PSYCHOMETRIKA—VOL. 89, NO. 2, 717–740
JUNE 2024
https://doi.org/10.1007/s11336-024-09961-w

SUFFICIENT AND NECESSARY CONDITIONS FOR THE IDENTIFIABILITY OF DINA MODELS WITH POLYTOMOUS RESPONSES

MENGQI LIN AND GONGJUN XU
UNIVERSITY OF MICHIGAN

Cognitive diagnosis models (CDMs) provide a powerful statistical and psychometric tool for researchers and practitioners to learn fine-grained diagnostic information about respondents' latent attributes. There has been a growing interest in the use of CDMs for polytomous response data, as more and more items with multiple response options become widely used. Similar to many latent variable models, the identifiability of CDMs is critical for accurate parameter estimation and valid statistical inference. However, the existing identifiability results are primarily focused on binary response models and have not adequately addressed the identifiability of CDMs with polytomous responses. This paper addresses this gap by presenting sufficient and necessary conditions for the identifiability of the widely used DINA model with polytomous responses, with the aim to provide a comprehensive understanding of the identifiability of CDMs with polytomous responses and to inform future research in this field.

Key words: identifiability, polytomous responses, **Q**-matrix, cognitive diagnosis models, DINA model.

Cognitive diagnosis models (CDMs), which serve as a powerful tool to infer subjects' latent attributes such as skills, knowledge, or psychological disorders based on their responses to some designed diagnostic items in the cognitive diagnosis assessment, have drawn increasing attention over the years. As a family of discrete latent variable models, its popularity is not limited to educational assessments (Junker & 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 & Henson, 2006; de la Torre et al., 2018), and epidemiological and medical measurement studies (Wu et al., 2017; O'Brien et al., 2019).

Various CDMs have been developed with different diagnostic assumptions and modeling goals, among which the Deterministic Input Noisy output "And" gate model (DINA; Junker and Sijtsma, 2001), which assumes that subjects are expected to complete an item correctly only when they possess all required attributes, is one of the most popular ones. Furthermore, the DINA model also serves as a basis for a larger range of more general CDMs, including the general diagnostic model (von Davier, 2008), the log linear CDM (LCDM; Henson et al., 2009), and the generalized DINA model (G-DINA; de la Torre, 2011). As tests with polytomous responses appear more frequently in practice, the study of CDMs with polytomous responses has also grown in popularity (Culpepper & Balamuta, 2021). Specifically, several models concerning polytomous responses were proposed, such as general diagnostic models (GDM; von Davier, 2008), general polytomous diagnosis models (GPDM; Chen and de la Torre, 2018), and sequential cognitive diagnosis models (Sequential CDM; Ma and de la Torre, 2016).

As is the case with many statistical methods, ensuring the models applied in the cognitive diagnosis are statistically *identifiable* is fundamental to achieve reliable and valid diagnostic assessment. Additionally, this is also a necessity for consistent estimation of the model parameters of interest and valid statistical inferences. The study of identifiability issue for CDMs has long

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11336-024-09961-w.

Correspondence should be made to Gongjun Xu, Department of Statistics, University of Michigan, 456 West Hall, 1085 South University, Ann Arbor 48109, MI, USA. Email: gongjun@umich.edu

been considered, such as DiBello et al. (1995), Maris and Bechger (2009), Tatsuoka (2009), DeCarlo (2011), and von Davier (2014). Considerable identifiability developments have been added to the CDM literature, such as DINA model and its generalizations in recent years. For instance, Xu and Zhang (2016) and Gu and Xu (2019b) discussed the sufficient and necessary condition for DINA model with binary responses. Xu (2017), Gu and Xu (2019a, 2020), Chen et al. (2020) and Culpepper (2022) discussed identifiability for more generally restricted latent class models. However, these results are targeted for dichotomous responses specifically, and the requirements for the identifiability of models with polytomous responses have sparingly been taken into consideration. For instance, Culpepper (2019) and Fang et al. (2019) discussed the sufficient condition for the identifiability of general CDMs with polytomous responses, while the necessity of those conditions remains an open problem.

Our paper fills this gap by providing sufficient and necessary conditions for the identifiability of CDMs with polytomous responses. In particular, we focus on two commonly used polytomous responses models under the DINA model setting: the GPDM (Chen & de la Torre, 2018) under the DINA model, which we refer as *GPDINA*, and the Sequential CDM (Ma & de la Torre, 2016) under the DINA model, which we refer as *Sequential DINA model*. There are several challenges in developing the identifiability of the polytomous responses models. Firstly, in binary responses DINA models, the uncertainty of each item is characterized by two item parameters, whereas in polytomous responses models, each item generally involves more than two parameters. Therefore, polytomous responses models have more parameters to identify, which makes its identifiability more challenging. What is more intricate is that the dependency structure between these parameters is different from that of the binary response models. This is because, in addition to accounting for dependencies across items, polytomous models must also consider the dependency of parameters within a single item. Moreover, the technical tool, **T**-matrix (Liu et al., 2013; Xu, 2017), which has been widely used in the identifiability literature, is restricted to binary responses models currently, to our knowledge.

To address these challenges, we generalize the **T**-matrix framework to the more complex polytomous model settings. Based on different dependency structures of the parameters of the two models, the generalizations of the **T**-matrix for the two considered models (i.e., GPDINA and sequential settings) are also different. In particular, there is a significant difference in the structure of the **T**-matrix for Sequential DINA model, as compared to the **T**-matrix for binary DINA models, since the sequential modeling introduces more complex and challenging structure than the binary DINA case. With this powerful tool, we establish sufficient and necessary conditions for the identifiability of the GPDINA and the Sequential DINA models. Our proposed conditions ensure the identifiability and also specify the practical requirements that the two models need to process to be identifiable. Through the duality of the DINA and DINO models (Chen et al., 2015), the identifiability finding can be immediately applied to the two models under the DINO setting. Moreover, our results not only extend many existing results aimed at binary data to the polytomous case, but also shed light on the study of more general polytomous CDMs, which cover the considered DINA models as submodels. Practically, the sufficient and necessary condition solely depend on the **Q**-matrix structure, and this easily verifiable requirement would serve as a practical guideline for developing cognitive tests that are both statistically valid and estimable.

The rest of the paper is organized as follows. Section 1 introduces the model setup and brings up the definition of identifiability. Section 2 introduces a powerful tool **T**-matrix, specific to the polytomous responses models and develops the identifiability results, examples are also provided for illustration. Section 3 gives further discussion, and the supplementary material provides the proofs for the main results.

1. Model Setup

Before we present our results, we first introduce some notations. Let $\boldsymbol{e}_j = (0, \ldots, 1, 0, \ldots, 0)^\top$ denote the vector where only the j-th entry is 1. Let $\mathbf{1} = (1, \ldots, 1)^\top$ denote the vector of all ones and $\mathbf{0} = (0, \ldots, 0)^\top$ denote the vector of all zeros. Let $\boldsymbol{\mathcal{I}}_K$ denote the K-dimensional identity matrix. For a positive integer m, we denote $[m] = \{1, \ldots, m\}$. Let $\circ$ denote the Hadamard product (element-wise product) of vectors. For instance, for $\boldsymbol{a} = (a_1, \ldots, a_m)^\top$ and $\boldsymbol{b} = (b_1, \ldots, b_m)^\top$, $\boldsymbol{a} \circ \boldsymbol{b} = (a_1 b_1, \ldots, a_m b_m)^\top$. Let $\otimes$ denote the Kronecker product between matrices. For example, for $\boldsymbol{c} = (c_1, \ldots, c_n)^\top \in \mathbb{R}^n$,

$$\boldsymbol{a} \otimes \boldsymbol{c} = \begin{pmatrix} a_1 \boldsymbol{c} \\ a_2 \boldsymbol{c} \\ \vdots \\ a_m \boldsymbol{c} \end{pmatrix} \in \mathbb{R}^{mn \times 1}, \quad \boldsymbol{a} \otimes \boldsymbol{\mathcal{I}}_K = \begin{pmatrix} a_1 \boldsymbol{\mathcal{I}}_K \\ a_2 \boldsymbol{\mathcal{I}}_K \\ \vdots \\ a_m \boldsymbol{\mathcal{I}}_K \end{pmatrix} \in \mathbb{R}^{mK \times K}.$$

Assume we have J polytomous items to measure K unobserved binary latent attributes, and a binary latent attribute profile can be written as $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K)^\top$, where $\alpha_k \in \{0, 1\}$. So there are 2^K attribute profiles in total. For $j \in [J]$, define positive integer H_j to be the number of nonzero categories (levels) the j-th polytomous item has, therefore, item j has $H_j + 1$ categories in total, i.e., $0, 1, \ldots, H_j$. Accordingly, we define the observed random variable response $\mathbf{R} = (R_1, \ldots, R_J)^\top$, with $R_j \in \{0, 1, \ldots, H_j\}$ and denote the set of all possible responses as $\mathcal{S} = \{\boldsymbol{r} = (r_1, \ldots r_J) : r_j \in \{0, 1, \ldots, H_j\}\}$.

In the CDM literature, the relationships between attributes and items are characterized by the $\mathbf{Q}$-matrix, which was proposed by Tatsuoka (1983). Different from CDMs with binary responses, for polytomous responses, the interpretations of the entries in the $\mathbf{Q}$-matrix differ according to different modelings. In the following, we focus on two popular models under the DINA assumption, the *general polytomous diagnosis model* (GPDINA) by Chen and de la Torre (2018) and the *Sequential DINA model* by Ma and de la Torre (2016) separately and introduce different ways of specifying the $\mathbf{Q}$-matrix for polytomous CDMs.

1.1. The GPDINA Model

In GPDINA (Chen & de la Torre, 2018) (the GPDM under the DINA assumption), for models with J items and K attributes, we define a $J \times K$ binary $\mathbf{Q}$-matrix. The entry q_{jk} of the $\mathbf{Q}$-matrix is interpreted as follows: $q_{jk} = 1$ means completing (responding) any nonzero category of item j requires attribute k, and $q_{jk} = 0$ means completing any nonzero categories does not require attribute k. So the j-th row of the $\mathbf{Q}$-matrix, $\mathbf{q}_j$, denotes the attributes required to complete any nonzero categories for item j. Therefore, any nonzero category of the same item requires the same attributes and shares the same $\mathbf{q}$-vector. In other words, nonzero categories of an item are indistinguishable and can be exchanged.

We consider the DINA assumption under the GPDINA framework. As in the DINA model for binary data, we denote the ideal response $\xi_{j,\boldsymbol{\alpha}} = I(\boldsymbol{\alpha} \succeq \mathbf{q}_j)$. To further quantify the uncertainty of the responses, define the item parameters as:

$$\theta^+_{j,l} := P(R_j = l \mid \xi_{j,\boldsymbol{\alpha}} = 1), \quad l \in [H_j], \tag{1}$$

$$\theta^-_{j,l} := P(R_j = l \mid \xi_{j,\boldsymbol{\alpha}} = 0), \quad l \in [H_j], \tag{2}$$

where $\theta^+_{j,l}$ means the probability of completing category l of item j given the attribute profile $\boldsymbol{\alpha}$ is capable of completing it and $\theta^-_{j,l}$ means the probability of completing category l of item

j given the attribute profile $\boldsymbol{\alpha}$ is not able to complete it. Then, $1-\theta^+_{j,l}$ can be interpreted as slipping parameter and $\theta^-_{j,l}$ interpreted as the guessing parameter (Junker & Sijtsma, 2001), and we assume that $\theta^+_{j,l} > \theta^-_{j,l}$ for $l \in [H_j]$ and $j \in [J]$. As we can see, although the attributes required by different categories of the same item are the same, here we allow the response uncertainty to be heterogeneous, i.e., $\theta^+_{j,l}$ and $\theta^-_{j,l}$ can be different across l. So in total we have $2\sum_{j=1}^{J} H_j$ item parameters, and the multiplicity of the item parameters is one of the aspects that makes polytomous responses models different from the binary DINA models. For notation convenience, we also let

$$P(R_j = 0 \mid \xi_{j,\boldsymbol{\alpha}} = 1) = 1 - \sum_{l=1}^{H_j} \theta^+_{j,l} \ := \theta^+_{j,0}, \tag{3}$$

$$P(R_j = 0 \mid \xi_{j,\boldsymbol{\alpha}} = 0) = 1 - \sum_{l=1}^{H_j} \theta^-_{j,l} \ := \theta^-_{j,0}. \tag{4}$$

