpha} : \Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}} = 1\}$ and $\{\boldsymbol{\alpha} : \Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}} = 0\}$, for which the item parameters quantify the noise levels of the response to item j that deviates from the ideal response. Note that $\theta_{j,\boldsymbol{\alpha}}$ is equal to either θ_j^+ or θ_j^-, depending on the ideal response $\Gamma_{j,\boldsymbol{\alpha}}$. Denote the item parameter vectors by $\boldsymbol{\theta}^+ = (\theta_1^+, \ldots, \theta_J^+)^\top$ and $\boldsymbol{\theta}^- = (\theta_1^-, \ldots, \theta_J^-)^\top$.

The second model is the deterministic input noisy output "Or" gate (DINO) model (Templin and Henson, 2006). The DINO model assumes the following ideal response,

$$\text{(DINO ideal response)} \qquad \Gamma^{\text{OR}}_{\boldsymbol{q}_j,\boldsymbol{\alpha}} = I(q_{j,k} = \alpha_k = 1 \text{ for at least one } k). \tag{2.4}$$

Such disjunctive relationships are often assumed in psychiatric measurements of mental disorders (Templin and Henson, 2006; de la Torre et al., 2018). With $\Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}}$ in (2.4), the uncertainty of the responses is modeled by the item-specific parameters defined in (2.3). In the Boolean matrix factorization literature, a similar model has been proposed (Ravanbakhsh et al., 2016). Adapted to the terminology here, Rukat et al. (2017) assumes the ideal response takes the form

$$\text{(equivalent to (2.4))} \qquad \Gamma^{\mathrm{OR}}_{\boldsymbol{q}_j,\boldsymbol{\alpha}} = 1 - \prod_{k=1}^{K}(1-\alpha_k q_{j,k}), \tag{2.5}$$

which is equivalent to the definition in (2.4), although their model constrains all item-level noise parameters to be the same.

We next focus on asymmetric DINA-based HLAMs, because these are popular and fundamental models widely used in the motivating applications of educational cognitive diagnosis. We also study the identifiability of DINO-based HLAMs and another type of HLAM in Section 5. For notational simplicity, we next write $\Gamma^{\mathrm{AND}}_{\boldsymbol{q}_j,\boldsymbol{\alpha}}$ simply as $\Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}}$. Denote by $\Gamma(\mathbf{Q},\mathcal{E})$ the $J\times|\mathcal{A}(\mathcal{E})|$ ideal response matrix with the $(j,\boldsymbol{\alpha})$th entry being $\Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}}$, for $\boldsymbol{\alpha}\in\mathcal{A}(\mathcal{E})$. Under the setup of DINA-based HLAMs, the probability mass function of the J-dimensional random response vector $\mathbf{R}$ takes the form

$$\begin{aligned} P(\mathbf{R}=\boldsymbol{r}\mid \mathbf{Q},\mathcal{E},\boldsymbol{\theta}^+,\boldsymbol{\theta}^-,\boldsymbol{p}) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}(\mathcal{E})} p_{\boldsymbol{\alpha}} \prod_{j=1}^{J} & [\Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}}\theta_j^+ + (1-\Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}})\theta_j^-]^{r_j} \\ & \times [1-\Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}}\theta_j^+ - (1-\Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}})\theta_j^-]^{1-r_j}, \end{aligned}$$

where $\boldsymbol{r}\in\{0,1\}^J$ is an arbitrary response pattern.

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

3. Identifiability of Q, Attribute Hierarchy, and Model Parameters: Establishing Sufficiency

This section presents a main result on the sufficient conditions for the identifiability of $\mathbf{Q}$, $\mathcal{E}$, and the model parameters $\boldsymbol{\theta}^+$, $\boldsymbol{\theta}^-$, and $\boldsymbol{p}$. Following the definition of identifiability in the statistics literature, we say that $(\mathbf{Q}, \mathcal{E}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$ of an HLAM are identifiable if for any $(\mathbf{Q}, \mathcal{E}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$ in the parameter space constrained by $\mathbf{Q}$ and $\mathcal{E}$, there exist no $(\bar{\mathbf{Q}}, \bar{\mathcal{E}}, \bar{\boldsymbol{\theta}}^+, \bar{\boldsymbol{\theta}}^-, \bar{\boldsymbol{p}}) \neq (\mathbf{Q}, \mathcal{E}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$ such that

$$\mathbb{P}(\mathbf{R} = \boldsymbol{r} \mid \bar{\mathbf{Q}}, \bar{\mathcal{E}}, \bar{\boldsymbol{\theta}}^+, \bar{\boldsymbol{\theta}}^-, \bar{\boldsymbol{p}}) = \mathbb{P}(\mathbf{R} = \boldsymbol{r} \mid \mathbf{Q}, \mathcal{E}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p}),\ \forall \boldsymbol{r} \in \{0,1\}^J. \tag{3.1}$$

In this definition, the alternative vector of the proportion parameters $\bar{\boldsymbol{p}}$ is not constrained to have support on $\mathcal{A}(\mathcal{E})$. Instead, the vector $\bar{\boldsymbol{p}}$ should be allowed to have an arbitrary support $\bar{\mathcal{A}}$ potentially resulting from an arbitrary $\bar{\mathcal{E}}$; the goal of establishing identifiability is indeed to develop conditions to ensure that as long as (3.1) holds, one must have $\bar{\boldsymbol{p}} = \boldsymbol{p}$ and $\bar{\mathcal{E}} = \mathcal{E}$ from the equations in (3.1).

Next, we introduce some additional notation and important concepts. Because an attribute hierarchy is a DAG, the K attributes $\{1, 2, \ldots, K\}$ can be arranged in a topological order such that the prerequisite relation "$\to$" only happens in one direction; in other words, we can assume without loss of generality that $k \to \ell$ only if $k < \ell$. Define the following *reachability matrix* $\mathbf{E}$ among the K attributes under the attribute hierarchy. Here $\mathbf{E} = (e_{k,\ell})$ is a $K \times K$ binary matrix, where $e_{k,k} = 1$ for all $k \in [K]$ and $e_{\ell,k} = 1$, if attribute k is a direct or indirect prerequisite for attribute ℓ. In cognitive diagnosis, the concept of a reachability matrix was first considered in Tatsuoka (1986) to represent the direct and indirect

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

relationships between attributes. It is not hard to see that if the attributes $1, 2, \ldots, K$ are in a topological order, the reachability matrix $\mathbf{E}$ is a lower-triangular matrix with all the diagonal entries being one.

Under DINA-based HLAMs, any nonempty attribute hierarchy $\mathcal{E}$ defines an equivalence relation on the set of all $\mathbf{Q}$-matrices. To see this, recall that $\Gamma(\mathbf{Q}, \mathcal{E})$ denotes the $J \times |\mathcal{A}(\mathcal{E})|$ ideal response matrix. If $\Gamma(\mathbf{Q}_1, \mathcal{E}) = \Gamma(\mathbf{Q}_2, \mathcal{E})$, then $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are said to be in the same $\mathcal{E}$-induced equivalence class and we denote this by $\mathbf{Q}_1 \overset{\mathcal{E}}{\sim} \mathbf{Q}_2$. To interpret, if under a certain hierarchy $\mathcal{E}$, two different $\mathbf{Q}$-matrices lead to identical ideal responses for all the permissible latent patterns in $\mathcal{A}(\mathcal{E})$, then these two $\mathbf{Q}$-matrices are indistinguishable based on the response data, and should be treated as equivalent. The following example illustrates how an attribute hierarchy determines a set of equivalent $\mathbf{Q}$-matrices.

Example 2. Consider the attribute hierarchy $\mathcal{E} = \{1 \to 2,\ 1 \to 3\}$ in Fig 2, which results in $\mathcal{A}(\mathcal{E}) = \{\mathbf{0}_3, (100), (110), (101), \mathbf{1}_3\}$. The identity matrix I_3 is equivalent to the reachability matrix $\mathbf{E}$ under $\mathcal{E}$ and

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \overset{\mathcal{E}}{\sim} \mathbf{E} = \begin{pmatrix} 1 & 0 & 0 \\ \mathbf{1} & 1 & 0 \\ \mathbf{1} & 0 & 1 \end{pmatrix} \overset{\mathcal{E}}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ * & 1 & 0 \\ * & 0 & 1 \end{pmatrix}, \tag{3.2}$$

where the "$*$"'s in the third matrix above indicate unspecified values, any of which can be either zero or one. This equivalence is because the attribute α_1 serves as a prerequisite for both α_2 and α_3, and any item loading vector $\boldsymbol{q}_j$ measuring α_2 or α_3 is equivalent to a modified one that also measures α_1, in terms of classifying the patterns in $\mathcal{A}$ into two categories, $\{\boldsymbol{\alpha} : \Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}} = 1\}$ and $\{\boldsymbol{\alpha} : \Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}} = 0\}$. Note that any $\mathbf{Q}$-matrix equivalent to I_K under $\mathcal{E} = \{1 \to 2,\ 1 \to 3\}$ must take the form of the third $\mathbf{Q}$-matrix in (3.2). Under a DINA-based HLAM, if the true $\mathbf{Q}$-matrix $\mathbf{Q}^{\text{true}}$ is not known, then any other $\mathbf{Q}$ with $\mathbf{Q} \overset{\mathcal{E}}{\sim} \mathbf{Q}^{\text{true}}$ cannot be distinguished from $\mathbf{Q}^{\text{true}}$ based on the observations, even if the

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

continuous parameters $(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$ are all known. This is because the ideal response matrix $\Gamma(\mathbf{Q}, \mathcal{E})$ is the key latent structure underlying a DINA-based HLAM, and if $\mathbf{Q} \overset{\mathcal{E}}{\sim} \mathbf{Q}^{\text{true}}$ (equivalently, $\Gamma(\mathbf{Q}, \mathcal{E}) = \Gamma(\mathbf{Q}^{\text{true}}, \mathcal{E})$), then $\mathbf{Q}$ and $\mathbf{Q}^{\text{true}}$ are inherently not distinguishable.

