Minimal Complete Q-matrix Design for Polytomous Scoring Based on Reachability Matrix

Authors: Tang Xiaojuan, Peng Zhixia, Qin Shanshan, Ding Shuliang, Mao Mengmeng, Li Yu, Li Yu

Date: 2024-09-16T00:00:00+00:00

Abstract

The completeness of the Q-matrix is crucial for the identifiability of cognitive diagnosis models. Polytomous scoring contains richer diagnostic information than dichotomous (0-1) scoring, yet research on the design of complete Q-matrices for polytomous scoring is rarely seen. Test designers aim to achieve the highest classification accuracy with the minimum number of items. Drawing on design methods for complete Q-matrices under dichotomous scoring, this paper proposes methods and algorithms to obtain structured/unstructured simplest complete Q-matrices (SSCQM/USCQM) for polytomous scoring from the reachability matrix. Simulation studies yielded the following conclusions: (1) The more SSCQMs/USCQMs a test contains, the higher the classification accuracy; (2) When the number of columns is the same, the classification accuracy of tests containing multiple SSCQMs or multiple USCQMs is very close to that of tests containing the reachability matrix; (3) For some structures, even when the number of columns of multiple SSCQMs/USCQMs is fewer than that of the reachability matrix, their classification accuracy is still not lower than that of the reachability matrix. In summary, SSCQM is preferred for short test designs, while USCQM is preferred for long test designs.

Full Text

Design of the Polytomous Simplest Complete Q-Matrix Based on the Reachability Matrix

Tang Xiaojuan[1], Peng Zhixia[2], Qin Shanshan[2], Ding Shuliang[3], Mao Mengmeng , Li Yu

[1]School of Education, Jiangxi Normal University, Nanchang 330022, China
[2]School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou 310018, China

[3]College of Computer Information Engineering, Jiangxi Normal University, Nanchang 330022, China
School of Public Policy and Administration, Nanchang University, Nanchang 330036, China
Mental Health Education Center, School of Marxism, Zhejiang Gongshang University, Hangzhou 310018, China

Abstract

The completeness of the Q-matrix is crucial for the identifiability of cognitive diagnosis models. Polytomous scoring provides richer diagnostic information than dichotomous scoring, yet research on designing complete Q-matrices for polytomous scoring remains scarce. Minimizing test length while maximizing classification accuracy represents a primary objective for test designers. Drawing upon design methods for dichotomous complete Q-matrices, this paper proposes a method and algorithm for extracting polytomous structured/unstructured simplest complete Q-matrices (SSCQM/USCQM) from the reachability matrix. Simulation experiments yield the following conclusions: (1) Tests containing more SSCQMs/USCQMs achieve higher classification accuracy; (2) When the number of columns is equal, the classification accuracy of tests containing multiple SSCQMs or multiple USCQMs closely approximates that of tests containing the reachability matrix; (3) For certain structures, even when multiple SSCQMs/USCQMs have fewer columns than the reachability matrix, their classification accuracy remains no lower than that of the reachability matrix. In summary, SSCQM should be prioritized for short test designs, while USCQM is preferable for long test designs.

Keywords: polytomous scoring, test design, structured simplest complete Q-matrix, unstructured simplest complete Q-matrix, algorithm
Classification Number: B841

1 Introduction

Cognitive diagnostic assessment provides diagnostic information about examinees' latent cognitive abilities (i.e., knowledge states, KS), facilitating individualized instruction and enhancing student competency (Leighton & Gierl, 2007; Tatsuoka, 2009). Against the backdrop of the "double reduction" policy, a practical challenge for test designers involves flexibly achieving precise diagnosis of students' cognitive abilities with minimal test length. In cognitive diagnostic testing, the Q-matrix, which characterizes the relationship between items and attributes, significantly influences classification precision (Chiu, 2013; De Carlo, 2011, 2012; Gross & George, 2014; Liu et al., 2012, 2013; Madison & Bradshaw, 2015). Through the Q-matrix, differential response patterns among examinees are induced, with greater differences facilitating more precise identification using cognitive diagnosis models. Consequently, the identifiability of cognitive diagnosis models based on Q-matrices has attracted widespread attention (Chiu et al., 2009; Chiu & Kohn 2015; Fang et al., 2019; Gu & Xu, 2019, 2021; Kohn & Chiu, 2017; Lin & Xu, 2023; Ouyang & Xu, 2022).