When $\mathbf{q}_j = \mathbf{0}$, $\xi_{j,\boldsymbol{\alpha}} \equiv 1$ for all $\boldsymbol{\alpha}$, then $\theta^-_{j,l}$ is not defined for all $l \in [H_j]$. In Proposition 1, we will show that excluding these zero $\mathbf{q}$-vectors does not affect our analysis. Let

$$\boldsymbol{\theta}^+_j = (\theta^+_{j,1},\ \theta^+_{j,2},\ \ldots \theta^+_{j,H_j})^\top \text{ and } \boldsymbol{\theta}^-_j = (\theta^-_{j,1},\ \theta^-_{j,2},\ \ldots \theta^-_{j,H_j})^\top,$$

$\boldsymbol{\theta}^+ = (\boldsymbol{\theta}^+_j)_{j=1}^J$ and $\boldsymbol{\theta}^- = (\boldsymbol{\theta}^-_j)_{j=1}^J$, where there are $\sum_{j=1}^J H_j$ entries in both $\boldsymbol{\theta}^+$ and $\boldsymbol{\theta}^-$. Denote $p_{\boldsymbol{\alpha}}$ as the proportion of attribute profile $\boldsymbol{\alpha}$ in the population and $\boldsymbol{p} := (p_{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \{0,1\}^K)^\top$, which satisfies $\sum_{\boldsymbol{\alpha}\in\{0,1\}^K} p_{\boldsymbol{\alpha}} = 1$, and we assume that $p_{\boldsymbol{\alpha}} > 0$ for all $\boldsymbol{\alpha}$. Given the attribute profile $\boldsymbol{\alpha}$, assume that a subject's responses to the J items are independent. For $\boldsymbol{r} = (r_1, \ldots, r_J)^\top \in \mathcal{S}$, we have

$$P(\mathbf{R} = \boldsymbol{r} \mid \mathbf{Q}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p}) = \sum_{\boldsymbol{\alpha}\in\{0,1\}^K} p_{\boldsymbol{\alpha}} \prod_{j=1}^{J} (\theta^+_{j,r_j})^{\xi_{j,\boldsymbol{\alpha}}} (\theta^-_{j,r_j})^{(1-\xi_{j,\boldsymbol{\alpha}})}. \tag{5}$$

We use the following example to further the illustration of the model setup.

Example 1. Suppose there are two polytomous items, each with two nonzero categories, so then $J = 2$ and $H_1 = H_2 = 2$. Suppose only two attributes α_1 and α_2 are involved, and the $\mathbf{Q}$-matrix takes the following formula:

$$\mathbf{Q} = \begin{pmatrix} item\ 1 \left\{ \begin{array}{c} \mathbf{q}_1 = [1\ \ 0] \\ \hdashline \mathbf{q}_2 = [0\ \ 1] \end{array} \right. \\ item\ 2 \end{pmatrix}.$$

The dashline "- - -" is used to separate different items. Therefore, the first and the second categories of the first item both require solely α_1, and the first and the second categories of the second item both require solely α_2. In particular, attribute profile $\boldsymbol{\alpha} = (1, 0)$ has $\xi_{1,\boldsymbol{\alpha}} = 1$ and $\xi_{2,\boldsymbol{\alpha}} = 0$. Thus,

$$P(R_1 = 1 \mid \boldsymbol{\alpha}) = \theta^+_{1,1};\quad P(R_1 = 2 \mid \boldsymbol{\alpha}) = \theta^+_{1,2};\quad P(R_2 = 1 \mid \boldsymbol{\alpha}) = \theta^-_{2,1};\quad P(R_2 = 2 \mid \boldsymbol{\alpha}) = \theta^-_{2,2},$$

whereas for attribute profile $\boldsymbol{\alpha} = (0, 1)$, $\xi_{1,\boldsymbol{\alpha}} = 0$, $\xi_{2,\boldsymbol{\alpha}} = 1$, and

$$P(R_1 = 1 \mid \boldsymbol{\alpha}) = \theta^-_{1,1};\quad P(R_1 = 2 \mid \boldsymbol{\alpha}) = \theta^-_{1,2};\quad P(R_2 = 1 \mid \boldsymbol{\alpha}) = \theta^+_{2,1};\quad P(R_2 = 2 \mid \boldsymbol{\alpha}) = \theta^+_{2,2}.$$

Therefore, attribute profile $\boldsymbol{\alpha} = (1, 0)$ has higher probability of completing the two nonzero categories of the first item but lower probability of completing the two nonzero categories of the second item. Distributions for profiles with $\boldsymbol{\alpha} = (1, 1)$ and $\boldsymbol{\alpha} = (0, 0)$ can be similarly obtained as well.

Under the above GPDINA model setup, the model parameters include $(\boldsymbol{\theta}^-, \boldsymbol{\theta}^+, \boldsymbol{p})$. To study the identifiability of these parameters, we formally introduce the definition in the following, and we defer the identifiability result in Sect. 2.

Identifiability. We say that the GPDINA parameters are identifiable if there is no $(\bar{\boldsymbol{\theta}}^+, \bar{\boldsymbol{\theta}}^-, \bar{\boldsymbol{p}}) \neq (\boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$ such that

$$P(\mathbf{R} = \boldsymbol{r} \mid \mathbf{Q}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p}) = P(\mathbf{R} = \boldsymbol{r} \mid \mathbf{Q}, \bar{\boldsymbol{\theta}}^+, \bar{\boldsymbol{\theta}}^-, \bar{\boldsymbol{p}}) \text{ for all } \boldsymbol{r} \in \mathcal{S}. \tag{6}$$

To simplify our discussion of the identifiability issue, we assume that $\mathbf{q}_j \neq 0$ for all $j \in [J]$ without compromising the validity of the analysis, thanks to the following proposition.

Proposition 1. *Let* $\Delta = \{j \in [J] : \mathbf{q}_j = \mathbf{0}\}$ *denote the set of items whose* $\mathbf{q}$*-vectors are zero, then the GPDINA model parameters with* $\mathbf{Q}$*-matrix are identifiable if and only if the GPDINA model parameters with* $\mathbf{Q}_{-\Delta}$*-matrix are identifiable, where* $\mathbf{Q}_{-\Delta}$ *is obtained by removing the* $\mathbf{q}$*-vectors in* $\mathbf{Q}$ *corresponding to the items in* Δ.

1.2. The Sequential DINA Model

Another popular modeling approach for polytomous responses is the Sequential DINA model, proposed by Ma and de la Torre (2016). In the Sequential DINA model with J items and K attributes, we assume that the subject's response to item $j \in [J]$, with $R_j = 0$, indicates that the subject fails to complete the first category, and $R_j = r_j$ for $0 < r_j < H_j$ indicates that the subject has completed categories $1, \ldots, r_j$ successfully and failed to complete category $r_j + 1$. $R_j = H_j$ simply means the subject successfully completed all the categories. Therefore, categories within one item are not exchangeable, and such ordered categories make it different from the previous GPDINA model setup.

Due to the sequential hierarchy of the categories, different categories could require different attributes. What's worth noticing is that though response categories are assumed to be attained sequentially, there is no required structure for the attributes required by different categories. For each item j, its different categories should have their corresponding $\mathbf{q}$-vectors. In Ma and de la Torre (2016), they refer such $\mathbf{Q}$-matrix as *restricted* $\mathbf{Q}$-matrix. As defined, the polytomous item j has H_j nonzero categories, so for the associations between the attributes and the polytomous item j, we have H_j rows in the $\mathbf{Q}$-matrix to characterize such information. With each row having K entries indicating which attributes are required by the category, the $\mathbf{Q}$-matrix can be summarized as a $(\sum_{j=1}^{J} H_j) \times K$ binary matrix. Specifically, we index the $\mathbf{Q}$-matrix in the following way: For $l \in [H_j]$, we define the (j, l)-th row of the matrix, $\mathbf{q}_{j,l}$, as a K dimensional binary vector indicating the association between the category l of item j and the K attributes. According to our model construction, the $\mathbf{q}_{j,l}$ vector indicates the attributes required to complete category l of item j, *given that* the subject has successfully completed the previous categories $1, \ldots, l-1$. To further illustrate the model setup, we present an example in the following.

Example 2. Suppose there are two polytomous items with $H_1 = H_2 = 2$ and two attributes α_1 and α_2, and

$$\mathbf{Q} = \begin{pmatrix} item\ 1 & \begin{cases} category\ 1 & \mathbf{q}_{1,1} = [1\ \ 0] \\ category\ 2 & \mathbf{q}_{1,2} = [0\ \ 1] \end{cases} \\ item\ 2 & \begin{cases} category\ 1 & \mathbf{q}_{2,1} = [0\ \ 1] \\ category\ 2 & \mathbf{q}_{2,2} = [1\ \ 1] \end{cases} \end{pmatrix}. \tag{7}$$

Therefore, to complete the first category of the first item, a subject needs to require the first attribute, and given that the subject has completed the first category, he/she needs to require the second attribute to complete the second category of the first item.

Since different categories require different attributes, the ideal response needs to be specified accordingly to different categories. We define the ideal response as $\xi_{j,l,\boldsymbol{\alpha}} = I(\boldsymbol{\alpha} \succeq \mathbf{q}_{j,l})$ for category l of item j. This is also different from the setup in GPDINA, for which we only need to define item-wise ideal response. To quantify the uncertainty of the response to different categories, we define the item parameters specific to the Sequential DINA model as:

$$\beta^{+}_{j,l} := P(R_j \geq l \mid R_j \geq l-1,\ \xi_{j,l,\boldsymbol{\alpha}} = 1), \quad l \in [H_j], \tag{8}$$

$$\beta^{-}_{j,l} := P(R_j \geq l \mid R_j \geq l-1,\ \xi_{j,l,\boldsymbol{\alpha}} = 0), \quad l \in [H_j], \tag{9}$$

and we assume that $0 \leq \beta^{-}_{j,l} < \beta^{+}_{j,l} \leq 1$. Note that the inequality $\beta^{-}_{j,l} < \beta^{+}_{j,l}$ is assumed to respect the monotonicity assumption of the latent attributes (Xu & Zhang, 2016), which is also needed to avoid the label switching issue of the DINA model. Consequently, $\beta^{-}_{j,l}$ is permitted to take on values within the range [0, 1), while $\beta^{+}_{j,l}$ can take on values within the range (0, 1]. These parameters characterize the probability of completing category l of item j *given* a subject with attributes $\boldsymbol{\alpha}$ has completed the previous categories. Furthermore, $1 - \beta^{+}_{j,l}$ can be interpreted as the slipping parameter and $\beta^{-}_{j,l}$ interpreted as the guessing parameter (Junker & Sijtsma, 2001). Also, notice that

$$P(R_j \geq 0 \mid \xi_{j,l,\boldsymbol{\alpha}} = 1) = 1, \quad l \in [H_j],$$
$$P(R_j \geq 0 \mid \xi_{j,l,\boldsymbol{\alpha}} = 0) = 1, \quad l \in [H_j],$$

and we let $\beta^{+}_{j,H_j+1} = \beta^{-}_{j,H_j+1} = 0$.