Given any hierarchy $\mathcal{E}$, the equivalence $I_K \overset{\mathcal{E}}{\sim} \mathbf{E}$ is always true; see, for example, Eq. (3.2) in Example 2. Before presenting the theorem on sufficient conditions for identifiability, we introduce two useful operations on a $\mathbf{Q}$-matrix, given an attribute hierarchy $\mathcal{E}$: the "densifying" operation $\mathcal{D}^{\mathcal{E}}(\cdot)$, and the "sparsifying" operation $\mathcal{S}^{\mathcal{E}}(\cdot)$, as follows.

Definition 1. *Given an attribute hierarchy $\mathcal{E}$ and a matrix $\mathbf{Q}$, do the following: for any $q_{j,h} = 1$ and $k \to h$, set $q_{j,k}$ to one and obtain a modified matrix $\mathcal{D}^{\mathcal{E}}(\mathbf{Q})$. This $\mathcal{D}^{\mathcal{E}}(\mathbf{Q})$ is said to be a "densified" version of $\mathbf{Q}$.*

Definition 2. *Given an attribute hierarchy $\mathcal{E}$ and a matrix $\mathbf{Q}$, do the following: for any $q_{j,h} = 1$ and $k \to h$, set $q_{j,k}$ to zero and obtain a modified matrix $\mathcal{S}^{\mathcal{E}}(\mathbf{Q})$. This $\mathcal{S}^{\mathcal{E}}(\mathbf{Q})$ is said to be a "sparsified" version of $\mathbf{Q}$.*

Under the above two definitions, given any attribute hierarchy $\mathcal{E}$, the two statements $\mathcal{D}^{\mathcal{E}}(I_K) = \mathbf{E}$ and $\mathcal{S}^{\mathcal{E}}(\mathbf{E}) = I_K$ always hold. Specifically, $\mathcal{D}^{\mathcal{E}}(I_K) = \mathbf{E}$ means that in the special case where $J = K$ and $\mathbf{Q}$ takes the form of an identity matrix I_K, densifying such a $\mathbf{Q}$ always gives a $K \times K$ reachability matrix $\mathbf{E}$ under the hierarchy $\mathcal{E}$. Similarly, $\mathcal{S}^{\mathcal{E}}(\mathbf{E}) = I_K$ means that in another special case where $J = K$ and $\mathbf{Q}$ takes the form of the reachability matrix $\mathbf{E}$, sparsifying $\mathbf{Q}$ always gives the identity matrix I_K. These two special examples illustrate the definitions of the sparsifying/densifying operations on $\mathbf{Q}$-matrices and the relationship between $\mathbf{E}$ and I_K. In cognitive diagnosis, the densified $\mathbf{Q}$-matrix with all row vectors respecting the attribute hierarchy $\mathcal{E}$ is also said to satisfy the "restricted $\mathbf{Q}$-matrix design" (e.g., Cai et al., 2018; Tu et al., 2019); for such $\mathbf{Q}$, it holds that $\mathbf{Q} = \mathcal{D}^{\mathcal{E}}(\mathbf{Q})$. Note that

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

the sparsifying and densifying operations modify $\mathbf{Q}$ only within the same equivalence class. Indeed, $\mathcal{D}^{\mathcal{E}}(\mathbf{Q})$ denotes the densest $\mathbf{Q}$ with the largest number of ones in the equivalence class, while $\mathcal{S}^{\mathcal{E}}(\mathbf{Q})$ denotes the sparsest $\mathbf{Q}$ with the largest number of zeros in the equivalence class. In the special case with an empty attribute hierarchy, each equivalence class of $\mathbf{Q}$ contains only one element, which is $\mathbf{Q}$ itself, so $\mathbf{Q} = \mathcal{D}^{\mathcal{E}}(\mathbf{Q}) = \mathcal{S}^{\mathcal{E}}(\mathbf{Q})$, for $\mathcal{E} = \varnothing$. As shown in the following theorem, our identifiability conditions are essentially requirements on the equivalence class of $\mathbf{Q}$ described using the densifying and sparsifying operations.

Theorem 1. *Consider an HLAM under the DINA model and an attribute hierarchy $\mathcal{E}$. Then, $(\Gamma(\mathbf{Q}, \mathcal{E})$, $\boldsymbol{\theta}^+$, $\boldsymbol{\theta}^-$, $\boldsymbol{p})$ are jointly identifiable if the true $\mathbf{Q}$ satisfies the following conditions:*

A. The $\mathbf{Q}$ contains the $K \times K$ submatrix $\mathbf{Q}^0$ that is equivalent to the identity matrix I_K under the hierarchy $\mathcal{E}$.

(Without loss of generality, assume the first K rows of $\mathbf{Q}$ form $\mathbf{Q}^0$, and denote the remaining submatrix of $\mathbf{Q}$ by $\mathbf{Q}^\star$.)

B. The $\mathcal{S}^{\mathcal{E}}(\mathbf{Q})$, the sparsified version of $\mathbf{Q}$, has at least three ones in each column.

C. The $\mathcal{D}^{\mathcal{E}}(\mathbf{Q}^\star)$, the densified version of the submatrix $\mathbf{Q}^\star$, contains K distinct column vectors.

Furthermore, Conditions A, B, and C are necessary and sufficient when no hierarchy exists with $p_{\boldsymbol{\alpha}} > 0$, for all $\boldsymbol{\alpha} \in \{0,1\}^K$.

We make several remarks on the relationship between the new theory and that of existing works.

Remark 2. In the cognitive diagnostic modeling literature, a $\mathbf{Q}$-matrix is said to be “complete” if it can distinguish all 2^K latent attribute profiles (Chiu et al., 2009). When the latent pattern space $\mathcal{A}$ is saturated with $\mathcal{A} = \{0,1\}^K$, the completeness of $\mathbf{Q}$ is a natural and necessary requirement for identifiability. When $\mathcal{A} = \{0,1\}^K$, the $\mathbf{Q}$-matrix is complete if it contains all K distinct standard basis vectors as row vectors, that is, $\mathbf{Q}$ contains an I_K. When a certain attribute hierarchy $\mathcal{E}$ exists leading to some $\mathcal{A} \subsetneq \{0,1\}^K$, the requirement for the "completeness" of $\mathbf{Q}$ changes. Recently, Köhn and Chiu (2019) and Cai et al. (2018) studied conditions for the completeness of $\mathbf{Q}$ under attribute hierarchies. However, these conditions cannot ensure that the entire probabilistic model structure involving $\mathbf{Q}$, $\mathcal{E}$, and the parameters $\boldsymbol{p}$, $\boldsymbol{\theta}^+$ and $\boldsymbol{\theta}^-$ are identifiable and estimable from the data. To the best of our knowledge, Theorem 1 establishes the first identifiability result under an arbitrary attribute hierarchy in the literature. Condition A in Theorem 1 is equivalent to requiring that the sparsified $\mathcal{S}^{\mathcal{E}}(\mathbf{Q})$ contains an I_K. Therefore, the combination of Conditions A and B is equivalent to the following statement about $\mathcal{S}^{\mathcal{E}}(\mathbf{Q})$: $\mathcal{S}^{\mathcal{E}}(\mathbf{Q})$ contains an I_K and each column contains at least three ones.

Remark 3. As stated in the last part of Theorem 1, when there is no attribute hierarchy with $\mathcal{E} = \varnothing$, Conditions A, B, and C become necessary and sufficient for the identifiability of both $\mathbf{Q}$ and ($\mathcal{E}$, $\boldsymbol{\theta}^+$, $\boldsymbol{\theta}^-$, $\boldsymbol{p}$). In such a special case with $\mathcal{E} = \varnothing$, Gu and Xu (2021) established the following necessary and sufficient identifiability conditions: "*completeness*", which requires that the true $\mathbf{Q}$ contain an identity submatrix I_K, "*repeated-measurement*", which requires that $\mathbf{Q}$ have at least three ones in each column, and "*distinctiveness*", which requires that in addition to containing an I_K, $\mathbf{Q}$ should contain distinct column vectors in the remaining submatrix; we denote these three requirements as Conditions A^0, B^0, and C^0, respectively. Our current conditions A, B, and C in Theorem 1 can be thought of as "*$\mathcal{E}$-completeness*," "*$\mathcal{E}$-repeated-measurement*," "*$\mathcal{E}$-distinctiveness*," given an attribute hierarchy

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

$\mathcal{E}$. When $\mathcal{E} = \varnothing$, $\mathcal{S}^{\mathcal{E}}(\mathbf{Q}) = \mathcal{D}^{\mathcal{E}}(\mathbf{Q}) = \mathbf{Q}$ holds; as a result, Condition A exactly requires that $\mathbf{Q}$ itself contain a submatrix I_K; similarly, Conditions B and C exactly reduce to the conditions B^0 and C^0, respectively, on $\mathbf{Q}$ itself. Indeed, in such cases with $\mathcal{E} = \varnothing$, the current conditions of "*$\mathcal{E}$-completeness*," "*$\mathcal{E}$-repeated-measurement*," and "*$\mathcal{E}$-distinctiveness*" reduce to the "*completeness*," "*repeated-measurement*," and "*distinctiveness*" conditions proposed in Gu and Xu (2021). Establishing identifiability under an arbitrary attribute hierarchy $\mathcal{E}$, as in Theorem 1, is technically much more challenging than the existing result for $\mathcal{E} = \varnothing$. Moreover, in Section 4, under a fixed $\mathbf{Q}$-matrix, we investigate how the necessity of the identifiability conditions changes when there is a nonempty hierarchy.