The completeness of Q-matrices constitutes a primary research focus within Q-matrix identifiability. A Q-matrix capable of identifying all examinees is termed a complete Q-matrix; otherwise, it is considered incomplete (Chiu et al., 2009; Chiu & Kohn 2015; Kohn & Chiu, 2017). Existing research demonstrates that incomplete Q-matrices misclassify some examinees into incorrect categories (Chiu et al., 2009). Therefore, Q-matrix completeness serves as a sufficient and/or necessary condition for identifiability and represents a critical condition for model identifiability (Gu & Xu, 2019, 2021; Lin & Xu, 2023; Ouyang & Xu, 2022). Based on whether item specifications align with attribute hierarchy structures, complete Q-matrices can be further categorized into structured complete Q-matrices and unstructured complete Q-matrices (Kohn & Chiu, 2021; Ding et al., 2022). In current research, structured complete Q-matrices primarily involve tests containing the reachability matrix, while unstructured complete Q-matrices (excluding independent structures) mainly consist of tests containing the identity matrix (Chiu et al., 2009; Fang et al., 2019; Lin & Xu, 2023; Ouyang & Xu, 2022; Xu, 2019; Xu & Zhang, 2016; Yu & Xu, 2021) or the novel method proposed by Kohn and Chiu (2021), who argued that for dichotomous scoring, adherence to a specific attribute hierarchy structure is not necessary for improving classification precision. They proposed and proved that unstructured R^* matrices satisfying $E \leq R^* < R$ constitute unstructured complete Q-matrices (where R and E represent the reachability matrix and identity matrix corresponding to given attributes and their hierarchical relationships, respectively).

The relationship among KS, ideal response patterns (IRP), and observed response patterns (ORP) forms the core of cognitive diagnosis (Ding et al., 2012). The IRP serves not only as the link between KS and ORP but also as the fundamental basis for their relationship. In the absence of item properties, motivational factors, or random effects, a complete Q-matrix can establish a one-to-one correspondence between IRPs and KSs. If KSs cannot be identified under these ideal conditions, they become even more difficult to identify from ORPs affected by the aforementioned factors. Previous research has extensively considered interfering factors while neglecting the core one-to-one correspondence between IRPs and KSs. Therefore, this paper's discussion of complete Q-matrices focuses on whether the matrix can establish such a one-to-one relationship between IRPs and KSs.

Compared to 0-1 scoring, polytomous scoring can more finely and deeply probe examinees' specific solution steps or processing procedures, thereby providing richer diagnostic information (Ma & de la Torre, 2016). In polytomous scoring tests, although the reachability matrix substantially improves classification accuracy (Ding et al., 2011; Ding et al., 2010), it does not represent the complete Q-matrix with the fewest items. For instance, Ding, Wang, et al. (2014) proposed that for linear structures with K attributes, a column of all 1s suffices to establish a one-to-one correspondence between IRPs and KSs, meaning the all-1 column constitutes the complete Q-matrix with the minimal number of items. This paper refers to structured and unstructured complete Q-matrices with the fewest items as structured simplest complete Q-matrices (SSCQM) and unstructured simplest complete Q-matrices (USCQM), respectively.

The simplest complete Q-matrix (SCQM) can be applied to short tests (e.g., classroom assessments), and long tests containing SCQMs as sub-matrices must be complete Q-matrices. Therefore, polytomous SCQMs hold broad application prospects in Q-matrix design.