To see how these item parameters are related to the model setup in Ma and de la Torre (2016), we formulate several concepts in their paper as the following. The *processing function* $S_j(l \mid \boldsymbol{\alpha})$ in Ma and de la Torre (2016), which denotes the probability of completing category l of item j provided that they have already completed category $l-1$ successfully, given the attribute profile $\boldsymbol{\alpha}$, can be written as

$$S_j(l \mid \boldsymbol{\alpha}) = (\beta^{+}_{j,l})^{\xi_{j,l,\boldsymbol{\alpha}}} (\beta^{-}_{j,l})^{1-\xi_{j,l,\boldsymbol{\alpha}}} = P(R_j \geq l \mid R_j \geq l-1,\ \boldsymbol{\alpha}), \quad l \in [H_j].$$

Let $S_j(0 \mid \boldsymbol{\alpha}) = P(R_j \geq 0 \mid \boldsymbol{\alpha}) = 1$ and $S_j(H_j + 1 \mid \boldsymbol{\alpha}) = 0$. Then, noticing that

$$\begin{aligned} P(R_j \geq r_j \mid \boldsymbol{\alpha}) &= \prod_{l=1}^{r_j} P(R_j \geq l \mid R_j \geq l-1,\ \boldsymbol{\alpha}) \cdot P(R_j \geq 0 \mid \boldsymbol{\alpha}) \\ &= \prod_{l=1}^{r_j} S_j(l \mid \boldsymbol{\alpha}) \\ &= \prod_{l=1}^{r_j} (\beta_{j,l}^{+})^{\xi_{j,l,\boldsymbol{\alpha}}} (\beta_{j,l}^{-})^{1-\xi_{j,l,\boldsymbol{\alpha}}}, \end{aligned}$$

given the attribute profile $\boldsymbol{\alpha}$, the probability of $R_j = r_j$ can be written as

$$\begin{aligned} P(R_j = r_j \mid \boldsymbol{\alpha}) &= P(R_j \geq r_j \mid \boldsymbol{\alpha}) - P(R_j \geq r_j + 1 \mid \boldsymbol{\alpha}) \\ &= \big[1 - S_j(r_j + 1 \mid \boldsymbol{\alpha})\big] \prod_{l=0}^{r_j} S_j(l \mid \boldsymbol{\alpha}). \end{aligned}$$

Similar to GPDINA, when $\mathbf{q}_{j,l} = \mathbf{0}$, $\xi_{j,l,\boldsymbol{\alpha}} \equiv 1$ for all $\boldsymbol{\alpha}$, and then $\beta_{j,l}^{-}$ is not defined. We will show later in Proposition 2 that excluding these categories with $\mathbf{q}_{j,l} = \mathbf{0}$ does not affect our analysis. Note that when $\beta_{j,l}^{-} = 0$ ($\mathbf{q}_{j,l}$ is not necessarily $\mathbf{0}$), some model parameters may not be well defined. Suppose category l^* is the first category in item j which appears to have $\beta_{j,l^*}^{-} = 0$, i.e., $\beta_{j,l}^{-} > 0$ for $l < l^*$. If we denote $\Gamma_{j,l^*}^{-} := \{\boldsymbol{\alpha} : \xi_{j,l^*,\boldsymbol{\alpha}} = 0\}$ as the set of attribute profiles that are not able to complete the l^*-th category of item j ideally, and if the probability of guessing correctly category l^* of item j is also 0, then there's no way for the subject to complete higher categories of item j. So we define for $\boldsymbol{\alpha} \in \Gamma_{j,l^*}^{-}$,

$$\beta_{j,l}^{+} = \beta_{j,l}^{-} = 0, \quad \text{for } l > l^*. \tag{10}$$

Assume that a subject's responses to the J items are conditionally independent given the attribute profiles. We let

$$\boldsymbol{\beta}_j^{+} = \left(\beta_{j,1}^{+},\ \beta_{j,1}^{+}\beta_{j,2}^{+},\ \ldots \prod_{l=1}^{H_j} \beta_{j,l}^{+}\right) \text{ and } \boldsymbol{\beta}_j^{-} = \left(\beta_{j,1}^{-},\ \beta_{j,1}^{-}\beta_{j,2}^{-},\ \ldots \prod_{l=1}^{H_j} \beta_{j,l}^{-}\right), \text{ for } j \in [J]$$

and $\boldsymbol{\beta}^{+} = (\boldsymbol{\beta}_1^{+}, \boldsymbol{\beta}_2^{+}, \ldots, \boldsymbol{\beta}_J^{+})$, $\boldsymbol{\beta}^{-} = (\boldsymbol{\beta}_1^{-}, \boldsymbol{\beta}_2^{-}, \ldots, \boldsymbol{\beta}_J^{-})$, then

$$P(\mathbf{R} = \boldsymbol{r} \mid \mathbf{Q}, \boldsymbol{\beta}^{+}, \boldsymbol{\beta}^{-}, \boldsymbol{p}) = \sum_{\boldsymbol{\alpha} \in \{0,1\}^K} p_{\boldsymbol{\alpha}} \prod_{j=1}^{J} P(R_j = r_j \mid \boldsymbol{\alpha}). \tag{11}$$

The Sequential DINA model parameters consist of $(\boldsymbol{\beta}^{+}, \boldsymbol{\beta}^{-}, \boldsymbol{p})$. Following the literature, we formally define the identifiability for the Sequential DINA model in the following.

Identifiability. We say that the Sequential DINA model parameters are identifiable if there is no $(\bar{\boldsymbol{\beta}}^+, \bar{\boldsymbol{\beta}}^-, \bar{\boldsymbol{p}}) \neq (\boldsymbol{\beta}^+, \boldsymbol{\beta}^-, \boldsymbol{p})$ such that

$$P(\mathbf{R} = \boldsymbol{r} \mid \mathbf{Q}, \boldsymbol{\beta}^+, \boldsymbol{\beta}^-, \boldsymbol{p}) = P(\mathbf{R} = \boldsymbol{r} \mid \mathbf{Q}, \bar{\boldsymbol{\beta}}^+, \bar{\boldsymbol{\beta}}^-, \bar{\boldsymbol{p}}) \text{ for all } \boldsymbol{r} \in \mathcal{S}. \tag{12}$$

Similar to GPDINA, in the following proposition, we show that excluding categories with $\mathbf{q}_{j,l} = \mathbf{0}$ does not influence our analysis of the identifiability. Therefore, for simplicity, we assume that $\mathbf{q}_{j,l} \neq 0$ for all $l \in [H_j]$, $j \in [J]$ in this paper.

Proposition 2. *Let* $\Delta^s = \left\{(j,l) : \mathbf{q}_{j,l} = \mathbf{0}\right\}$ *denote the set of categories whose* $\mathbf{q}$*-vectors are zero, then the Sequential DINA model parameters with* $\mathbf{Q}$*-matrix are identifiable if and only if the Sequential DINA model parameters with* $\mathbf{Q}_{-\Delta^s}$*-matrix are identifiable, where* $\mathbf{Q}_{-\Delta^s}$ *is obtained by removing the* $\mathbf{q}$*-vectors in* $\mathbf{Q}$ *corresponding to the categories in* Δ^s.

1.3. Relationship Between GPDINA and Sequential DINA Models

In this section, we briefly discuss the relation between the GPDINA model and the Sequential DINA model.

Fundamentally, GPDINA and Sequential DINA models differ by the hierarchy of the categories of items. In GPDINA, different nonzero categories of the same item can be exchanged and share the same $\mathbf{q}$-vector, whereas in Sequential DINA model, different nonzero categories are generally not exchangeable and need to be completed sequentially, and different nonzero categories are allowed to have arbitrarily different $\mathbf{q}$-vectors. However, when all the nonzero categories of an item share the same $\mathbf{q}$-vector, the Sequential DINA model becomes equivalent to GPDINA.

Formally, in Sequential DINA model, when $\mathbf{q}_{j,1} = \cdots = \mathbf{q}_{j,H_j}$, such $\mathbf{Q}$-matrix is referred to as *unrestricted* $\mathbf{Q}$-matrix (Ma & de la Torre, 2016), we have $\xi_{j,1,\boldsymbol{\alpha}} = \cdots = \xi_{j,H_j,\boldsymbol{\alpha}}$ for all $\boldsymbol{\alpha}$ and $j \in [J]$. Under this $\mathbf{Q}$-matrix, attribute profile $\boldsymbol{\alpha}$ is either capable of completing all the nonzero categories of an item or unable to complete any nonzero category. With these constraints, such $\mathbf{Q}$-matrix is also applicable to GPDINA, and we show that the two models are equivalent by presenting a bijective mapping from the item parameters of GPDINA to the parameters of the Sequential DINA model when the parameters are well defined. Specifically, for each item $j \in [J]$, we have the following relation between the two models' parameters.

From Sequential DINA model to GPDINA:

$$\begin{cases} P(R_j = l \mid \xi_{j,\boldsymbol{\alpha}} = 1) = \theta^+_{j,l} = (1 - \beta^+_{j,l+1}) \prod_{h=1}^{l} \beta^+_{j,h}, & \text{for } l \geq 1; \\ P(R_j = l \mid \xi_{j,\boldsymbol{\alpha}} = 0) = \theta^-_{j,l} = (1 - \beta^-_{j,l+1}) \prod_{h=1}^{l} \beta^-_{j,h}, & \text{for } l \geq 1. \end{cases}$$

From GPDINA to Sequential DINA model:

$$\begin{cases} P(R_j \geq l \mid R_j \geq l-1, \xi_{j,\boldsymbol{\alpha}} = 1) = \beta^+_{j,l} = \dfrac{\sum_{h=l}^{H_j} \theta^+_{j,h}}{\sum_{h=l-1}^{H_j} \theta^+_{j,h}}, & \text{for } l \geq 1; \\ P(R_j \geq l \mid R_j \geq l-1, \xi_{j,\boldsymbol{\alpha}} = 0) = \beta^-_{j,l} = \dfrac{\sum_{h=l}^{H_j} \theta^-_{j,h}}{\sum_{h=l-1}^{H_j} \theta^-_{j,h}}, & \text{for } l \geq 1. \end{cases}$$

By examining the above equations, it becomes apparent that there is a one-to-one correspondence between the parameters of the two models, demonstrating the equivalence of the two models under the considered $\mathbf{Q}$-matrix constraints.

2. Identifiability

This section introduces our identifiability results for the GPDINA model and the Sequential DINA model. To provide a foundation for these results, we first generalize the **T**-matrix, a powerful tool in the literature to establish the identifiability of CDMs with binary responses (Liu et al., 2013; Xu & Zhang, 2016; Xu, 2017), to polytomous models in Sect. 2.1. Since the structure of the two polytomous models differs, the **T**-matrix generalizations also differ, and we provide examples to illustrate this. We then formally present our identifiability results for the two models in Sects. 2.2 and 2.3, respectively.

2.1. Generalized **T***-Matrix for CDMs with Polytomous Responses*

Directly working on Eqs. (6) and (12) from the definitions of identifiability is challenging. Alternatively, we work on the marginal probability matrix, the **T**-matrix, firstly introduced by Liu et al. (2013), which sets up a linear dependence between attribute distribution and the response distribution. However, under the DINA model setting, most existing literature only focuses on the **T**-matrix for binary responses. For polytomous response DINA models, there are more parameters involved for each item, and these parameters cannot be naively treated separately. Our aim in this section is to generalize this powerful **T**-matrix tool to polytomous response models adjusted accordingly to the model setup.

2.1.1. **T***-Matrix for GPDINA* The **T**-matrix for GPDINA $\mathbf{T}(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-)$ is a $\prod_{j=1}^{J}(H_j+1) \times 2^K$ matrix, where the entries are indexed by row index $\boldsymbol{r} \in \mathcal{S}$ with $r_j \in \{0, 1, \ldots, H_j\}$ and column index $\boldsymbol{\alpha} \in \{0,1\}^K$. The $\boldsymbol{r}$-th row and $\boldsymbol{\alpha}$-th column entry of $\mathbf{T}(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-)$, denoted by $t_{\boldsymbol{r},\boldsymbol{\alpha}}(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-)$, is defined as

$$t_{\boldsymbol{r},\boldsymbol{\alpha}}(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-) = P\left(\bigcap_{j:r_j\neq 0} \{R_j = r_j\} \mid \mathbf{Q}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{\alpha}\right) = \prod_{j:r_j\neq 0} P\left(R_j = r_j \mid \mathbf{Q}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{\alpha}\right).$$

When $\boldsymbol{r} = \mathbf{0}$, $t_{\mathbf{0},\boldsymbol{\alpha}}(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-) = 1$ for any $\boldsymbol{\alpha}$. When $\boldsymbol{r} = r_j \cdot \boldsymbol{e}_j$,

$$t_{\boldsymbol{r},\boldsymbol{\alpha}}(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-) = P(R_j = r_j \mid \mathbf{Q}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{\alpha}).$$

Let $\mathbf{T}_{\boldsymbol{r}}(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-)$ be the row vector in the **T**-matrix corresponding to $\boldsymbol{r}$. Then for any $\boldsymbol{r} \neq \mathbf{0}$, we can write $\mathbf{T}_{\boldsymbol{r}}(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-) = \underset{j:r_j\neq 0}{\circ} \mathbf{T}_{r_j\cdot\boldsymbol{e}_j}(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-)$, where $\circ$ is the element-wise product of the row vectors. Since there exists a one-to-one mapping between $\mathbf{T}_{\boldsymbol{r}}$ and $P(\mathbf{R} = \boldsymbol{r} \mid \mathbf{Q}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$ for all $\boldsymbol{r} \in \mathcal{S}$, we may substitute the original identifiability problem with an equivalent statement as follows.