Theorem 1 ensures that the discrete ideal response structure $\Gamma(\mathbf{Q}, \mathcal{E})$ and all associated model parameters ($\boldsymbol{\theta}^+$, $\boldsymbol{\theta}^-$, $\boldsymbol{p}$) are identifiable. The following proposition complements this conclusion and further establishes the identifiability of $\mathcal{E}$ and $\mathbf{Q}$ based on Theorem 1.

Proposition 1. *Consider a DINA-based HLAM. In addition to Conditions A–C in Theorem 1, if the true $\mathbf{Q}$ is known to contain an I_K, then $(\mathcal{E}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$ are identifiable. On the other hand, for $\mathbf{Q}$ must contain an I_K to ensure that an arbitrary $\mathcal{E}$ is identifiable.*

Proposition 2. *Consider a DINA-based HLAM. If Conditions A–C in Theorem 1 are satisfied and the true $\mathbf{Q}$ is known in part to contain a submatrix I_K for certain K items, then the equivalence class of $\mathbf{Q}$ defined by the attribute hierarchy $\mathcal{E}$ is identifiable. That is, the specific $\mathbf{Q}$ is not strictly identifiable within its equivalence class under any $\mathcal{E} \neq \varnothing$, but the densified $\mathcal{D}^{\mathcal{E}}(\mathbf{Q})$ and the sparsified $\mathcal{S}^{\mathcal{E}}(\mathbf{Q})$ are identifiable.*

The statement in Proposition 2 that $\mathbf{Q}$ is identifiable only up to its equivalence class is inherent to all DINA- or DINO-type HLAMs, and is an inevitable consequence of any nonempty attribute hierarchy $\mathcal{E} \neq \varnothing$; see Example 2. However, this statement does not undermine the efficacy of the identifiability conclusion, because $\mathcal{D}^{\mathcal{E}}(\mathbf{Q})$ and $\mathcal{S}^{\mathcal{E}}(\mathbf{Q})$ themselves are still identifiable and provide practical interpretability of the structural matrix. The following toy example shows how to apply Theorem 1 to check identifiability.

Example 3. Consider the attribute hierarchy $\{\alpha_1 \to \alpha_2,\, \alpha_1 \to \alpha_3\}$ among $K = 3$ attributes, as in Fig 2. The following 8×3 structural matrix $\mathbf{Q}$ satisfies Conditions A, B, and C in Theorem 1. In particular, the first three rows of $\mathbf{Q}$ serve as $\mathbf{Q}^0$ in Condition A, and the last five rows serve as $\mathbf{Q}^\star$. In the following display, the matrix entries modified by the sparsifying operation in Condition B and the densifying operation in Condition C are highlighted and in bold. The resulting $\mathcal{S}^{\mathcal{E}}(\mathbf{Q})$ and $\mathcal{D}^{\mathcal{E}}(\mathbf{Q})$ satisfy the requirements in Conditions B and C. Thus, the HLAM associated with $\mathbf{Q}$ is identifiable.

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}^0 \\ \hline \mathbf{Q}^\star \end{pmatrix} = \begin{pmatrix} & I_3 & \\ \hline 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \overset{\text{Sparsify}}{\Longrightarrow} \mathcal{S}^{\mathcal{E}}(\mathbf{Q}) = \begin{pmatrix} & I_3 & \\ \hline 1 & 0 & 0 \\ 1 & 0 & 0 \\ \mathbf{0} & 1 & 0 \\ 0 & 0 & 1 \\ \mathbf{0} & 1 & 1 \end{pmatrix}; \tag{3.3}$$

$$\overset{\text{Densify}}{\Longrightarrow} \mathcal{D}^{\mathcal{E}}(\mathbf{Q}) = \begin{pmatrix} & \mathbf{E} & \\ \hline 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ \mathbf{1} & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}. \tag{3.4}$$

When estimating an HLAM with the goal of recovering the ideal response structure $\Gamma(\mathbf{Q}, \mathcal{E})$ and the continuous parameters, Theorem 1 guarantees that Conditions A, B, and C suffice and are close to being necessary. If the goal is to uniquely determine the attribute hierarchy from the identified $\Gamma(\mathbf{Q}, \mathcal{E})$, the additional condition that $\mathbf{Q}$ contains an I_K becomes necessary. This phenomenon can be better understood by relating it to the identification of a

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

factor loading matrix in the factor analysis (Anderson, 2009; Bai and Li, 2012); in this case, the loading matrix is often required to include an identity submatrix or satisfy certain rank constraints, because otherwise it cannot be identifiable, owing to rotational indeterminacy.

Existing results for the identifiability of *non-hierarchical* latent attribute models (i.e., with an empty graph $\mathcal{E} = \varnothing$ in Gu and Xu, 2021) adopt a key assumption that $p_{\boldsymbol{\alpha}} > 0$ for any possible binary pattern $\boldsymbol{\alpha} \in \{0,1\}^K$, and the proofs in Gu and Xu (2021) rely heavily on this assumption on $\boldsymbol{p}$ to derive the identifiability conditions. Importantly, when the assumption that "$p_{\boldsymbol{\alpha}} > 0$, for any $\boldsymbol{\alpha} \in \{0,1\}^K$" fails to hold, the proof in Gu and Xu (2021) no longer holds, and we cannot simply modify their conclusions. Rather, we need to carefully analyze the polynomial systems arising from the probability mass function (PMF) of the observed $\mathbf{R}$ to derive suitable identifiability conditions.

Note that dealing with a degenerate parameter space of $\boldsymbol{p}$ under an attribute hierarchy $\mathcal{E}$ requires quite delicate algebraic work. Specifically, our proof technique of identifiability is based on investigating the conditions under which the highly complex and $\mathbf{Q}$-matrix-constrained polynomial equations given by the PMF of the vector $\mathbf{R}$ has unique roots; the uniqueness of the roots indicates the identifiability of the parameters. When using this proof technique, we start by inspecting the polynomial equations $\mathbb{P}(\mathbf{R} = \boldsymbol{r} \mid \bar{\mathbf{Q}}, \bar{\mathcal{E}}, \bar{\boldsymbol{\theta}}^+, \bar{\boldsymbol{\theta}}^-, \bar{\boldsymbol{p}}) = \mathbb{P}(\mathbf{R} = \boldsymbol{r} \mid \mathbf{Q}, \mathcal{E}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$, $\forall \boldsymbol{r} \in \{0,1\}^J$ for the unknown true parameters $(\mathbf{Q}, \mathcal{E}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$ and arbitrary alternative parameters $(\bar{\mathbf{Q}}, \bar{\mathcal{E}}, \bar{\boldsymbol{\theta}}^+, \bar{\boldsymbol{\theta}}^-, \bar{\boldsymbol{p}})$, and investigate which conditions guarantee that the alternative parameters are identical to the true ones. Under an unknown attribute hierarchy $\mathcal{E}$, certain true proportions $p_{\boldsymbol{\alpha}}$ are equal to zero, but we do not know which ones these are. Therefore, complex constraints on polynomial equations occur, because certain terms vanish from the one side of the equation (corresponding to the true parameters

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

$p_{\boldsymbol{\alpha}}$) but do not vanish from the other side (corresponding to the unknown alternative parameters $\bar{p}_{\boldsymbol{\alpha}}$; we do not know which $\bar{p}_{\boldsymbol{\alpha}} = 0$ out of all the possible $\boldsymbol{\alpha} \in \{0,1\}^K$). This fact makes studying the identifiability issues in the current work considerably harder and quite different from existing results for non-hierarchical latent attribute models (e.g., Gu and Xu, 2021).

As stated at the end of Theorem 1, Conditions A, B, and C become sufficient and necessary for identifiability when there is no hierarchy among the attributes. Interestingly, the necessity of these conditions changes subtly when there is a nonempty attribute hierarchy.

4. Identifiability of Attribute Hierarchy and Model Parameters: Pushing Toward Necessity

In order to close the gap between necessity and sufficiency, we thoroughly investigate the necessity of the identifiability conditions for $(\mathcal{E}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$ under the assumption that $\mathbf{Q}$ is known and fixed. In Subsection 4.1, we first investigate the necessity of the conditions proposed in Section 3 individually, to gain insight into how the necessity changes as the attribute hierarchy changes. Then, in Subsection 4.2, we establish the general necessary and sufficient conditions for identifying the attribute hierarchy and other parameters under an arbitrary hierarchy graph $\mathcal{E}$.

4.1 Investigating the Necessity of Conditions A, B, and C Individually

Our first result establishes the necessity of Condition A in Theorem 1.

Proposition 3. *Consider a DINA-based HLAM. Condition A that the sparsified* $\mathcal{S}^{\mathcal{E}}(\mathbf{Q})$ *contains an* I_K *is necessary for the identifiability of* $(\Gamma(\mathbf{Q}, \mathcal{E}),\ \boldsymbol{\theta}^+, \boldsymbol{\theta}^-,\ \boldsymbol{p})$.