The design of polytomous SCQMs is often more challenging than that of dichotomous complete Q-matrices. To the best of our knowledge, only Ding, Luo, et al. (2014) and Ding, Wang, et al. (2014) have addressed polytomous complete Q-matrix design with theoretical proofs. For different attribute hierarchy structures, they respectively provided design methods and proofs for several basic attribute hierarchy structures (linear, convergent, divergent, unstructured, and independent). While offering new approaches to improve polytomous cognitive diagnostic test precision and efficiency, these studies suffer from several limitations: (1) Their methods for designing complete Q-matrices across different attribute hierarchy structures lack compatibility and cannot be integrated. In practical applications, attribute hierarchy structures are more complex, typically composed of combinations of these basic structures, yet research on obtaining complete Q-matrices for such complex structures remains lacking; (2) The effectiveness of their methods lacks simulation verification. For example, regarding the all-1 column for linear structures, since this complete Q-matrix yields short tests with potentially large measurement errors, and considering that repeated measurement can reduce such errors, a critical question emerges: How many all-1 columns are needed to achieve measurement precision comparable to a K-order reachability matrix? This question must be addressed; (3) Research on polytomous unstructured complete Q-matrices is absent. While Kohn and Chiu (2021) attempted to extend structured complete Q-matrix research to unstructured Q-matrices and found a broader selection range for unstructured complete Q-matrices, their work was limited to dichotomous scoring and did not address polytomous scoring. Existing literature (Fang et al., 2019; Lin & Xu, 2023; Ouyang & Xu, 2022) proposes that polytomous complete Q-matrices contain identity matrices, but test designs containing only identity matrices are overly simplistic. Do other unstructured complete Q-matrices exist that enable test diversity? These issues require further discussion.

Building upon previous research, this paper aims to address two problems based on IRPs: First, to propose a general design method and algorithm for polytomous SSCQMs and USCQMs; Second, to examine how attribute hierarchy structures, number of attributes, and item parameters affect the classification accuracy of polytomous SSCQMs and USCQMs. All Q-matrices discussed below are at the item level, meaning one item corresponds to one vector.

2 Design and Algorithm for Polytomous SSCQM

Considering the complexity of design and proof in Ding, Luo, et al. (2014) and Ding, Wang, et al. (2014), and their failure to provide a unified design method applicable to various attribute hierarchy structures, this study adopts an alternative approach. First, guided by the principle that complete Q-matrices can establish a one-to-one correspondence between IRPs and KSs, we construct SSCQMs and validate them. Then, we propose an operational, user-friendly, and generalizable SSCQM algorithm applicable to all attribute hierarchy structures.

2.1 Design of Polytomous SSCQM

For dichotomous scoring, using the DINA ideal scoring rule (which requires non-compensatory attributes), employing the reachability matrix as a test submatrix can establish a one-to-one correspondence between examinees' IRPs and KSs (Ding et al., 2012; Ding et al., 2011; Ding et al., 2010; Kohn & Chiu, 2021), making the reachability matrix a complete Q-matrix.

For polytomous scoring, we adopt the scoring rule where an examinee's ideal score increases by one point for each attribute mastered in an item (Tatsuoka, 1995):

$$\eta_{ij} = \alpha_i \circ q_j = \sum_{k=1}^{K} \alpha_{ik} q_{jk}$$

where $\alpha_i = \{\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{iK}\}$, $Q = \{q_1, q_2, \dots, q_J\}$, and $q_j = \{q_{j1}, q_{j2}, \dots, q_{jK}\}$. The maximum score value is:

$$m_j = \sum_{k=1}^{K} q_{jk}$$

The reachability matrix remains a complete Q-matrix (Ding & Wang et al., 2014; Sun et al., 2013). However, except for independent structures, the reachability matrix for other attribute hierarchy structures is not the simplest complete Q-matrix. Therefore, using the reachability matrix as the research object, we compare the ideal response patterns of all KSs on the reachability matrix and delete all unnecessary columns to obtain a sub-matrix of the reachability matrix as the simplest complete Q-matrix. Below, we use a linear structure with six attributes as an example (see Figure 1), and based on formulas (1) and (2), construct and validate the SSCQM. Due to space limitations, we directly present the SSCQMs for the other four attribute hierarchy structures (convergent, divergent, unstructured, and independent).