Lemma 1. *Following the definition in* (6) *and letting the attribute* $\boldsymbol{\alpha}$ *index of* $\boldsymbol{p}$ *be consistent with the* $\boldsymbol{\alpha}$ *index in* **T***, the GPDINA parameters are identifiable if and only if there is no* $(\bar{\boldsymbol{\theta}}^+, \bar{\boldsymbol{\theta}}^-, \bar{\boldsymbol{p}}) \neq (\boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$ *such that*

$$\mathbf{T}\boldsymbol{p} = \bar{\mathbf{T}}\bar{\boldsymbol{p}}. \tag{13}$$

To illustrate the construction of the **T**-matrix, we provide an example in the following.

Example 3. For the **Q**-matrix given in Example 1, the **T**-matrix for this **Q**-matrix is

$$\mathbf{T} = \begin{pmatrix} \boldsymbol{\alpha}: & (0,0) & (1,0) & (0,1) & (1,1) \\ \mathbf{T}_{\boldsymbol{r}=(0,0)} & 1 & 1 & 1 & 1 \\ \mathbf{T}_{\boldsymbol{r}=(1,0)} & \theta_{1,1}^{-} & \theta_{1,1}^{+} & \theta_{1,1}^{-} & \theta_{1,1}^{+} \\ \mathbf{T}_{\boldsymbol{r}=(2,0)} & \theta_{1,2}^{-} & \theta_{1,2}^{+} & \theta_{1,2}^{-} & \theta_{1,2}^{+} \\ \mathbf{T}_{\boldsymbol{r}=(0,1)} & \theta_{2,1}^{-} & \theta_{2,1}^{-} & \theta_{2,1}^{+} & \theta_{2,1}^{+} \\ \mathbf{T}_{\boldsymbol{r}=(0,2)} & \theta_{2,2}^{-} & \theta_{2,2}^{-} & \theta_{2,2}^{+} & \theta_{2,2}^{+} \\ \mathbf{T}_{\boldsymbol{r}=(1,1)} & \theta_{1,1}^{-}\theta_{2,1}^{-} & \theta_{1,1}^{+}\theta_{2,1}^{-} & \theta_{1,1}^{-}\theta_{2,1}^{+} & \theta_{1,1}^{+}\theta_{2,1}^{+} \\ \mathbf{T}_{\boldsymbol{r}=(2,1)} & \theta_{1,2}^{-}\theta_{2,1}^{-} & \theta_{1,2}^{+}\theta_{2,1}^{-} & \theta_{1,2}^{-}\theta_{2,1}^{+} & \theta_{1,2}^{+}\theta_{2,1}^{+} \\ \mathbf{T}_{\boldsymbol{r}=(1,2)} & \theta_{1,1}^{-}\theta_{2,2}^{-} & \theta_{1,1}^{+}\theta_{2,2}^{-} & \theta_{1,1}^{-}\theta_{2,2}^{+} & \theta_{1,1}^{+}\theta_{2,2}^{+} \\ \mathbf{T}_{\boldsymbol{r}=(2,2)} & \theta_{1,2}^{-}\theta_{2,2}^{-} & \theta_{1,2}^{+}\theta_{2,2}^{-} & \theta_{1,2}^{-}\theta_{2,2}^{+} & \theta_{1,2}^{+}\theta_{2,2}^{+} \end{pmatrix},$$

where $\mathbf{T}_{\boldsymbol{r}=(1,1)} = \mathbf{T}_{\boldsymbol{r}=(1,0)} \circ \mathbf{T}_{\boldsymbol{r}=(0,1)}$, $\mathbf{T}_{\boldsymbol{r}=(2,1)} = \mathbf{T}_{\boldsymbol{r}=(2,0)} \circ \mathbf{T}_{\boldsymbol{r}=(0,1)}$, $\mathbf{T}_{\boldsymbol{r}=(1,2)} = \mathbf{T}_{\boldsymbol{r}=(1,0)} \circ \mathbf{T}_{\boldsymbol{r}=(0,2)}$, $\mathbf{T}_{\boldsymbol{r}=(2,2)} = \mathbf{T}_{\boldsymbol{r}=(2,0)} \circ \mathbf{T}_{\boldsymbol{r}=(0,2)}$. We can see that the **T**-matrix's structure is the same as the classic **T**-matrix for binary DINA model, where the entries of the **T**-matrix involve at most two parameters.

2.1.2. **T**-*Matrix for Sequential DINA Model* Similarly, we generalize the **T**-matrix for the Sequential DINA model. However, due to the special structure of the Sequential DINA model, the generalization of the **T**-matrix here is slightly different from the literature, which we denote as $\mathbf{T}^s$-matrix, where the "s" stands for Sequential DINA model. Let the entries of $\mathbf{T}^s$-matrix $\mathbf{T}^s(\boldsymbol{\beta}^+, \boldsymbol{\beta}^-)$ be indexed by row index $\boldsymbol{r} \in \mathcal{S}$ and column index $\boldsymbol{\alpha} \in \{0,1\}^K$. The $\boldsymbol{r}$-th row and $\boldsymbol{\alpha}$-th column entry of $\mathbf{T}^s(\boldsymbol{\beta}^+, \boldsymbol{\beta}^-)$, denoted by $t^s_{\boldsymbol{r},\boldsymbol{\alpha}}(\boldsymbol{\beta}^+, \boldsymbol{\beta}^-)$, is defined as

$$\begin{aligned} t^s_{\boldsymbol{r},\boldsymbol{\alpha}}(\boldsymbol{\beta}^+, \boldsymbol{\beta}^-) &= P\left(\bigcap_{j:r_j\neq 0} \{R_j \geq r_j\} \mid \mathbf{Q}, \boldsymbol{\beta}^+, \boldsymbol{\beta}^-, \boldsymbol{\alpha}\right) \\ &= \prod_{j:r_j\neq 0} P\left(R_j \geq r_j \mid \mathbf{Q}, \boldsymbol{\beta}^+, \boldsymbol{\beta}^-, \boldsymbol{\alpha}\right) \\ &= \prod_{j:r_j\neq 0} \prod_{l=1}^{r_j} (\beta_{j,l}^+)^{\xi_{j,l,\boldsymbol{\alpha}}} (\beta_{j,l}^-)^{1-\xi_{j,l,\boldsymbol{\alpha}}}. \end{aligned}$$

Apparently, $t^s_{\mathbf{0},\boldsymbol{\alpha}}(\boldsymbol{\beta}^+, \boldsymbol{\beta}^-) = \mathbf{1}$ for any $\boldsymbol{\alpha}$. When $\boldsymbol{r} = r_j \cdot \boldsymbol{e}_j$,

$$t^s_{r_j\cdot\boldsymbol{e}_j,\boldsymbol{\alpha}}(\boldsymbol{\beta}^+, \boldsymbol{\beta}^-) = P(R_j \geq r_j \mid \mathbf{Q}, \boldsymbol{\beta}^+, \boldsymbol{\beta}^-, \boldsymbol{\alpha}) = \prod_{l=1}^{r_j} (\beta_{j,l}^+)^{\xi_{j,l,\boldsymbol{\alpha}}} (\beta_{j,l}^-)^{1-\xi_{j,l,\boldsymbol{\alpha}}}.$$

Let $\mathbf{T}^s_{\boldsymbol{r}}(\boldsymbol{\beta}^+, \boldsymbol{\beta}^-)$ be the row vector in the $\mathbf{T}^s$-matrix corresponding to $\boldsymbol{r}$. Then for any $\boldsymbol{r} \neq \mathbf{0}$, we can write $\mathbf{T}^s_{\boldsymbol{r}}(\boldsymbol{\beta}^+, \boldsymbol{\beta}^-) = \underset{j:r_j\neq 0}{\circ} \mathbf{T}^s_{r_j\cdot\boldsymbol{e}_j}(\boldsymbol{\beta}^+, \boldsymbol{\beta}^-)$. Similarly, due to the one-to-one mapping between $\mathbf{T}^s_{\boldsymbol{r}}$ and $P(\mathbf{R} \geq \boldsymbol{r} \mid \mathbf{Q}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$ for all $\boldsymbol{r} \in \mathcal{S}$, we may substitute the original identifiability problem using the $\mathbf{T}^s$-matrix technique, and we state this consequence in the following lemma.

Lemma 2. *Following the definition in* (12) *and letting the attribute* $\boldsymbol{\alpha}$ *index in* $\boldsymbol{p}$ *be consistent with the* $\boldsymbol{\alpha}$ *index in* $\mathbf{T}$*, the Sequential DINA model parameters are identifiable if and only if there is no* $(\bar{\boldsymbol{\beta}}^{+}, \bar{\boldsymbol{\beta}}^{-}, \bar{\boldsymbol{p}}) \neq (\boldsymbol{\beta}^{+}, \boldsymbol{\beta}^{-}, \boldsymbol{p})$ *such that*

$$\mathbf{T}^{s}\boldsymbol{p} = \bar{\mathbf{T}}^{s}\bar{\boldsymbol{p}}. \tag{14}$$

In the following, we present the $\mathbf{T}^{s}$-matrix for the model given in Example 2. Due to the unique structure of the Sequential DINA model, the $\mathbf{T}^{s}$-matrix is designed in a very different way from a standard $\mathbf{T}$-matrix for the DINA model.

Example 4. For the $\mathbf{Q}$-matrix given in Example 2, the $\mathbf{T}^{s}$-matrix for this $\mathbf{Q}$-matrix is

$$\mathbf{T}^{s} = \begin{pmatrix} \boldsymbol{\alpha}: & (0,0) & (1,0) & (0,1) & (1,1) \\ \mathbf{T}^{s}_{\boldsymbol{r}=(0,0)} & 1 & 1 & 1 & 1 \\ \mathbf{T}^{s}_{\boldsymbol{r}=(1,0)} & \beta^{-}_{1,1} & \beta^{+}_{1,1} & \beta^{-}_{1,1} & \beta^{+}_{1,1} \\ \mathbf{T}^{s}_{\boldsymbol{r}=(2,0)} & \beta^{-}_{1,1}\beta^{-}_{1,2} & \beta^{+}_{1,1}\beta^{-}_{1,2} & \beta^{-}_{1,1}\beta^{+}_{1,2} & \beta^{+}_{1,1}\beta^{+}_{1,2} \\ \mathbf{T}^{s}_{\boldsymbol{r}=(0,1)} & \beta^{-}_{2,1} & \beta^{-}_{2,1} & \beta^{+}_{2,1} & \beta^{+}_{2,1} \\ \mathbf{T}^{s}_{\boldsymbol{r}=(0,2)} & \beta^{-}_{2,1}\beta^{-}_{2,2} & \beta^{-}_{2,1}\beta^{-}_{2,2} & \beta^{+}_{2,1}\beta^{-}_{2,2} & \beta^{+}_{2,1}\beta^{+}_{2,2} \\ \mathbf{T}^{s}_{\boldsymbol{r}=(1,1)} & \beta^{-}_{1,1}\beta^{-}_{2,1} & \beta^{+}_{1,1}\beta^{-}_{2,1} & \beta^{-}_{1,1}\beta^{+}_{2,1} & \beta^{+}_{1,1}\beta^{+}_{2,1} \\ \mathbf{T}^{s}_{\boldsymbol{r}=(2,1)} & \beta^{-}_{1,1}\beta^{-}_{1,2}\beta^{-}_{2,1} & \beta^{+}_{1,1}\beta^{-}_{1,2}\beta^{-}_{2,1} & \beta^{-}_{1,1}\beta^{+}_{1,2}\beta^{+}_{2,1} & \beta^{+}_{1,1}\beta^{+}_{1,2}\beta^{+}_{2,1} \\ \mathbf{T}^{s}_{\boldsymbol{r}=(1,2)} & \beta^{-}_{1,1}\beta^{-}_{2,1}\beta^{-}_{2,2} & \beta^{+}_{1,1}\beta^{-}_{2,1}\beta^{-}_{2,2} & \beta^{-}_{1,1}\beta^{+}_{2,1}\beta^{-}_{2,2} & \beta^{+}_{1,1}\beta^{+}_{2,1}\beta^{+}_{2,2} \\ \mathbf{T}^{s}_{\boldsymbol{r}=(2,2)} & \beta^{-}_{1,1}\beta^{-}_{1,2}\beta^{-}_{2,1}\beta^{-}_{2,2} & \beta^{+}_{1,1}\beta^{-}_{1,2}\beta^{-}_{2,1}\beta^{-}_{2,2} & \beta^{-}_{1,1}\beta^{+}_{1,2}\beta^{+}_{2,1}\beta^{-}_{2,2} & \beta^{+}_{1,1}\beta^{+}_{1,2}\beta^{+}_{2,1}\beta^{+}_{2,2} \end{pmatrix}$$

where $\mathbf{T}^{s}_{\boldsymbol{r}=(1,1)} = \mathbf{T}^{s}_{\boldsymbol{r}=(1,0)} \circ \mathbf{T}^{s}_{\boldsymbol{r}=(0,1)}$, $\mathbf{T}^{s}_{\boldsymbol{r}=(2,1)} = \mathbf{T}^{s}_{\boldsymbol{r}=(2,0)} \circ \mathbf{T}^{s}_{\boldsymbol{r}=(0,1)}$, $\mathbf{T}^{s}_{\boldsymbol{r}=(1,2)} = \mathbf{T}^{s}_{\boldsymbol{r}=(1,0)} \circ \mathbf{T}^{s}_{\boldsymbol{r}=(0,2)}$, and $\mathbf{T}^{s}_{\boldsymbol{r}=(2,2)} = \mathbf{T}^{s}_{\boldsymbol{r}=(2,0)} \circ \mathbf{T}^{s}_{\boldsymbol{r}=(0,2)}$.