Proposition 3 shows that Condition A cannot be relaxed under any attribute hierarchy. On the other hand, Conditions B and C are more "local" in the sense that they relate to individual attributes (equivalently, individual columns of the **Q**-matrix). Interestingly, the necessity of these two conditions depends greatly on the role of each attribute in the attribute hierarchy graph. We next characterize the fine boundary between the sufficiency and the necessity of the identifiability conditions for various types of attributes. Given any attribute hierarchy graph $\mathcal{E}$, we define the following four types of attributes.

Definition 3 (Singleton Attribute)**.** *An attribute k is a "singleton attribute" if no attribute h exists such that $k \to h$, and no attribute ℓ exists such that $\ell \to k$.*

Definition 4 (Root Attribute)**.** *An attribute k is a "root attribute" if there exists some attribute h such that $k \to h$, but there is no attribute ℓ such that $\ell \to k$.*

Definition 5 (Leaf Attribute)**.** *An attribute k is a "leaf attribute" if there* ***exists*** *some attribute ℓ such that $\ell \to k$, but there is no attribute h such that $k \to h$.*

Definition 6 (Intermediate Attribute)**.** *An attribute k is an "intermediate attribute" if there* ***exists*** *some attribute ℓ with $\ell \to k$ and some attribute h with $k \to h$.*

The above four definitions together describe a full categorization of attributes, given any attribute hierarchy. In other words, given any $\mathcal{E}$, an attribute is either a singleton, a root, a leaf, or an intermediate attribute. As a special case, when the attribute pattern space $\mathcal{A} = \{0, 1\}^K$ is saturated, all K attributes are singleton attributes.

Example 4. Leighton *et al.* (Leighton et al., 2004) were among the first to consider the attribute hierarchy method for the purpose of cognitive diagnosis. In particular, they present and name the four types of hierarchies among $K = 6$ attributes, as shown in our Fig 3. In our terminology, in plot (a), attribute 1 is a root attribute, attribute 6 is a leaf attribute, and the remaining attributes 2, 3, 4, and 5 are intermediate attributes; in plot (b), the roles of the six attributes are the same as those in plot (a); in plot (c), attribute 1 is a root attribute, attributes 2 and 3 are intermediate attributes, and attributes 4, 5, and 6 are leaf attributes; in plot (d), attribute 1 is a root, and the remaining attributes are leaves.

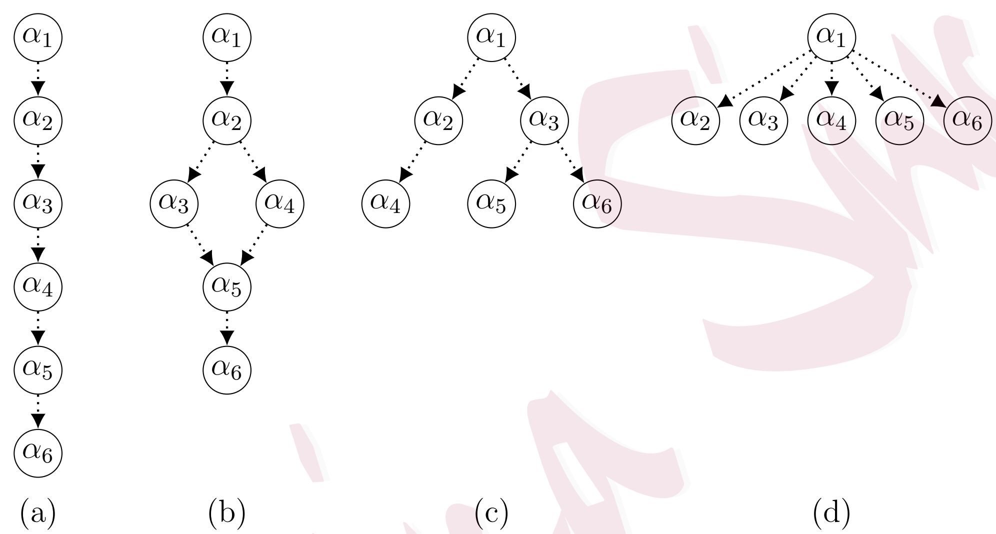


Figure 3: The four attribute hierarchies presented in Leighton et al. (2004): (a) *linear*, (b) *convergent*, (c) *divergent*, and (d) *unstructured*. For example, in (b), α_1 is a root attribute, $\alpha_2, \ldots, \alpha_6$ are intermediate attributes, α_7 is a leaf attribute, and there are no singleton attributes.

For ease of discussion, in the following conclusions on the necessity of the identifiability conditions, we focus on **Q**-matrices that satisfy the restricted **Q**-matrix design, that is, each row vector is a permissible attribute pattern under the hierarchy $\mathcal{E}$. In the literature on cognitive diagnostic modeling, the restricted **Q**-matrix design is shown empirically to improve the clustering accuracy of diagnostic test takers (Tu et al., 2019). Our theoretical findings reveal that, in addition to the restricted **Q**-matrix design, certain other requirements are necessary to ensure identifiability.

Before presenting the next identifiability result, we first introduce a new notion for the identifiability of the attribute hierarchy $\mathcal{E}$ and the proportion parameters $\boldsymbol{p}$, that is, identifiability up to the equivalence classes $[\mathcal{E}]$ and $[\boldsymbol{p}]$, respectively. Under an unknown nonempty hierarchy $\mathcal{E} \neq \varnothing$, if all row vectors of $\mathbf{Q}$ respect the attribute hierarchy, then there exists a trivial nonidentifiability issue that can be resolved by introducing an equivalence relation, similar in spirit to that in Gu and Xu (2020). To see this, consider $K = 2$ and $\mathcal{E} = \{1 \to 2\}$. Then, both rows of a $\mathbf{Q}$-matrix $\mathbf{Q} = \mathbf{E} = (1,0;\ 1,1)$ respect the attribute hierarchy. Further, consider the simplest special case without any item-level noise, $1-\theta_1^+ = 1-\theta_2^+ = \theta_1^- = \theta_2^- = 0$. Now if $\mathcal{E}$ is unknown, then it is not hard to see that any alternative proportion parameters $\bar{\boldsymbol{p}}$ satisfying the following equations are nondistinguishable from the true parameters $\boldsymbol{p}$:

$$p_{(00)} = \bar{p}_{(00)} + \bar{p}_{(01)}; \quad p_{(10)} = \bar{p}_{(10)}; \quad p_{(11)} = \bar{p}_{(11)}. \tag{4.1}$$

The above phenomenon is closely related to the $\boldsymbol{p}$-partial identifiability defined in Gu and Xu (2020), which means that when $\mathbf{Q}$ does not contain an identity submatrix I_K (often called "incomplete" in cognitive diagnosis models), the proportion parameters can at best be identified up to the equivalence classes induced by $\mathbf{Q}$. In the earlier toy example in the last paragraph, the attribute patterns (00) and (01) are equivalent under the incomplete $\mathbf{Q} = (1,0;\ 1,1)$ because $\Gamma_{\mathbf{Q},(00)} = \Gamma_{\mathbf{Q},(01)}$; thus, $\bar{p}_{(00)}$ and $\bar{p}_{(01)}$ can be identified up to their sum, at best, as illustrated in (4.1). Therefore, we say $(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-, [\mathcal{E}], [\boldsymbol{p}])$ are identifiable if the continuous parameters $(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-)$ are identifiable in the usual sense, and $\mathcal{E}$ and $\boldsymbol{p}$ are identifiable up to each of their respective equivalence classes; that is, the only nonidentifiability issue related to $\mathcal{E}$ (and hence $\boldsymbol{p}$) is due to the equivalence relation induced by $\mathbf{Q}$, as in the example in (4.1). The following are formal definitions of $[\mathcal{E}]$ and $[\boldsymbol{p}]$: given an attribute hierarchy $\mathcal{E}$ and a $\mathbf{Q}$-matrix with all row vectors respecting the hierarchy $\mathcal{E}$ (i.e., satisfying the restricted $\mathbf{Q}$-matrix design), define the equivalence class of the attribute hierarchies $[\mathcal{E}]$

and that of the proportion parameters $[\boldsymbol{p}]$ as

$$
\begin{aligned}
&[\mathcal{E}] = \{\bar{\mathcal{E}} : \Gamma(\mathbf{Q}, \mathcal{E}) = \Gamma(\mathbf{Q}, \bar{\mathcal{E}})\} \text{ and} \\
&[\boldsymbol{p}] = \{\bar{\boldsymbol{p}} : \bar{\boldsymbol{p}} \text{ is associated with some } \bar{\mathcal{E}} \in [\mathcal{E}], \text{ with } (\boldsymbol{\theta}^+ = \mathbf{1}_{J\times 1}, \boldsymbol{\theta}^- = \mathbf{0}_{J\times 1}, \mathcal{E}, \boldsymbol{p}) \text{ and} \\
&\qquad\qquad (\boldsymbol{\theta}^+ = \mathbf{1}_{J\times 1}, \boldsymbol{\theta}^- = \mathbf{0}_{J\times 1}, \bar{\mathcal{E}}, \bar{\boldsymbol{p}}) \text{ giving rise to the same distribution of } \mathbf{R}\} \\
&\quad = \{\bar{\boldsymbol{p}} = (\bar{p}_{\boldsymbol{\alpha}}, \boldsymbol{\alpha} \in \{0,1\}^K) : \forall \boldsymbol{\alpha} \text{ that respects the hierarchy } \mathcal{E}, \sum_{\boldsymbol{\alpha}' :\, \Gamma_{\mathbf{Q},\boldsymbol{\alpha}'} = \Gamma_{\mathbf{Q},\boldsymbol{\alpha}}} \bar{p}_{\boldsymbol{\alpha}'} = p_{\boldsymbol{\alpha}}\},
\end{aligned}
$$

respectively. Note that the above nonidentifiability of a specific $\mathcal{E}$ within its equivalence class $[\mathcal{E}]$ is somewhat trivial, and can be easily resolved by simply defining the final $\bar{\mathcal{E}}^\star$ as the hierarchy with the most directed edges among all possible hierarchies in the equivalence class $[\bar{\mathcal{E}}]$. It is easy to see that $\bar{\mathcal{E}}^\star$ is equal to the true $\mathcal{E}$ in the toy example, because in order for $\bar{\mathcal{E}}$ to have the most directed edges, one needs to set $\bar{p}_{(01)} = 0$ under (4.1), which makes the resulting $\bar{\boldsymbol{p}} = \boldsymbol{p}$ and $\bar{\mathcal{E}}^\star = \mathcal{E} = \{1 \to 2\}$. By similar reasoning, this procedure also works more generally for any hierarchy $\mathcal{E}$. Therefore, when a fixed $\mathbf{Q}$-matrix has all rows respecting the hierarchy, it is still meaningful and useful to study the identifiability of $(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-, [\mathcal{E}], [\boldsymbol{p}])$ and to investigate the minimal identifiability conditions. Our results in this section establish the necessary and sufficient identifiability conditions in this regard.

In the following Propositions 4–6, we show how Condition B can be relaxed, in general, depending on whether the attribute is a root, a leaf, or an intermediate attribute.

Proposition 4 (Necessary Condition for Singleton Attribute)**.** *Consider a DINA-based HLAM with a fixed* $\mathbf{Q}$*-matrix whose row vectors respect the hierarchy* $\mathcal{E}$*. The following hold for a singleton attribute* k *in any attribute hierarchy:*

(a) $\sum_{j=1}^J q_{j,k} \geq 3$ *is necessary for the identifiability of* $(\mathcal{E}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$*.*

(b) There exist scenarios in which the equality in part (a) is achieved with $\sum_{j=1}^J q_{j,k} = 3$

and the identifiability of $(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-, [\mathcal{E}], [\boldsymbol{p}])$ *is guaranteed.*

Proposition 5 (Necessary Condition for Root and Leaf Attributes)**.** *Consider a DINA-based HLAM with a fixed* $\mathbf{Q}$*-matrix whose row vectors respect the hierarchy* $\mathcal{E}$*. Denote the* (j,k)*th entry of* $\mathcal{S}^{\mathcal{E}}(\mathbf{Q})$ *by* $q_{j,k}^{\text{sparse}}$*. The following conclusions hold for* k *if the attribute* k *is either a root or a leaf attribute:*

(a) $\sum_{j=1}^{J} q_{j,k}^{\text{sparse}} \geq 2$ *is necessary for the identifiability of* $(\mathcal{E}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$.

(b) There exist scenarios in which the equality in part (a) is achieved with $\sum_{j=1}^{J} q_{j,k}^{\text{sparse}} = 2$ *and the identifiability of* $(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-, [\mathcal{E}], [\boldsymbol{p}])$ *is guaranteed.*

Proposition 6 (Necessary Condition for Intermediate Attribute)**.** *Consider a DINA-based HLAM with a fixed* $\mathbf{Q}$*-matrix whose row vectors respect the hierarchy* $\mathcal{E}$*. Denote the* (j,k)*th entry of* $\mathcal{S}^{\mathcal{E}}(\mathbf{Q})$ *by* $q_{j,k}^{\text{sparse}}$*. The following statements hold for an intermediate attribute* k*:*

(a) $\sum_{j=1}^{J} q_{j,k}^{\text{sparse}} \geq 1$ *is necessary for the identifiability of* $(\mathcal{E}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$.

(b) There exist scenarios in which the equality in part (a) is achieved with $\sum_{j=1}^{J} q_{j,k}^{\text{sparse}} = 1$ *and the identifiability of* $(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-, [\mathcal{E}], [\boldsymbol{p}])$ *is guaranteed.*

Propositions 4–6 together characterize the identifiability phenomena caused by different types of attributes in the attribute hierarchy graph. An intuitive explanation for these conclusions is as follows. For a singleton attribute k that is not connected to any other attribute in the attribute hierarchy graph, no additional information is provided by the other attributes. Therefore, the requirement of k being measured by at least three items in the $\mathbf{Q}$-matrix is necessary. This aligns well with the conclusion established in Xu and Zhang (2016) and Gu and Xu (2019b) for a latent attribute model without any hierarchy, where all the attributes are singletons and each needs to be measured by at least three items. However, this requirement can be relaxed for any other type of attribute which is somewhat connected in the attribute hierarchy graph. In particular, fewer measurements are needed for k in the $\mathbf{Q}$-matrix, because more information is available for this attribute in the attribute hierarchy graph. For a root attribute k with some "child" or a leaf attribute with some "parent" as one-sided information, the requirement is relaxed to k being measured by at least two items in $\mathcal{S}^{\mathcal{E}}(\mathbf{Q})$. For an intermediate attribute k with some child and some parent as two-sided information, the requirement is further relaxed to k being measured by at least one item in $\mathcal{S}^{\mathcal{E}}(\mathbf{Q})$.

We next discuss the necessity of Condition C. Given $\mathbf{Q}$, we denote by $\mathbf{Q}_{1:K,:}$ the submatrix consisting of its first K rows, and by $\mathbf{Q}_{(K+1):J,:}$ the submatrix consisting of its last $J-K$ rows. For $\mathbf{Q}$ with rows respecting the attribute hierarchy, Condition C requires $\mathbf{Q}_{(K+1):J,k} \neq \mathbf{Q}_{(K+1):J,\ell}$, for any $k \neq \ell$, when $\mathbf{Q}_{1:K,:} = \mathbf{E}$. We have the following result.

Proposition 7 (Necessity of Condition C)**.** *Consider a DINA-based HLAM with a fixed* $\mathbf{Q}$ *whose row vectors respect the hierarchy* $\mathcal{E}$*. The condition that* $\mathbf{Q}_{(K+1):J,k} \neq \mathbf{Q}_{(K+1):J,\ell}$ *(when* $\mathbf{Q}_{1:K,:} = \mathbf{E}$*) is* ***necessary*** *for identifiability if both* α_k *and* α_ℓ *are singleton attributes.*

4.2 Bridging the Necessity and Sufficiency of the Identifiability Conditions

Still under a fixed and known $\mathbf{Q}$-matrix, as in Section 4.1, we next investigate how the sufficient identifiability conditions for $(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$ can meet the necessary identifiability conditions proposed earlier in Propositions 5–7. In the next theorem, we establish that the combination of the individual necessary conditions established in Section 4.1 are actually sufficient to guarantee the identifiability in fully general scenarios. This result therefore establishes the general necessary and sufficient condition on the $\mathbf{Q}$-matrix for identifiability under an arbitrary attribute structure.

Theorem 2 (Necessary and Sufficient Conditions under a Fully General $\mathcal{E}$)**.** *Consider a DINA-based HLAM with a fixed* $\mathbf{Q}$*-matrix whose row vectors respect the hierarchy* $\mathcal{E}$*. Then, Condition A and the following condition* $B^\star$ *and* $C^\star$ *are necessary and sufficient for the identifiability of* $(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-, [\mathcal{E}], [\boldsymbol{p}])$*.*

- $B^\star$*. In* $\mathcal{S}^{\mathcal{E}}(\mathbf{Q})$*, any intermediate attribute is measured by at least one item, any root and leaf attributes are measured by at least two items, and any singleton attribute is measured by at least three items.*

- $C^\star$*. For any two singleton attributes* α_k *and* α_ℓ*,* $\mathbf{Q}_{(K+1):J,k} \neq \mathbf{Q}_{(K+1):J,\ell}$*. (Assume* $\mathbf{Q}_{1:K,:} = \mathbf{E}$ *under Condition A.)*

Theorem 2 covers any type of attribute structure and allows for any type of attributes in the attribute hierarchy graph. In the special case in which there are no singleton attributes in the graph, the necessary and sufficient identifiability conditions in Theorem 2 can be simplified. We refer to a family of hierarchies without any singleton attributes as a *connected-graph hierarchy.*

Corollary 1 (Necessary and Sufficient Condition under a Connected-Graph Hierarchy)**.** *Consider a DINA-based HLAM with a fixed* $\mathbf{Q}$*-matrix whose row vectors respect the hierarchy* $\mathcal{E}$*. Suppose the* K *attributes form a connected graph. Then, Condition A and the following Condition D are necessary and sufficient for the identifiability of* $(\mathcal{E}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, [\boldsymbol{p}])$*.*

- *D. In* $\mathcal{S}^{\mathcal{E}}(\mathbf{Q})$*, any root and leaf attributes are measured by at least two items, and any intermediate attribute is each measured by at least one item.*

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

Remark 4. In the first extreme case, if $\mathcal{E} = \varnothing$, without any true hierarchy among the attributes, then Conditions A, B*, and C* in Theorem 2 become Conditions A, B, and C in Theorem 1 in Section 3. In the second extreme case, if there are no singleton attributes in the attribute hierarchy graph, then Condition B* in Theorem 2 reduces to Condition D in Corollary 1; here Condition C* in Theorem 2 should be understood as being always satisfied, and hence can be omitted. That is, under a connected-graph hierarchy without any singleton attributes, the Conditions A, B*, and C* in Theorem 2 reduce to Conditions A and D in Corollary 1. Therefore, Theorem 2 covers Corollary 1 as a special case, and is indeed fully general. We state these two results separately to highlight the most general form of the result, and to show how the necessary and sufficient conditions simplify under the popular family of connected-graph hierarchies, as in Corollary 1.