Unlike the $\mathbf{T}$-matrix for GPDINA, the entries of the $\mathbf{T}^{s}$-matrix for the Sequential DINA model usually involve more than two parameters, making identifying them technically more challenging. For instance, $\mathbf{T}^{s}_{\boldsymbol{r}=(2,2)}$ in the Sequential DINA model has four parameters in each entry, whereas $\mathbf{T}_{\boldsymbol{r}=(2,2)}$ in GPDINA only has two parameters in each entry. The following sections give a more detailed discussion of the identifiability issue.

2.2. *Identifiability of the GPDINA Model*

In this section, we develop the sufficient and necessary condition for the identifiability of the GPDINA model. To begin with, we introduce the terminology "completeness" for $\mathbf{Q}$-matrix, which was firstly proposed by Chiu et al. (2009). A $\mathbf{Q}$-matrix is said to be complete if it can differentiate all latent attribute profiles. Under the DINA model with binary responses, it requires that for each attribute, there exists some item requiring solely that attribute; that is, a complete $\mathbf{Q}$-matrix must contain an identity matrix $\mathcal{I}_K$ up to some row permutations, which can be written as

$$\mathbf{Q} = \begin{pmatrix} \mathcal{I}_K \\ \mathbf{Q}^{*} \end{pmatrix}_{J \times K}. \tag{15}$$

Similar to the binary response case (Xu & Zhang, 2016), the completeness of the $\mathbf{Q}$-matrix is necessary for the identifiability of the population proportion parameter $\boldsymbol{p}$. Additionally, each attribute must be required by a certain amount of items, and formally we state these conditions as follows.

Condition C1. *The* $\mathbf{Q}$*-matrix must be complete, taking the form (15).*

Condition C2. *Each of the K attributes is required by at least three items.*

Condition C3. *Any two different columns of the submatrix $\mathbf{Q}^*$ in (15) are distinct.*

Theorem 1. *Conditions C1–C3 are sufficient and necessary for the identifiability of the parameters of the GPDINA model.*

Remark 1. When $H_j = 1$ for all $j \in [J]$, the model is reduced to binary DINA model, and the result we develop here is consistent with the result in Gu and Xu (2019b).

Remark 2. While the identifiability conditions are the same as those for the DINA model with binary responses Gu and Xu (2019b), we would like to emphasize several significant distinctions. In the case of the DINA model with binary responses, the uncertainty of each item is characterized by two parameters—the slipping and guessing parameters. In contrast, the GPDINA model with polytomous responses introduces more than two parameters for each item, significantly complicating the models and rendering the study of identifiability more challenging. In particular, as discussed in Sect. 2.1, one crucial theoretical tool commonly employed in the literature to investigate the identifiability of the DINA model is the **T**-matrix, which is primarily designed for binary response models (Xu, 2027; Gu and Xu, 2019). However, when extending our focus to the polytomous response scenario such as the GPDINA model, it cannot be directly applied and a generalization of this tool becomes necessary. The first contribution of our work, detailed in Sect. 2.1, lies in this generalization, extending the applicability of these analytical techniques to a broader class of cognitive diagnosis models. Moreover, with the newly developed **T**-matrix tool, significant efforts and new techniques are involved in the establishment of our new results. From the sufficient condition perspective, although conditions C1–C3 are also the counterparts of those of the DINA model, it is not immediately evident if these conditions, transposed from the binary model, are still capable of capturing the complexity and ensuring the identifiability of the more parameter-rich GPDINA model. Additionally, from the necessary condition perspective, evaluating the necessity of conditions C1–C3 for the GPDINA model is more challenging than that of the DINA model with binary responses, due to the increased complexity of the GPDINA model, as illustrated in the following example and our proof of the theorem.

The completeness of the **Q**-matrix is necessary for the identifiability of the population proportion parameters, which follows from a similar argument as the binary DINA model (Gu & Xu, 2019b). See our proof in Supplementary Material for more details. To illustrate the necessity of the second condition C2 and the third condition C3, we consider a simple case when $K = 2$ in the following example.

Example 5. We illustrate the necessity of the conditions C2 and C3 with an example with $K = 2$. We first consider the necessity of the second condition. Suppose the **Q**-matrix is complete, but does not satisfy condition C2, i.e., there exists some attribute which is required by at most two items. Without loss of generality (WLOG), assume that this is the first attribute. According to Proposition 1, $\mathbf{q}_j \neq \mathbf{0}$ for all $j \in [J]$, so the **Q**-matrix can be written as one of the following:

$$\mathbf{Q} = \begin{pmatrix} 1 & 0 \\ \hdashline 1 & 0 \\ \hdashline 0 & 1 \\ \hdashline \vdots & \vdots \\ \hdashline 0 & 1 \end{pmatrix}_{J\times 2} \quad \text{or} \quad \mathbf{Q} = \begin{pmatrix} 1 & 0 \\ \hdashline 1 & 1 \\ \hdashline 0 & 1 \\ \hdashline \vdots & \vdots \\ \hdashline 0 & 1 \end{pmatrix}_{J\times 2}, \tag{16}$$

where the dashline "- - -" indicates the separation of different items. For simplicity, we may assume that the **Q**-matrix takes the first formula. The case when the **Q**-matrix takes the other formula can be

similarly obtained. So then only the first and the second item require α_1. Under this $\mathbf{Q}$-matrix, we show that the model parameters $(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$ are not identifiable by constructing a set of parameters $(\bar{\boldsymbol{\theta}}^+, \bar{\boldsymbol{\theta}}^-, \bar{\boldsymbol{p}}) \neq (\boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$ which satisfy (6). Take $\bar{\boldsymbol{\theta}}^+ = \boldsymbol{\theta}^+$ and $\bar{\boldsymbol{\theta}}_j^- = \boldsymbol{\theta}_j^-$ for $j > 2$, and $\bar{p}_{(11)} + \bar{p}_{(01)} = p_{(11)} + p_{(01)}$. Next we show that the remaining parameters $(\boldsymbol{\theta}_1^-, \boldsymbol{\theta}_2^-, p_{(00)}, p_{(01)}, p_{(10)})$ are not identifiable. Using the $\mathbf{T}$-matrix tool, it can be shown that the non-identifiability occurs if the following equations hold: $P\big((R_1, R_2) = (r_1, r_2) \mid \mathbf{Q}, \bar{\boldsymbol{\theta}}^+, \bar{\boldsymbol{\theta}}^-, \bar{\boldsymbol{p}}\big) = P\big((R_1, R_2) = (r_1, r_2) \mid \mathbf{Q}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p}\big)$ for all $r_1 \in \{0, 1, \ldots H_1\}$, $r_2 \in \{0, 1, \ldots, H_2\}$, where (R_1, R_2) are the first two entries of the random response $\mathbf{R}$. These equations can be further expressed as the following equations:

$$\begin{cases} \bar{p}_{(00)} + \bar{p}_{(10)} + p_{(01)} + p_{(11)} = p_{(00)} + p_{(10)} + p_{(01)} + p_{(11)}; \\ \bar{\theta}_{1,l_1}^-[\bar{p}_{(00)} + \bar{p}_{(01)}] + \theta_{1,l_1}^+[\bar{p}_{(10)} + \bar{p}_{(11)}] = \theta_{1,l_1}^-[p_{(00)} + p_{(01)}] + \theta_{1,l_1}^+[p_{(10)} + p_{(11)}]; \\ \bar{\theta}_{2,l_2}^-[\bar{p}_{(00)} + \bar{p}_{(01)}] + \theta_{2,l_2}^+[\bar{p}_{(10)} + \bar{p}_{(11)}] = \theta_{2,l_2}^-[p_{(00)} + p_{(01)}] + \theta_{2,l_2}^+[p_{(10)} + p_{(11)}]; \\ \bar{\theta}_{1,l_1}^-\bar{\theta}_{2,l_2}^-[\bar{p}_{(00)} + \bar{p}_{(01)}] + \theta_{1,l_1}^+\theta_{2,l_2}^+[\bar{p}_{(10)} + \bar{p}_{(11)}] = \theta_{1,l_1}^-\theta_{2,l_2}^-[p_{(00)} + p_{(01)}] + \theta_{1,l_1}^+\theta_{2,l_2}^+[p_{(10)} + p_{(11)}]; \end{cases} \tag{17}$$

where $l_1 \in [H_1]$, $l_2 \in [H_2]$. Then, there are $(1 + H_1 + H_2 + H_1H_2)$ equations above in total. If we further let some $\kappa \in (0, 1)$ s.t.

$$\begin{pmatrix} \theta_{1,l_1}^- \\ \bar{\theta}_{1,l_1}^- \\ \theta_{1,l_1}^+ \\ \bar{\theta}_{1,l_1}^+ \end{pmatrix} = \kappa^{l_1 - 1} \begin{pmatrix} \theta_{1,1}^- \\ \bar{\theta}_{1,1}^- \\ \theta_{1,1}^+ \\ \bar{\theta}_{1,1}^+ \end{pmatrix}, \text{ and } \begin{pmatrix} \theta_{2,l_2}^- \\ \bar{\theta}_{2,l_2}^- \\ \theta_{2,l_2}^+ \\ \bar{\theta}_{2,l_2}^+ \end{pmatrix} = \kappa^{l_2 - 1} \begin{pmatrix} \theta_{2,1}^- \\ \bar{\theta}_{2,1}^- \\ \theta_{2,1}^+ \\ \bar{\theta}_{2,1}^+ \end{pmatrix} \quad \text{for } l_1 \in [H_1],\ l_2 \in [H_2], \tag{18}$$

then Eq. (17) can be reduced to four equations

$$\begin{cases} \bar{p}_{(00)} + \bar{p}_{(10)} + p_{(01)} + p_{(11)} = p_{(00)} + p_{(10)} + p_{(01)} + p_{(11)}; \\ \bar{\theta}_{1,1}^-[\bar{p}_{(00)} + \bar{p}_{(01)}] + \theta_{1,1}^+[\bar{p}_{(10)} + \bar{p}_{(11)}] = \theta_{1,1}^-[p_{(00)} + p_{(01)}] + \theta_{1,1}^+[p_{(10)} + p_{(11)}]; \\ \bar{\theta}_{2,1}^-[\bar{p}_{(00)} + \bar{p}_{(01)}] + \theta_{2,1}^+[\bar{p}_{(10)} + \bar{p}_{(11)}] = \theta_{2,1}^-[p_{(00)} + p_{(01)}] + \theta_{2,1}^+[p_{(10)} + p_{(11)}]; \\ \bar{\theta}_{1,1}^-\bar{\theta}_{2,1}^-[\bar{p}_{(00)} + \bar{p}_{(01)}] + \theta_{1,1}^+\theta_{2,1}^+[\bar{p}_{(10)} + \bar{p}_{(11)}] = \theta_{1,1}^-\theta_{2,1}^-[p_{(00)} + p_{(01)}] + \theta_{1,1}^+\theta_{2,1}^+[p_{(10)} + p_{(11)}]. \end{cases} \tag{19}$$

For any $(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$, there are 4 constraints in (19) but 5 parameters $(\bar{\theta}_{1,1}^-, \bar{\theta}_{2,1}^-, \bar{p}_{(00)}, \bar{p}_{(10)}, \bar{p}_{(01)})$ to solve. Therefore, there are infinitely many solutions and $(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$ are non-identifiable. As for the case when the $\mathbf{Q}$-matrix takes the other formula, the proof can be easily obtained with minor change of notation.