The following example illustrates the minimal requirements on $\mathbf{Q}$ under the attribute hierarchies considered in Leighton et al. (2004).

Example 5. Under the linear hierarchy $\mathcal{E} = \mathcal{E}^{\text{linear}}$ in Fig 4(b), the 8×6 matrix $\mathbf{Q}^{\text{linear}}_{8\times 6}$ shown in Fig 4(a) encodes the minimal requirement for identifiability. Fig 4(b) visualizes the sparsified version of $\mathbf{Q}^{\text{linear}}_{8\times 6}$ as directed solid edges from the latent attributes to the observed item responses. Under the so-called convergent hierarchy and divergent hierarchy presented in Fig 3, the minimal requirement on $\mathbf{Q}$ for model identifiability is presented in parts (c)–(d) and parts (e)–(f) of Figure 4, repectively. For the divergent hierarchy $\mathcal{E} = \mathcal{E}^{\text{div}}$ in Fig 4(f), $\mathbf{Q}^{\text{div}}_{10\times 6}$ in Fig 4(c) gives an identifiable model under minimal conditions.

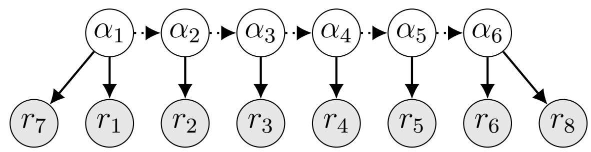
5. Identifiability of HLAMs other than DINA-based HLAMs

We also study the identifiability of HLAMs other than DINA-based HLAMs.

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

$$\mathbf{Q}^{\text{linear}}_{8\times 6} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

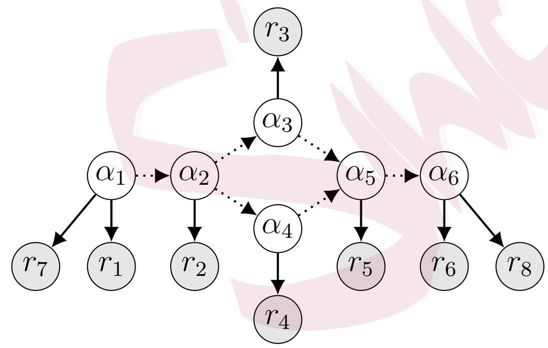
(a) $\mathbf{Q}^{\text{linear}}_{8\times 6}$



(b) visualization of the sparsified $\mathcal{S}^{\mathcal{E}}(\mathbf{Q}^{\text{linear}}_{8\times 6})$

$$\mathbf{Q}^{\text{conv}}_{8\times 6} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

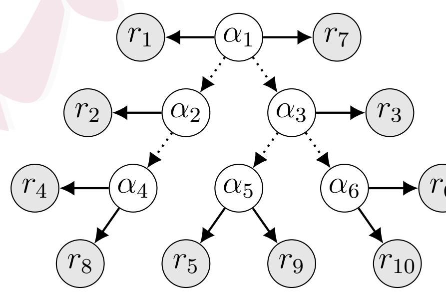
(c) $\mathbf{Q}^{\text{conv}}_{8\times 6}$



(d) visualization of the sparsified $\mathcal{S}^{\mathcal{E}}(\mathbf{Q}^{\text{conv}}_{8\times 6})$

$$\mathbf{Q}^{\text{div}}_{10\times 6} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

(e) $\mathbf{Q}^{\text{div}}_{10\times 6}$



(f) visualization of the sparsified $\mathcal{S}^{\mathcal{E}}(\mathbf{Q}^{\text{div}}_{10\times 6})$

Figure 4: Minimally sufficient requirements on $\mathbf{Q}$ for identifiability under the *linear* hierarchy, *convergent* hierarchy, and *divergent* hierarchy proposed in Leighton et al. (2004), respectively.

5.1 DINO-based HLAMs

As introduced in Section 2, the DINO model is a popular type of latent attribute model, often used for the psychiatric and clinical measurement of mental disorders (Templin and

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

Henson, 2006; de la Torre et al., 2018). A careful examination of the definitions of the ideal responses Γ^{AND} and Γ^{OR} in (2.1) and (2.5) reveals the relationship $\Gamma^{\text{OR}}_{\boldsymbol{q}_j,\boldsymbol{\alpha}} = 1-\Gamma^{\text{AND}}_{\boldsymbol{q}_j,\mathbf{1}_K-\boldsymbol{\alpha}}$, where $\mathbf{1}_K - \boldsymbol{\alpha} = (1-\alpha_1,\dots,1-\alpha_K)^\top$ also denotes an attribute pattern. Building on the duality between DINA and DINO, the following proposition characterizes how the identifiability results obtained under a DINA-based HLAM can be translated to the case of a DINO-based HLAM.

Proposition 8. *Consider a DINO-based HLAM with a fixed* $\mathbf{Q}$*-matrix and an unknown attribute hierarchy* $\mathcal{E}$*. Define the reversed attribute hierarchy* $\mathcal{E}^{\text{reverse}}$ *as*

$$\mathcal{E}^{\text{reverse}} = \{\ell \to k : \textit{ if } k \to \ell \textit{ under the original hierarchy } \mathcal{E}\}. \tag{5.1}$$

(a) For any $\boldsymbol{\alpha} \in \{0,1\}^K$, $\boldsymbol{\alpha} \in \mathcal{A}(\mathcal{E})$ *if and only if* $\mathbf{1}_K - \boldsymbol{\alpha} \in \mathcal{A}(\mathcal{E}^{\text{reverse}})$*. That is, any attribute pattern* $\boldsymbol{\alpha}$ *is allowable under the original hierarchy* $\mathcal{E}$ *if and only if another attribute pattern* $\boldsymbol{\alpha}' = \mathbf{1} - \boldsymbol{\alpha}$ *is allowable under the reversed hierarchy* $\mathcal{E}^{\text{reverse}}$*.*

(b) The attribute hierarchy $\mathcal{E}$ *and model parameters under a DINO-based HLAM are identifiable* ***if and only if*** *the reversed attribute hierarchy* $\mathcal{E}^{\text{reverse}}$ *and model parameters are identifiable under a DINA-based HLAM with the same* $\mathbf{Q}$*-matrix.*

For any attribute hierarchy graph $\mathcal{E}$, the reversed hierarchy $\mathcal{E}^{\text{reverse}}$ in (5.1) is another directed graph among the attributes, where the direction of each arrow in $\mathcal{E}$ is reversed. Therefore, for the same set of K attributes, any root attribute in $\mathcal{E}$ becomes a leaf attribute in $\mathcal{E}^{\text{reverse}}$, and any leaf in $\mathcal{E}$ becomes a root in $\mathcal{E}^{\text{reverse}}$. Any intermediate or singleton attributes remain the same type when $\mathcal{E}$ is reversed to $\mathcal{E}^{\text{reverse}}$. Proposition 8 provides guidelines on how to check the identifiability for a DINO-based HLAM using the identifiability results established earlier for DINA-based HLAMs. In particular, we have the following necessary and sufficient conditions for the identifiability of $(\mathcal{E}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$ under a DINO-based HLAM with a fixed $\mathbf{Q}$-matrix.

Corollary 2 (Necessary and Sufficient Conditions under a General $\mathcal{E}$ for a DINO-based HLAM)**.** *Consider a DINO-based HLAM with an attribute hierarchy $\mathcal{E}$ and a fixed $\mathbf{Q}$-matrix whose rows respect the reversed hierarchy $\mathcal{E}^{\text{reverse}}$. Consider the following condition.*

$A^\star$. *The $\mathcal{E}^{\text{reverse}}$-densified matrix $\mathcal{D}^{\mathcal{E}^{\text{reverse}}}(\mathbf{Q})$ contains a submatrix that is the reachability matrix under the reversed hierarchy $\mathcal{E}^{\text{reverse}}$.*

Then, Condition $A^\star$ and Conditions $B^\star$–$C^\star$ in Theorem 2 are necessary and sufficient for the identifiability of $(\mathcal{E}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$.

5.2 Main-effect-based HLAMs

Another family of HLAMs in the literature (e.g., DiBello et al., 1995; von Davier, 2008; Henson et al., 2009) incorporate the main effects of the latent attributes into the model. We review these main-effect-based HLAMs in Example 6, and then provide the identifiability result for them.

Example 6 (HLAMs that Model the Main Effects of Attributes)**.** *Main-effect HLAMs* assume that the main effects of the attributes measured by each item indicated by $\boldsymbol{q}_j$ play a role in distinguishing the item parameters. Under a main-effect HLAM, the Bernoulli parameter $\theta_{j,\boldsymbol{\alpha}}$ can be written as

$$\theta^{\text{main-eff}}_{j,\boldsymbol{\alpha}} = f\Big(\beta_{j,0} + \sum\nolimits_{k=1}^{K} \beta_{j,k} q_{j,k} \alpha_k\Big), \tag{5.2}$$

where $f(\cdot)$ is a link function. Note not we do not need all the β-coefficients in the above display in the model specification; instead, only when $q_{j,k} = 1$ will $\beta_{j,k}$ be needed and truly incorporated in the model. Different link functions $f(\cdot)$ in (5.2) lead to different models, including the linear logistic model (LLM; Maris, 1999), with $f(\cdot)$ being the sigmoid function, and the additive cognitive diagnosis model (ACDM; de la Torre, 2011), with $f(\cdot)$ being the identity. When $f(\cdot)$ is a monotonically increasing function, it is usually assumed in practice that each $\beta_{j,k} > 0$ wherever $q_{j,k} = 1$, for interpretability.

All-effect HLAMs model the main effects and the interaction effects of the attributes. The Bernoulli parameter $\theta_{j,\boldsymbol{\alpha}}$ of an all-effect model is

$$\theta^{\text{all-eff}}_{j,\boldsymbol{\alpha}} = f\Big(\beta_{j,0} + \sum\nolimits_{k=1}^{K}\beta_{j,k}(q_{j,k}\alpha_k) + \sum\nolimits_{1\le k_1<k_2\le K}\beta_{j,k_1k_2}(q_{j,k_1}\alpha_{k_1})(q_{j,k_2}\alpha_{k_2}) + \cdots + \beta_{j,12\cdots K}\prod\nolimits_{k=1}^{K}(q_{j,k}\alpha_k)\Big). \tag{5.3}$$

Similarly to (5.2), we do not need all the β-coefficients in the model specification. When $f(\cdot)$ in (5.3) is the identity function, (5.3) gives the generalized DINA (GDINA) model in de la Torre (2011); and when $f(\cdot)$ is the sigmoid function, (5.3) gives the log-linear cognitive diagnosis models (LCDMs) in Henson et al. (2009); see also the general diagnostic models (GDMs) in von Davier (2008). In general, we call the main-effect HLAMs in (5.2) and the all-effect HLAMs in (5.3) *main-effect-based HLAMs*, because they both incorporate the main effects of the latent attributes in the model.

Under main-effect-based HLAMs, the probability mass function of the J-dimensional random response vector $\mathbf{R}$ can be written generally as

$$P(\mathbf{R} = \boldsymbol{r} \mid \mathbf{Q}, \mathcal{E}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p}) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}(\mathcal{E})} p_{\boldsymbol{\alpha}} \prod_{j=1}^{J} \theta_{j,\boldsymbol{\alpha}}^{r_j} \times (1-\theta_{j,\boldsymbol{\alpha}})^{1-r_j},$$

where $\boldsymbol{r} \in \{0,1\}^J$ is an arbitrary response pattern. Notably, these main-effect-based HLAMs have quite different algebraic structures to those two-parameter HLAMs, namely the DINA and DINO models. The key structure of any two-parameter HLAM is captured by the ideal response $\Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}}$ in (2.1) or (2.4), under the "AND" or "OR" operations, respectively. Intuitively, two-parameter HLAMs are characterized by a probabilistic version of the *Boolean product* of two groups of binary vectors, the group of $\boldsymbol{q}_j$ and the group of $\boldsymbol{\alpha}$; however, this is not the case for any HLAM in Example 6, owing to the incorporation of the main effects of the attributes. Indeed, incorporating main effects in the form of $\sum_{k=1}^{K}\beta_{j,k}q_{j,k}\alpha_k$ in (5.2) or (5.3) is taking an *inner product* of the vectors $\boldsymbol{q}_j$ and $\boldsymbol{\alpha}$ and an additional β-coefficient vector, rather than the Boolean product. As such, the necessary and sufficient identifiability conditions derived for two-parameter HLAMs in Sections 3-4 do not apply to main-effect-based HLAMs.

Next, we give a set of sufficient conditions for the identifiability of main-effect-based HLAMs. The technical concept of $\Gamma(\mathbf{Q},\mathcal{E})$ (specifically, with $\Gamma = \Gamma^{\text{AND}}$ defined in (2.1)) introduced in Section 3 is still useful here. Denote the collection of all per-item Bernoulli parameters by $\boldsymbol{\Theta} = (\theta_{j,\boldsymbol{\alpha}})$. We have the following theorem.

Theorem 3 (Identifiability of HLAMs that Model the Main Effects of Attributes)**.** *Consider an HLAM that incorporates the main effects of the attributes, with* $\mathbf{Q}$ *and* $\mathcal{E}$ *both unknown. Suppose* $\boldsymbol{\Theta}$ *satisfies a natural inequality constraint* $\theta_{j,\boldsymbol{\alpha}} \neq \theta_{j,\boldsymbol{\alpha}'}$ *if* $\Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}} \neq \Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}'}$. *If* $\Gamma(\mathbf{Q},\mathcal{E})$ *satisfies the following conditions with the number of columns known, then* $(\boldsymbol{\Theta},\boldsymbol{p})$ *and* $\Gamma(\mathbf{Q},\mathcal{E})$ *are identifiable.*

E. There exist two disjoint sets of items S_1, $S_2 \subseteq [J]$, *such that* $\Gamma(\mathbf{Q}_{S_1,:}\,,\mathcal{E})$ *and* $\Gamma(\mathbf{Q}_{S_2,:}\,,\mathcal{E})$ *each has distinct column vectors.*

F. For any $\boldsymbol{\alpha} \neq \boldsymbol{\alpha}' \in \mathcal{A}(\mathcal{E})$, *there exists some* $j \notin S_1 \cup S_2$ *such that* $\Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}} \neq \Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}'}$.

G. For any $\boldsymbol{\alpha} \in \mathcal{A}(\mathcal{E})$, $\boldsymbol{\alpha}' \in \{0,1\}\backslash\mathcal{A}(\mathcal{E})$, *there exists some* $j \in [J]$ *such that* $\Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}} \neq \Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}'}$.

In addition to the above three conditions, if $\mathbf{Q}$ *is known in part to contain an identity submatrix* I_K*, then the attribute hierarchy* $\mathcal{E}$ *is identifiable from* $\Gamma(\mathbf{Q}, \mathcal{E})$.

For main-effect-based HLAMs, the ideal response matrix $\Gamma(\mathbf{Q}, \mathcal{E})$ may not sharply characterize the entire latent structure, owing to the incorporation of the main effects, on contrast to the DINA-based HLAMs. To see this, consider two latent patterns $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}'$, with $\Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}} = \Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}'} = 0$. Then, the specification in (5.2) or (5.3) implies that it is possible that $\theta_{j,\boldsymbol{\alpha}} \neq \theta_{j,\boldsymbol{\alpha}'}$. Therefore it is difficult, if at all possible, to explicitly characterize the necessary identifiability conditions in terms of $\Gamma(\mathbf{Q}, \mathcal{E})$ for main-effect-based HLAMs. However, $\Gamma(\mathbf{Q}, \mathcal{E})$ is still useful for deriving the sufficient conditions for identifiability, as in Theorem 3. This is because if $\Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}} = \Gamma_{\boldsymbol{q}_j,\boldsymbol{\alpha}'} = 1$, the two attribute patterns $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}'$ both satisfy $\boldsymbol{\alpha} \succeq \boldsymbol{q}_j$ and $\boldsymbol{\alpha}' \succeq \boldsymbol{q}_j$, by the definition in (2.1). This implies that the patterns $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}'$ both possess all the attributes measured by the vector $\boldsymbol{q}_j$. As a result, the definition of main-effect-based models in (5.2) or in (5.3) shows that there must be $\theta_{j,\boldsymbol{\alpha}} = \theta_{j,\boldsymbol{\alpha}'}$ for these two patterns. This intuitively explains why $\Gamma(\mathbf{Q}, \mathcal{E})$ can be used to describe a set of sufficient identifiability conditions for main-effect-based HLAMs.

We make a remark on the relationship between the main-effect-based HLAMs and the DINA-based HLAMs studied in Sections 3–4. On the one hand, the main-effect-based HLAMs are more general than DINA-based HLAMs in the sense that the formulation of $\theta^{\text{main-eff}}_{j,\boldsymbol{\alpha}}$ in (5.2) or $\theta^{\text{all-eff}}_{j,\boldsymbol{\alpha}}$ in (5.3) can generally allow for more than two Bernoulli parameters for each j, whereas DINA-based HLAMs always have two parameters θ_j^+ and θ_j^- for each j. However, we focus on DINA-based two-parameter HLAMs, which are widely used in the motivating applications of cognitive diagnosis in educational settings. Indeed, these are the settings in which attribute hierarchies receive the most attention in terms of modeling the

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0350

sequential acquisition of skill attributes (e.g., Leighton et al., 2004; Gierl et al., 2007; Wang and Lu, 2020). On the practical side, assuming a conjunctive relationship among the attributes, as in DINA, is often believed to be suitable for modeling the response mechanism of diagnostic test items in such settings (e.g., Junker and Sijtsma, 2001; de la Torre and Douglas, 2004). On the theoretical side, the identifiability of two-parameter DINA-based HLAMs is more intriguing because of the Boolean product involved. The rich combinatorial nature of such models gives the opportunity to close the gap between the necessity and sufficiency of identifiability requirements; interestingly, these minimal requirements are explicit conditions on the discrete structure: the $\mathbf{Q}$-matrix and attribute types, as depicted in Section 4. Therefore, we believe that closely examining DINA-based two-parameter HLAMs and establishing the minimal identifiability conditions for them (as done in Sections 3–4) are highly desirable, for both theoretical interest and practical relevance.