Next we prove the necessity of the third condition C3. Suppose the $\mathbf{Q}$-matrix is complete, according to Proposition 1, we may assume that the $\mathbf{Q}$-matrix has the following form up to some permutation:

$$\mathbf{Q} = \begin{pmatrix} \mathcal{I}_2 \\ \mathbf{1}_{J-2} \quad \mathbf{1}_{J-2} \end{pmatrix}_{J \times 2}. \tag{20}$$

Take $\bar{\boldsymbol{\theta}}^+ = \boldsymbol{\theta}^+$ and $\bar{\boldsymbol{\theta}}_j^- = \boldsymbol{\theta}_j^-$ for $j > 2$, and $\bar{p}_{(11)} = p_{(11)}$. Next we show the remaining parameters $(\boldsymbol{\theta}_1^-, \boldsymbol{\theta}_2^-, p_{(00)}, p_{(10)}, p_{(01)})$ are not identifiable. Using the $\mathbf{T}$-matrix tool, again we can show that the non-identifiability occurs if the following equations hold: $P\big((R_1, R_2) = (r_1, r_2) \mid \mathbf{Q}, \bar{\boldsymbol{\theta}}^+, \bar{\boldsymbol{\theta}}^-, \bar{\boldsymbol{p}}\big) = P\big((R_1, R_2) = (r_1, r_2) \mid \mathbf{Q}, \boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p}\big)$ for all $r_1 \in \{0, 1, \ldots H_1\}$, $r_2 \in \{0, 1, \ldots, H_2\}$, where (R_1, R_2) are the first two entries of the random response $\mathbf{R}$. These equations can be further expressed into $(1 + H_1 + H_2 + H_1H_2)$ equations similar to Eq. (17) with minor notation modification. Similarly, if we further let some $\kappa \in (0, 1)$ s.t. Eq. (18) hold, then

these equations can be reduced to only four equations. For any $(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$, there are four constraints but five parameters $(\bar{\theta}^-_{1,1}, \bar{\theta}^-_{2,1}, \bar{p}_{(00)}, \bar{p}_{(10)}, \bar{p}_{(01)})$ to solve. Therefore, there are infinitely many solutions and $(\boldsymbol{\theta}^+, \boldsymbol{\theta}^-, \boldsymbol{p})$ are non-identifiable. Thus, we have shown that that the conditions C2 and C3 are indeed necessary. For the proofs of more general cases and the sufficiency of the conditions, see Supplementary Material for more details.

2.3. Identifiability of the Sequential DINA Model

To study the identifiability of the Sequential DINA model, different techniques need to be developed. From the discussion in Sects. 2.1 and 2.2, the structure of the $\mathbf{T}^s$-matrix for the Sequential DINA model is different from the $\mathbf{T}$-matrix defined for the GPDINA model, since the rows of the $\mathbf{T}^s$-matrix of Sequential DINA corresponding to higher response categories generally involve more than two item parameters, making it different from the usual DINA model structure.

To address this issue, note that

$$\mathbf{T}^s_{\boldsymbol{e}_j} = P(R_j \geq 1 \mid \mathbf{Q}, \boldsymbol{\beta}^+, \boldsymbol{\beta}^-, \boldsymbol{\alpha}) = (\beta^+_{j,1})^{\xi_{j,1,\boldsymbol{\alpha}}} (\beta^-_{j,1})^{1-\xi_{j,1,\boldsymbol{\alpha}}} \tag{21}$$

only involves two parameters $\beta^+_{j,1}$ and $\beta^-_{j,1}$, which has a similar algebraic structure to that for the DINA model with binary responses, and thus working on these parameters firstly would be a good strategy to consider. The focus of these quantities can be interpreted as follows: Consider "binary" responses of the form $I(\text{item j} \geq 1)$, the Sequential DINA model is then reduced to a binary DINA model. According to equation (21), the uncertainty parameters for this model are $\{\beta^+_{j,1},\ \beta^-_{j,1}\}_{j\in[J]}$. The corresponding $\mathbf{T}$-matrix for this reduced model consists of exactly vectors $\left(\mathbf{T}^s_{\boldsymbol{e}_j}\right)_{j\in[J]}$ and their element-wise products. That is, let $\mathbf{T}^1$ denote the $\mathbf{T}$-matrix for the reduced model (here we compress the notation "s" in $\mathbf{T}^s$), which is a submatrix of $\mathbf{T}^s$-matrix, then

$$\mathbf{T}^1 = \begin{pmatrix} \overset{i_j}{\underset{l=i_1}{\circ}} \mathbf{T}^s_{\boldsymbol{e}_l} \end{pmatrix} \text{ for } i_1 < \cdots < i_j,\ j \in [J].$$

Furthermore, let $\mathbf{Q}^1$ denote the submatrix of the $\mathbf{Q}$-matrix for the first category of each item, i.e., $\mathbf{Q}^1 = (\mathbf{q}_{j,1})_{j\in J}$. Then, the $\mathbf{Q}$-matrix for the above reduced model is $\mathbf{Q}^1$, as only the attributes required for completing the first categories are in scope. For notation convenience, we let $\mathbf{Q}^1_{1:K}$ denote the submatrix of the $\mathbf{Q}^1$-matrix that consists of the $\mathbf{q}$-vectors for the first categories of the first K items, and $\mathbf{Q}^1_{K+1:J}$ denote the submatrix of $\mathbf{Q}^1$ that consists of the $\mathbf{q}$-vectors for the first categories of items $(K+1), \ldots, J$, i.e.,

$$\mathbf{Q}^1 = \begin{pmatrix} \mathbf{Q}^1_{1:K} \\ \mathbf{Q}^1_{K+1:J} \end{pmatrix}.$$

To better illustrate this idea, we present an example in the following.

Example 6. The $\mathbf{Q}^1$-matrix and the $\mathbf{T}^1$-matrix for the reduced model of Example 2 are:

$$\mathbf{Q}^1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{T}^1 = \begin{pmatrix} \boldsymbol{\alpha}: & (0,0) & (1,0) & (0,1) & (1,1) \\ \mathbf{T}^1_{\boldsymbol{r}=(0,0)} & 1 & 1 & 1 & 1 \\ \mathbf{T}^1_{\boldsymbol{r}=(1,0)} & \beta^-_{1,1} & \beta^+_{1,1} & \beta^-_{1,1} & \beta^+_{1,1} \\ \mathbf{T}^1_{\boldsymbol{r}=(0,1)} & \beta^-_{2,1} & \beta^-_{2,1} & \beta^+_{2,1} & \beta^+_{2,1} \\ \mathbf{T}^1_{\boldsymbol{r}=(1,1)} & \beta^-_{1,1}\beta^-_{2,1} & \beta^+_{1,1}\beta^-_{2,1} & \beta^-_{1,1}\beta^+_{2,1} & \beta^+_{1,1}\beta^+_{2,1} \end{pmatrix},$$

where $\mathbf{T}^1_{\boldsymbol{r}=(1,1)} = \mathbf{T}^1_{\boldsymbol{r}=(1,0)} \circ \mathbf{T}^1_{\boldsymbol{r}=(0,1)}$.

It turns out that the first category of each item plays a crucial role in the identifiability of the Sequential DINA model. Provided the first categories of the items are informative enough, based on the identifiability results for the DINA model with binary responses, we can identify the item parameters of the first categories and the population proportion parameters. More interestingly, we can show that the item parameters of the other categories can be identified based on these identified parameters without additional requirements. Motivated by this, we have the following sufficient condition for the identifiability of the Sequential DINA model.

Theorem 2. *The Sequential DINA model parameters are identifiable if the* $\mathbf{Q}^1$ *matrix satisfies the following conditions S1–S3.*

Condition S1. $\mathbf{Q}^1$*-matrix is complete, i.e., under some permutation,* $\mathbf{Q}^1_{1:K} = \mathcal{I}_K$.

Condition S2. *Each of the K attributes is required by at least three items' first categories.*

Condition S3. *Suppose* $\mathbf{Q}^1_{1:K} = \mathcal{I}_K$*, then any two different columns of* $\mathbf{Q}^1_{K+1:J}$ *are distinct.*

Remark 3. Conditions S1–S3 are similar to conditions C1–C3, with different target. S1–S3 are stated for $\mathbf{Q}^1$-matrix in the Sequential DINA model, whereas C1–C3 are stated for $\mathbf{Q}$-matrix in GPDINA. When $H_j \equiv 1$, both polytomous models are reduced to binary DINA model, and conditions C1–C3 are equivalent to S1–S3.

The conditions S1–S3, as sufficient conditions for identifying the Sequential DINA model, provide guidelines for practitioners to design $\mathbf{Q}$-matrix that validates identifiability. Based on the theorem, it is suggested to design $\mathbf{Q}$-matrix with informative first categories (satisfying S1–S3) to ensure identifiability.

On the other hand, sufficient these conditions are: Their requirements only rely on the model's first categories. With polytomous response data involving more categories, it is natural to ask whether other categories can aid in relaxing these conditions. It turns out that relaxing these conditions necessitates careful consideration. In the following, we examine the necessity of each condition, and our primary finding is that while these conditions are challenging to relax, with certain constraints that allow for other informative categories to help, they might be possible to be relaxed. The finding that these conditions are challenging to relax comes from the intrinsic sequential structure of the model. Specifically, we will show that condition S1 cannot be relaxed and conditions S2 and S3 are hard to relax as non-identifiable examples do exist with the absence of these conditions.

Our first claim is that without additional constraints, the first condition S1 cannot be relaxed, i.e., S1 is necessary.

Proposition 3. (Necessity of Condition S1) *Condition S1 is necessary for the identifiability of the parameters of the Sequential DINA model.*

For the convenience of the following discussion, we present the proof of Proposition 3.

Proof of Proposition 3. Suppose that the $\mathbf{Q}$-matrix does not satisfy condition S1, i.e., $\mathbf{Q}^1$ is not complete, then there exists some attribute that is not solely required by any item's first category. WLOG, assume that this is the first attribute, and thus any item's first category that requires the first attribute also requires some other attributes. We claim that the model parameters are not identifiable for such an incomplete $\mathbf{Q}^1$-matrix. Specifically, take $\beta^-_{j,1} \equiv 0$, for $j \in [J]$. Then, subjects with attribute profiles $\mathbf{0}$ and $\boldsymbol{e}_1$ are not able to complete the first categories of all the items. Since $\beta^-_{j,1} \equiv 0$, according to the model construction in Sect. 1.2, subjects with attribute

profiles $\mathbf{0}$ and $\boldsymbol{e}_1$ cannot complete the other categories either, and for attribute profiles $\mathbf{0}$ and $\boldsymbol{e}_1$, $\beta^+_{j,l} \equiv \beta^-_{j,l} \equiv 0$ for $l > 1$. Therefore, the the two attribute profiles $\mathbf{0}$ and $\boldsymbol{e}_1$ share the same probability of completing all the categories of all the items, which is zero, i.e., $t_{\boldsymbol{r},\mathbf{0}} = t_{\boldsymbol{r},\boldsymbol{e}_1} \equiv 0, \forall \boldsymbol{r}$. Thus, parameters $p_{\mathbf{0}}$ and $p_{\boldsymbol{e}_1}$ are not identifiable. □

In the above proof, we constructed a Sequential DINA model with $\beta^-_{j,1} \equiv 0$ so that the parameters of higher categories are defined to be zero for attribute profiles $\mathbf{0}$ and $\boldsymbol{e}_1$. Note that the identifiability definition requires any set of the parameters in the parameter space to be identifiable. With the model parameters space including $0 \leq \beta^-_{j,l} < \beta^+_{j,l} \leq 1$, in the proof of Proposition 3, showing the non-identifiability of the case $\beta^-_{j,1} = 0$ would be enough to establish our claim on the necessity of the completeness condition.

However, this example is tender and may no longer be valid if we add additional constraints for the model parameters; that is, we only focus on the identifiability of a subset of the model parameters space. For instance, if we restrict our model parameters to the subset $0 < \beta^-_{j,l} < \beta^+_{j,l} \leq 1$, then the necessity of S1 may not hold anymore. This is because by constraining $0 < \beta^-_{j,l} < \beta^+_{j,l} \leq 1$, we allow more categories to help identifying the parameters. The following gives an example of the model with identifiable parameters whose $\mathbf{Q}$-matrix does not satisfy condition S1 under the assumption that $0 < \beta^-_{j,l} < \beta^+_{j,l} \leq 1$.