6. Conclusion

We have provided a first study on the identifiability of the hierarchical latent attribute model, a complex-structured latent variable model popular for modeling modern assessment data. We propose sufficient identifiability conditions that explicitly depend on the attribute hierarchy graph and the structural $\mathbf{Q}$-matrix. We also discuss the necessity of the identifiability conditions and sharply characterize the effects on identifiability of different types of attributes in the attribute hierarchy graph. We focus mainly on the basic and popular HLAMs, namely the DINA-based HLAMs, where each item is modeled using two parameters. We also extend the theory to other types of HLAMs in Section 5.

One nice implication of identifiability is the estimability of both the latent structure and the parameters that define the probabilistic model. When the proposed conditions are satisfied, all the components of an HLAM can be uniquely and consistently estimated from data based on the maximum likelihood. In a practical data analysis under the HLAM framework, if $\mathbf{Q}$ and $\mathcal{E}$ are specified by domain experts or applied researchers, then before seeing any data, one can check whether $\mathbf{Q}$ and $\mathcal{E}$ satisfy our proposed conditions for assessing model identifiability. On the other hand, if $\mathbf{Q}$ and $\mathcal{E}$ are not known and one hopes to estimate them exploratorily from data, our identifiability results can also be useful. In such scenarios, one can check whether the estimated $\widehat{\mathbf{Q}}$ and $\widehat{\mathcal{E}}$ satisfy necessary identifiability conditions; if not, then a more careful investigation of the diagnostic test design may be needed. Therefore, this study provides useful insights into designing valid diagnostic tests and drawing valid scientific conclusions from assessment data under a potentially complicated attribute hierarchy.

Supplementary Materials. The supplementary material contains proofs of the theorems and some illustrative examples.

Acknowledgments. This research was partially supported by NSF CAREER SES-1846747, DMS-1712717, SES-1659328; and NIH R01ES027498, R01ES028804. The authors are grateful to the editor, associate editor, and two reviewers for their constructive comments.

References

Allman, E. S., Matias, C., and Rhodes, J. A. (2009). Identifiability of parameters in latent structure models with many observed variables. *The Annals of Statistics*, 37:3099–3132.

Anandkumar, A., Ge, R., Hsu, D., Kakade, S. M., and Telgarsky, M. (2014). Tensor decompositions for learning latent variable models. *Journal of Machine Learning Research*, 15(1):2773–2832.

Anderson, T. W. (2009). *An introduction to multivariate statistical analysis*. John Wiley & Sons, New York.

Bai, J. and Li, K. (2012). Statistical analysis of factor models of high dimension. *The Annals of Statistics*, 40(1):436–465.

Cai, Y., Tu, D., and Ding, S. (2018). Theorems and methods of a complete Q matrix with attribute hierarchies under restricted Q-matrix design. *Frontiers in Psychology*, 9:1413.

Chen, Y., Culpepper, S., and Liang, F. (2020). A sparse latent class model for cognitive diagnosis. *Psychometrika*, pages 1–33.

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

Choi, M. J., Tan, V. Y., Anandkumar, A., and Willsky, A. S. (2011). Learning latent tree graphical models. *Journal of Machine Learning Research*, 12(May):1771–1812.

de la Torre, J. (2011). The generalized DINA model framework. *Psychometrika*, 76:179–199.

de la Torre, J. and Douglas, J. A. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika*, 69(3):333–353.

de la Torre, J., van der Ark, L. A., and Rossi, G. (2018). Analysis of clinical data from a cognitive diagnosis modeling framework. *Measurement and Evaluation in Counseling and Development*, 51(4):281–296.

DiBello, L. V., Stout, W. F., and Roussos, L. A. (1995). Unified cognitive/psychometric diagnostic assessment likelihood-based classification techniques. *Cognitively diagnostic assessment*, pages 361–389.

Fang, G., Liu, J., and Ying, Z. (2019). On the identifiability of diagnostic classification models. *Psychometrika*, 84(1):19–40.

Fraley, C. and Raftery, A. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association*, 97(458):611–631.

Ghahramani, Z. and Griffiths, T. (2006). Infinite latent feature models and the Indian buffet process. In *Advances in Neural Information Processing Systems*, pages 475–482.

Gierl, M. J., Leighton, J. P., and Hunka, S. M. (2007). Using the attribute hierarchy method to make diagnostic inferences about respondents' cognitive skills. *Cognitive diagnostic assessment for education: Theory and applications, Cambridge, UK: Cambridge University Press*, pages 242 – 274.

Gu, Y. and Xu, G. (2019a). Learning attribute patterns in high-dimensional structured latent attribute models. *Journal of Machine Learning Research*, 20(115):1–58.

Gu, Y. and Xu, G. (2019b). The sufficient and necessary condition for the identifiability and estimability of the DINA model. *Psychometrika*, 84(2):468–483.

Gu, Y. and Xu, G. (2020). Partial identifiability of restricted latent class models. *Annals of Statistics*, 48(4):2082–2107.

Gu, Y. and Xu, G. (2021). Sufficient and necessary conditions for the identifiability of the Q-matrix. *Statistica Sinica*, 31:449–472.

Hamada, M. and Wu, C. J. (1992). Analysis of designed experiments with complex aliasing. *Journal of Quality Technology*, 24(3):130–137.

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

Hernán, M. A. and Robins, J. M. (2006). Instruments for causal inference: an epidemiologist's dream? *Epidemiology*, pages 360–372.

Hinton, G. E. (2002). Training products of experts by minimizing contrastive divergence. *Neural computation*, 14(8):1771–1800.

Junker, B. W. and Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement*, 25(3):258–272.

Köhn, H.-F. and Chiu, C.-Y. (2019). Attribute hierarchy models in cognitive diagnosis: identifiability of the latent attribute space and conditions for completeness of the Q-matrix. *Journal of Classification*, 36(3):541–565.

Leighton, J. P., Gierl, M. J., and Hunka, S. M. (2004). The attribute hierarchy method for cognitive assessment: a variation on Tatsuoka's rule-space approach. *Journal of Educational Measurement*, 41(3):205–237.

Maris, E. (1999). Estimating multiple classification latent class models. *Psychometrika*, 64(2):187–212.

McLachlan, G. and Peel, D. (2004). *Finite mixture models*. John Wiley & Sons.

O'Brien, K. L., Baggett, H. C., Brooks, W. A., et al. (2019). Causes of severe pneumonia requiring hospital admission in children without hiv infection from africa and asia: the perch multi-country case-control study. *The Lancet*.

Pearl, J. (1986). Fusion, propagation, and structuring in belief networks. *Artificial intelligence*, 29(3):241–288.

Ravanbakhsh, S., Póczos, B., and Greiner, R. (2016). Boolean matrix factorization and noisy completion via message passing. In *Proceedings of the 33rd International Conference on Machine Learning-Volume 48*, pages 945–954.

Rukat, T., Holmes, C. C., Titsias, M. K., and Yau, C. (2017). Bayesian Boolean matrix factorisation. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 2969–2978. JMLR. org.

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

Swanson, S. A., Miller, M., Robins, J. M., and Hernán, M. A. (2015). Definition and evaluation of the monotonicity condition for preference-based instruments. *Epidemiology (Cambridge, Mass.)*, 26(3):414.

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

Tatsuoka, M. M. (1986). Graph theory and its applications in educational research: A review and integration. *Review of Educational Research*, 56(3):291–329.

Templin, J. and Bradshaw, L. (2014). Hierarchical diagnostic classification models: A family of models for estimating and testing attribute hierarchies. *Psychometrika*, 79(2):317–339.

Templin, J. L. and Henson, R. A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological methods*, 11(3):287.

Tu, D., Wang, S., Cai, Y., Douglas, J., and Chang, H.-H. (2019). Cognitive diagnostic models with attribute hierarchies: Model estimation with a restricted q-matrix design. *Applied Psychological Measurement*, 43(4):255–271.

von Davier, M. (2008). A general diagnostic model applied to language testing data. *British Journal of Mathematical and Statistical Psychology*, 61:287–307.

von Davier, M. (2014). The DINA model as a constrained general diagnostic model: Two variants of a model equivalency. *British Journal of Mathematical and Statistical Psychology*, 67(1):49–71.

Wang, C. and Lu, J. (2020). Learning attribute hierarchies from data: two exploratory approaches. *Journal of Educational and Behavioral Statistics*.

Wang, S., Yang, Y., Culpepper, S. A., and Douglas, J. A. (2018). Tracking skill acquisition with cognitive diagnosis models: a higher-order, hidden markov model with covariates. *Journal of Educational and Behavioral Statistics*, 43(1):57–87.

Wu, Z., Deloria-Knoll, M., Hammitt, L. L., Zeger, S. L., and for Child Health Core Team, P. E. R. (2016). Partially latent class models for case–control studies of childhood pneumonia aetiology. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 65(1):97–114.

Wu, Z., Deloria-Knoll, M., and Zeger, S. L. (2017). Nested partially latent class models for dependent binary data; estimating disease etiology. *Biostatistics*, 18(2):200–213.

Xu, G. (2017). Identifiability of restricted latent class models with binary responses. *The Annals of Statistics*, 45:675–707.

Xu, G. and Shang, Z. (2018). Identifying latent structures in restricted latent class models. *Journal of the American Statistical Association*, 113(523):1284–1295.

Xu, G. and Zhang, S. (2016). Identifiability of diagnostic classification models. *Psychometrika*, 81:625–649.

Yuan, M., Joseph, V. R., and Zou, H. (2009). Structured variable selection and estimation. *The Annals of Applied Statistics*, pages 1738–1757.