Example 7. Assume that $0 < \beta^-_{j,l} < \beta^+_{j,l} \leq 1$, and consider the case when $K = 2$ where the $\mathbf{Q}$-matrix takes the following form:

$$\mathbf{Q} = \begin{pmatrix} item1 \begin{cases} 1 & 1 \\ 0 & 1 \end{cases} \\ item2 \begin{cases} 1 & 1 \end{cases} \\ item3 \begin{cases} 1 & 1 \end{cases} \\ item4 \begin{cases} 1 & 0 \end{cases} \\ item5 \begin{cases} 1 & 0 \end{cases} \\ item6 \begin{cases} 1 & 0 \end{cases} \end{pmatrix} \text{ and } \mathbf{Q}^1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}. \tag{22}$$

Clearly, the $\mathbf{Q}^1$-matrix does not satisfy the completeness condition, but the model parameters with this $\mathbf{Q}$-matrix are identifiable, whose proof is presented in Supplementary Material.

Remark 4. Through the above analysis, we can see that condition S1 is necessary in a strict sense, which may impose overly stringent requirements for practical cognitive diagnostic tests. Statistically, "strict sense" in this context refers to the standard identifiability definition of the model parameters that requires any set of the parameters in the parameter space to be identifiable (Gu & Xu, 2020). Contrary to the notion of strict identifiability is the notion of generic identifiability (Allman et al., 2009; Gu & Xu, 2020), where we allow for non-identifiability to happen within a zero-measure set.

This slightly weaker notion can often suffice for real data analysis purposes (Allman et al., 2009; Gu & Xu, 2020) and is therefore useful in practice. The extent to which our necessary conditions can be relaxed for generic identifiability of the Sequential DINA model needs further explorations in the future, and the above case with $\beta^-_{j,l} = 0$ in the Sequential DINA model is one of such example.

Next we study the necessity of conditions S2 and S3. It turns out that the analysis for conditions S2 and S3 is more complicated. We start by presenting two examples to illustrate that.

Example 8. Consider the case when $K = 2$ with two attributes α_1 and α_2, $J = 4$ items, and the Q-matrix takes the following form:

$$\mathbf{Q} = \begin{pmatrix} item\ 1 \begin{cases} 1 & 0 \end{cases} \\ item\ 2 \begin{cases} 0 & 1 \end{cases} \\ item\ 3 \begin{cases} 0 & 1 \end{cases} \\ item\ 4 \begin{cases} 1 & 1 \\ 1 & 0 \end{cases} \end{pmatrix} \text{ and } \mathbf{Q}^1 = \begin{pmatrix} 1\ 0 \\ 0\ 1 \\ 0\ 1 \\ 1\ 1 \end{pmatrix}. \tag{23}$$

The above **Q**-matrix satisfies conditions S1 and S3, but does not satisfy condition S2, and the model parameters are not identifiable.

Example 9. Consider the case when $K = 2$ with two attributes α_1 and α_2, $J = 4$ items, and the Q-matrix takes the following form:

$$\mathbf{Q} = \begin{pmatrix} item\ 1 \begin{cases} 1 & 0 \end{cases} \\ item\ 2 \begin{cases} 0 & 1 \end{cases} \\ item\ 3 \begin{cases} 1 & 1 \\ 1 & 0 \end{cases} \\ item\ 4 \begin{cases} 1 & 1 \end{cases} \end{pmatrix} \text{ and } \mathbf{Q}^1 = \begin{pmatrix} 1\ 0 \\ 0\ 1 \\ 1\ 1 \\ 1\ 1 \end{pmatrix}. \tag{24}$$

The above **Q**-matrix satisfies conditions S1 and S2, but does not satisfy condition S3, and the model parameters are not identifiable.

We defer the proofs of the non-identifiability of the above two examples in Supplementary Material. The preceding examples illustrate the difficulty in relaxing conditions S2 and S3, even in simple cases such as $J = 4$ and $K = 2$, where non-identifiable examples exist when these conditions are violated. For more general cases, relaxing these conditions could be even more challenging.

However, the existence of these examples does not necessarily mean that conditions S2 and S3 are always necessary. In fact, we construct two identifiable examples that do not satisfy conditions S2 and S3 in the following, which indicates that conditions S2 and S3 may not be necessary in general. The identifiability of the following two examples relies on other additional categories, which carry relevant information in place of the first categories. This is also aligned with intuition, as we expect other categories to contribute to the identification of the model parameters. In other words, with the help of other categories, the model parameters could possibly be identified.

Example 10. Consider the case when $K = 2$ with two attributes α_1 and α_2, and $J = 4$ items. Each item contains two categories, and the **Q**-matrix takes the following form:

$$\mathbf{Q} = \begin{pmatrix} item\ 1 \begin{cases} 1 & 0 \\ 0 & 1 \end{cases} \\ item\ 2 \begin{cases} 1 & 0 \\ 0 & 1 \end{cases} \\ item\ 3 \begin{cases} 0 & 1 \\ 1 & 0 \end{cases} \\ item\ 4 \begin{cases} 0 & 1 \\ 1 & 0 \end{cases} \end{pmatrix} \text{ and } \mathbf{Q}^1 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}. \tag{25}$$

The above $\mathbf{Q}^1$ matrix does not satisfy the condition S2, yet the model parameters are identifiable, whose proof is deferred to Supplementary Material.

Remark 5. Condition S2 assumes each attribute is required by three items' first categories. In the above example, both attributes α_1 and α_2 are required by only two items' first categories, yet the two attributes are also required by the second categories of other items, which provides additional information and eventually makes the model parameters identifiable. This suggests that the information provided by higher categories would also be helpful for the model identifiability.

Similarly, as illustrated in the following example, the role of the first category in condition S3 could also be replaced by other categories, which may make the model identifiable as well.

Example 11. Consider the case when $K = 2$ with two attributes α_1 and α_2, and $J = 5$ items, and the **Q**-matrix takes the following form:

$$\mathbf{Q} = \begin{pmatrix} item\ 1 \begin{cases} 1 & 0 \\ 0 & 1 \end{cases} \\ item\ 2 \begin{cases} 0 & 1 \\ 1 & 0 \end{cases} \\ item\ 3 \begin{cases} 1 & 1 \end{cases} \\ item\ 4 \begin{cases} 1 & 1 \end{cases} \\ item\ 5 \begin{cases} 1 & 1 \end{cases} \end{pmatrix} \text{ and } \mathbf{Q}^1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

The above $\mathbf{Q}^1$ matrix does not satisfy the condition S3, yet the model parameters are identifiable, whose proof is presented in Supplementary Material.

While the above two examples imply that the conditions S2 and S3 may not be necessary for the identifiability of the parameters for the Sequential DINA model, the following weaker versions of S2 and S3 (denoted as conditions S2* and S3*) are necessary for the model identifiability. This proposition is summarized as follows.

Proposition 4. (Necessity of Conditions S2* and S3*) *The Sequential DINA model parameters are identifiable only if the* **Q***-matrix satisfies the following conditions S2* and S3*.*

Condition S2* *Each of the K attributes is required by at least three categories (not necessarily the first categories), and the three categories must come from at least two different items.*
Condition S3* *Suppose* $\mathbf{Q}$*-matrix satisfies S1, i.e.,* $\mathbf{Q}^{1}_{1:K} = \mathcal{I}_K$*, and any two different columns of the following matrix (which removes the identity matrix of* $\mathbf{Q}^{1}_{1:K}$ *from* $\mathbf{Q}$*)*

$$\begin{pmatrix} \mathbf{Q}^{-1}_{1:K} \\ \mathbf{Q}_{K+1:J} \end{pmatrix}$$

are distinct, where $\mathbf{Q}^{-1}_{1:K}$ *denotes the remaining submatrix of* $\mathbf{Q}_{1:K}$ *after removing* $\mathbf{Q}^{1}_{1:K}$.

We can see that conditions S2 and S3 are stronger versions of S2* and S3*, which means that any $\mathbf{Q}$-matrix satisfying condition S2 (S3) will satisfy condition S2* (S3*). We can also see that the two identifiable models in Example 10 and Example 11 that do not satisfy conditions S2 and S3 both satisfy condition S2* and condition S3*. For instance, the $\mathbf{Q}$-matrix in Example 10 does not satisfy condition S2 since there are only two items' first categories require α_1 and only two items' first categories require α_2. However, it does satisfy condition S2*, since two other items' second categories require α_1 and other two items' second categories require α_2. Similarly, the $\mathbf{Q}$-matrix in Example 11, not satisfying condition S3, does satisfy condition S3*, as the second category of the first item requires only α_2 and the second category of the second item requires only α_1.

In summary, from the above discussions, we conclude that the sufficient conditions S1–S3 are challenging to relax. Specifically, condition S1 cannot be relaxed unless additional constraints are imposed. While conditions S2 and S3 are also difficult to relax, we found that other categories may assist in identifying the parameters.

In spite of the fact that the sufficient condition and the necessary condition proposed in this section are different, filling the gap is not an easy task, as the model structure is more subtle and the interactions between parameters are more complex. For instance, the $\mathbf{T}^s$-matrix structure is different from the $\mathbf{T}$-matrix structure for the binary DINA model except for the first categories. The $\mathbf{T}^s_{\boldsymbol{r}}$-vectors for higher categories behave more similar to the $\mathbf{T}_{\boldsymbol{r}}$-vectors for G-DINA model (de la Torre, 2011), as the uncertainty for these categories is characterized by more than two parameters. Therefore, to study the identifiability of the Sequential DINA model requires more techniques beyond the DINA setting.

3. Data Examples

In this section, we demonstrate the application of our proposed results by examining two educational assessment datasets: a PISA 2000 reading assessment dataset using the GPDINA model (Chen & de la Torre, 2018) and a TIMSS 2007 fourth-grade mathematics assessment dataset using the Sequential DINA model (Ma & de la Torre, 2016).

Identifiability of the GPDINA model: a PISA 2000 data example. We consider a dataset from the PISA 2000 reading assessment, which was previously studied in Chen and de la Torre (2018). This assessment, released by the OECD (1999, 2006), comprised both polytomous and binary items. The dataset for this application comprises responses from 1,039 English examinees to 20 specific items from a designated test booklet. Out of these 20 items, five are polytomous. Following Chen and de la Torre (2018), the attribute definitions for the PISA dataset are given in Table 1 and the $\mathbf{Q}$-matrix for this application is presented in Table 2. Since in the GPDINA model, different categories within the same item share the same $\mathbf{q}$-vectors, it suffices to provide one $\mathbf{q}$-vector for each item.

TABLE 1.
Attribute definitions for the PISA data (Chen & de la Torre, 2018).

Symbol	Description
c	Number of categories
α_1	Retrieving information
α_2	Forming a broad general understanding
α_3	Developing an interpretation
α_4	Reflecting on and evaluating the content of a text
α_5	Reflecting on and evaluating the form of a text

TABLE 2.
Items and **Q**-matrix for the PISA data (Chen & de la Torre, 2018).

No.	Item Code	c	α_1	α_2	α_3	α_4	α_5	No.	Item Code	c	α_1	α_2	α_3	α_4	α_5
1	R040Q02	2	1	0	1	0	0	11	R088Q04T	3	1	0	1	0	0
2	R040Q03A	2	1	0	1	1	0	12	R088Q05T	2	0	1	1	1	0
3	R040Q04	2	0	1	1	1	0	13	R088Q07	2	0	1	0	0	1
4	R040Q06	2	1	0	1	0	0	14	R216Q01	2	0	1	0	0	0
5	R077Q03	3	0	1	0	1	1	15	R216Q02	2	1	0	0	0	1
6	R077Q04	2	1	1	1	0	0	16	R216Q03T	2	0	1	1	0	0
7	R077Q05	3	0	1	1	1	0	17	R216Q04	2	0	1	1	0	0
8	R077Q06	2	0	1	0	0	1	18	R216Q06	2	0	1	0	1	0
9	R088Q01	2	0	1	1	0	0	19	R236Q01	2	1	0	1	0	0
10	R088Q03	3	1	0	1	0	0	20	R236Q02	3	0	0	1	1	0

According to our Theorem 1, this **Q**-matrix does not contain an identity matrix, and thus, the model parameters are not identifiable. Specifically, since the matrix does not contain $\boldsymbol{e}_1^\top$, $\boldsymbol{e}_3^\top$, $\boldsymbol{e}_4^\top$ and $\boldsymbol{e}_5^\top$, attribute profiles $\mathbf{0}$, $\boldsymbol{e}_1$, $\boldsymbol{e}_3$, $\boldsymbol{e}_4$ and $\boldsymbol{e}_5$ have the same conditional response distributions. Therefore, the parameters $p_{\mathbf{0}}$, $p_{\boldsymbol{e}_1}$, $p_{\boldsymbol{e}_3}$, $p_{\boldsymbol{e}_4}$ and $p_{\boldsymbol{e}_5}$ cannot be identified.

Identifiability of the Sequential DINA model: a TIMSS 2007 data example. We consider the dataset in Ma and de la Torre (2016), which is derived from booklets 4 and 5 of the TIMSS 2007 fourth-grade mathematics assessment. This subset, originally utilized by Lee et al. (2011), includes responses from 823 students to 12 items, which are linked to eight of the original 15 attributes. Notably, items 3 and 9 are constructed-response items scored polytomously across three response categories (0, 1, and 2). The dataset also features items like 7a and 7b which, due to their heavy interdependence, can be treated as a single polytomous item. We consider the Sequential DINA model in this example. Following Ma and de la Torre (2016), the attribute definitions for the TIMSS data are given in Table 3 and the **Q**-matrix is in Table 4. The corresponding $\mathbf{Q}^1$-matrix is also presented below.

TABLE 3.
Attribute definitions for TIMSS 2007 data (Ma & de la Torre, 2016).

Attribute	Description
α_1	Representing, comparing, and ordering whole numbers as well as demonstrating knowledge of place value
α_2	Recognizing multiples, computing with whole numbers using the four operations, and estimating computations
α_3	Solving problems, including those set in real-life contexts
α_4	Finding the missing number or operation and modeling simple situations involving unknowns in number sentence or expression
α_5	Describing relationships in patterns and their extensions; generating pairs of whole numbers by a given rule and identifying a rule for every relationship given pairs of whole numbers
α_6	Reading data from tables, pictographs, bar graphs, and pie charts
α_7	Comparing and understanding how to use information from data
α_8	Understanding different representations and organizing data using tables, pictographs, and bar graphs

TABLE 4.
Q-matrix for TIMSS 2007 data (Ma & de la Torre, 2016).

Item	TIMSS item no.	Category	Attributes							
			α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8
1	M041052	1	1	1	0	0	0	0	0	0
2	M041281	1	0	1	1	0	1	0	0	0
3**a**	M041275	1	1	0	0	0	0	1	0	1
3**b**	M041275	2	1	0	0	0	0	1	0	1
4	M031303	1	0	1	1	0	0	0	0	0
5	M031309	1	0	1	1	0	0	0	0	0
6	M031245	1	0	1	0	1	0	0	0	0
7**a**	M031242A	1	0	1	1	0	1	0	0	0
7**b**	M031242B	2	0	0	0	0	0	0	1	0
8	M031242C	1	0	1	1	0	1	0	1	0
9**a**	M031247	1	0	1	1	1	0	0	0	0
9b	M031247	2	0	1	1	1	0	0	0	0
10	M031173	1	0	1	1	0	0	0	0	0
11	M031172	1	1	1	0	0	0	1	0	1

$$\mathbf{Q}^1 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

According to Proposition 3, since the $\mathbf{Q}^1$-matrix does not contain an identity matrix, the model parameters are not identifiable. Specifically, since the matrix does not contain any $\boldsymbol{e}_j$ for $j = 1, 2, \ldots, 8$, if we take $\beta^-_{j,1} = 0$ for $j = 1, 2, \ldots 20$, then subjects with attribute profiles $\mathbf{0}$ and $\boldsymbol{e}_j$ for $j = 1, 2, \ldots, 8$ are not able to complete the first categories of all the items. Since $\beta^-_{j,1} \equiv 0$, according to the model construction in Sect. 1.2, these attribute profiles cannot complete other categories either. Therefore, attribute profiles $\mathbf{0}$, $\boldsymbol{e}_j$ for $j = 1, 2, \ldots, 8$ have the same probability of completing all the categories of all the items, which is zero. Therefore, the parameters $p_{\mathbf{0}}$, $p_{\boldsymbol{e}_j}$ for $j = 1, 2, \ldots, 8$ cannot be identified.

Remark 6. For the above educational assessment examples, while the analysis shows non-identifiability issues for the two considered models, this should not overshadow the potential for analyzing these data using polytomous DINA or more general cognitive diagnosis models. First, as discussed in Sect. 2.3, although the two models in our application data fail to satisfy the completeness condition, if we consider the more relaxed generic identifiability of the model parameters, that is, allowing non-identifiability of parameters in a negligible zero-measure set of the parameter space, the stringent completeness condition may not be necessary, as discussed in Gu and Xu (2020). Second, the investigation of partial identifiability, as proposed by Gu and Xu (2020), could also be extended to the current situation. Specifically, when the completeness condition is violated, partial identifiability may be established to partially identify the non-identifiable proportion parameters $\boldsymbol{p}$ up to their equivalent classes. For example, in the first example, since attribute profiles $\mathbf{0}$, $\boldsymbol{e}_1$, $\boldsymbol{e}_3$, $\boldsymbol{e}_4$ and $\boldsymbol{e}_5$ have the same conditional response distributions, they can be grouped and considered as an equivalent latent class. Partial identifiability then seeks to identify parameter $(p_{\mathbf{0}} + p_{\boldsymbol{e}_1} + p_{\boldsymbol{e}_3} + p_{\boldsymbol{e}_4} + p_{\boldsymbol{e}_5})$ as a whole, instead of treating each proportion parameter separately. Under such relaxation, the models applied to the data examples may be partially identifiable. Finally, beyond the DINA models considered in this paper, general cognitive diagnosis models (Chen & de la Torre, 2018; Ma & de la Torre, 2016) may be more appropriate for the two datasets, and studying the identifiability (Gu & Xu, 2020) of these models could be also of great interest. Further explorations of these interesting extensions are promising future research directions.

4. Discussion

This paper presents the sufficient and necessary conditions for the identifiability of CDMs with polytomous responses. Our results focus on two popular models under the DINA assumption: the GPDINA model and the Sequential DINA model. For both models, we provide the sufficient and necessary conditions for their identifiability. The results can be easily extended to the DINO (deterministic input; noisy "or" gate) model (Templin & Henson, 2006) through the duality between the DINA and DINO models. While the minimum requirements for more general CDMs are still unknown, our proposed necessary conditions remain necessary for them since our polytomous DINA models are submodels of the general CDMs. Therefore, our results would also shed light on the study of their identifiability.

The popularity of polytomous data is not restricted to response data, and polytomous attributes data are also receiving more and more attention (Haberman et al., 2008; von Davier, 2008; Chen & Torre, 2013; de la Torre et al., 2022). Yet the discussion on the identifiability of such models has sparingly been considered. More interestingly, we may further study the identifiability results under the general CDM framework with polytomous responses and polytomous attributes.

The $\mathbf{Q}$-matrix in this paper is assumed to be correctly specified. In practice, the $\mathbf{Q}$-matrix is usually constructed by the designers, which can be subjective and may not be accurate. For this reason, researchers have proposed to estimate and validate the design $\mathbf{Q}$-matrix based on the

response data, which motivates the study of the identifiability of the **Q**-matrix (e.g., Liu et al., 2013; Chen et al., 2015; Xu and Shang, 2018; Culpepper, 2019; Chen et al., 2020; Gu and Xu, 2021). Nevertheless, most of these existing works focus on dichotomous responses, and only few have explored the identifiability of **Q**-matrix in the polytomous data setting, which would also be an interesting future research topic.

Acknowledgments

This work is partially supported by NSF grants SES-1846747 and SES-2150601. We are grateful to the editor, an associate editor, and anonymous referees for their helpful comments and suggestions.

Data Availability Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

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

Chen, J., & de la Torre, J. (2018). Introducing the general polytomous diagnosis modeling framework. *Frontiers in Psychology, 9*, 1474.

Chen, J., & Torre, J. (2013). A general cognitive diagnosis model for expert-defined polytomous attributes. *Applied Psychological Measurement, 37*, 419–437.

Chen, Y., Culpepper, S., & Liang, F. (2020). A sparse latent class model for cognitive diagnosis. *Psychometrika, 85*, 121–153.

Chen, Y., Liu, J., Xu, G., & Ying, Z. (2015). Statistical analysis of *Q*-matrix based diagnostic classification models. *Journal of the American Statistical Association, 110*(510), 850–866.

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

Culpepper, S. A. (2019). An exploratory diagnostic model for ordinal responses with binary attributes: Identifiability and estimation. *Psychometrika, 84*(4), 921–940.

Culpepper, S. A. (2022). A note on weaker conditions for identifying restricted latent class models for binary responses. *Psychometrika*, pages 1–17.

Culpepper, S. A., & Balamuta, J. J. (2021). Inferring latent structure in polytomous data with a higher-order diagnostic model. *Multivariate Behavioral Research, 58*, 368–386.

de la Torre, J. (2011). The generalized DINA model framework. *Psychometrika, 76*(2), 179–199.

de la Torre, J., Qiu, X.-L.S., & Carl, K. (2022). An empirical Q-matrix validation method for the polytomous G-DINA model. *Psychometrika, 87*(2), 693–724.

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

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

DiBello, L. V., Stout, W. F., & Roussos, L. A. (1995). *Unified cognitive psychometric diagnostic assessment likelihood-based classification techniques*. Hillsdale, NJ: Erlbaum Associates.

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

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

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

Gu, Y., & Xu, G. (2020). Partial identifiability of restricted latent class models. *Annals of Statistics, 48*(4), 2082–2107.
Gu, Y., & Xu, G. (2021). Sufficient and necessary conditions for the identifiability of the Q-matrix. *Statistica Sinica, 31*, 449–472.
Haberman, S. J., von Davier, M., & Lee, Y.-H. (2008). Comparison of multidimensional item response models: Multivariate normal ability distributions versus multivariate polytomous ability distributions. *ETS Research Report Series*. https://doi.org/10.1002/j.2333-8504.2008.tb02131.x
Henson, R. A., Templin, J. L., & Willse, J. T. (2009). Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika, 74*(2), 191.
Junker, B. W., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement, 25*, 258–272.
Lee, Y.-S., Park, Y. S., & Taylan, D. (2011). A cognitive diagnostic modeling of attribute mastery in massachusetts, minnesota, and the u.s. national sample using the TIMSS 2007. *International Journal of Testing, 11*(2), 144–177.
Liu, J., Xu, G., & Ying, Z. (2013). Theory of self-learning *Q*-matrix. *Bernoulli, 19*(5A), 1790–1817.
Ma, W., & de la Torre, J. (2016). A sequential cognitive diagnosis model for polytomous responses. *British Journal of Mathematical and Statistical Psychology, 69*(3), 253–275.
Maris, G., & Bechger, T. M. (2009). Equivalent diagnostic classification models. *Measurement, 7*, 41–46.
O'Brien, K. L., Baggett, H. C., Brooks, W. A., Feikin, D. R., Hammitt, L. L., Higdon, M. M., 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, 394*, 757–779.
OECD. (1999). *Measuring Student Knowledge and Skills: A New Framework for Assessment*. Paris: Organisation for Economic Co-operation and Development.
OECD. (2006). *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006*. Paris: Organisation for Economic Co-operation and Development.
Rupp, A. A., Templin, J., & Henson, R. A. (2010). *Diagnostic measurement: Theory, methods, and applications*. New York City: Guilford Press.
Tatsuoka, C. (2009). Diagnostic models as partially ordered sets. *Measurement, 7*, 49–53.
Tatsuoka, K. K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement, 20*, 345–354.
Templin, J. L., & Henson, R. A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological Methods, 11*, 287–305.
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, S., Yang, Y., Culpepper, S. A., & 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., & 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., & Shang, Z. (2018). Identifying latent structures in restricted latent class models. *Journal of the American Statistical Association, 113*(523), 1284–1295.
Xu, G., & Zhang, S. (2016). Identifiability of diagnostic classification models. *Psychometrika, 81*, 625–649.

Manuscript Received: 7 APR 2023
Published Online Date: 22 MAR 2024