(1) Linear Structure

There are 7 KS types, and the reachability matrix R has 6 columns. The ideal response patterns based on R are shown in Table 1:

Table 1 Ideal Responses of KS on R

$\alpha(\in KS)$	$(1,0,0,0,0,0)^T$	$(1,1,0,0,0,0)^T$	$(1,1,1,0,0,0)^T$	$(1,1,1,1,0,0)^T$	$(1,1,1,1,1,0)^T$	$(1,1,1,1,1,1)^T$
1. (000000)	0	0	0	0	0	0
2. (100000)	1	0	0	0	0	0
3. (110000)	1	1	0	0	0	0
4. (111000)	1	1	1	0	0	0
5. (111100)	1	1	1	1	0	0
6. (111110)	1	1	1	1	1	0
7. (111111)	1	1	1	1	1	1

From the last column of Table 1, we see that $Q_L = (1,1,1,1,1,1)^T$ in the reachability matrix enables a one-to-one correspondence between IRP and KS, meaning Q_L is the SSCQM.

(2) Convergent Structure

The SSCQM for convergent structure (a) is $Q_{C11} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$. The SSCQM for convergent structure (b) is $Q_{C21} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$, $Q_{C22} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$. The SSCQM for convergent structure (c) is $Q_{C31} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$, $Q_{C32} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$, $Q_{C33} =$

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Although convergent structures have multiple SSCQMs, each matrix selects columns corresponding to one or several different branches of the convergent structure. Structurally, these branches are parallel relationships. In test design, any of these SSCQMs can be chosen, or if multiple SSCQMs are needed for a test, these matrices can be mixed to minimize exposure rates.

(3) Divergent Structure

The SSCQM is $Q_D = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

(4) Unstructured

The SSCQM is $Q_U = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$.

(5) Independent Structure

The SSCQM is the reachability matrix.

From the SSCQMs of these five attribute hierarchy structures, we observe that the number of branches in the attribute hierarchy diagram equals the number of columns in the SSCQM, with each column vector corresponding to one branch. The SSCQMs for linear and convergent structures contain the maximum column of the reachability matrix (i.e., the all-1 column), while the columns in SSCQMs for the other three structures are incomparable with each other.

2.2 Algorithm for Polytomous SSCQM

Definition: For two K-dimensional vectors x and y, we say $x \leq y$ if and only if $\forall k \leq K, x_k \leq y_k$, where "$\leq$" denotes a partial order relation. If x and y do not satisfy this relationship, x and y are said to be incomparable.

Based on whether columns in the reachability matrix are comparable and their relative magnitudes, we present the following algorithm. The column obtained through comparison is termed the "pillar column." The algorithm for polytomous SSCQM is as follows:

Step 1: Input the reachability matrix R, calculate the column sum (the sum of attributes measured) for each column of R, and sort columns by column sum in ascending order. If column sums are equal, all permutations of these columns must be examined. Let $W = R = (r_1, r_2, \ldots, r_K)$, and generate two empty matrices $Q1$ and $Q2$.

Step 2: Loop through columns to find pillar columns in W:

```
for i = 1 to [col(W)-1] {
    // First, starting from the last column of W, if this column is greater than all other c
    // designate it as a pillar column and place it in Q1. If similar columns remain,
    // continue placing them in Q1 (during this process, W must have more than 2 columns;
    // otherwise, directly compare column magnitudes).

    // Next, if r_i < r_{i+1}, delete r_i from W, compare r_{i+1} with subsequent columns,
    // delete smaller ones until finding the maximum column, designate it as a pillar column
    // and place it in Q2.

    // If r_i and r_{i+1} are incomparable, skip to compare with the next column.
    // If r_i is incomparable with all subsequent columns, place r_i in Q2.

    // If only 1 column remains in W, place it directly in Q2.

    // Finally, recalculate col(W) until no columns remain in W.
}
```

Step 3: Find pillar columns in $Q1$; delete non-pillar columns.

Step 4: Calculate the number of branches in the attribute hierarchy diagram (denoted as n). If $Q1$ has m columns, randomly select $n - m$ columns from $Q2$ and merge them with $Q1$ to form a Q-matrix. If $Q1$ contains no columns, randomly select n columns from $Q2$ to form matrix Q.

Step 5: Verify the completeness of the obtained Q-matrix. If Q is complete, output Q; otherwise, discard it.

We illustrate the algorithm using convergent structure (c) as an example. **Step 1**: First, sort the columns of R, yielding $W = R_1 = (r_1, r_2, \ldots, r_6) = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$. Since columns 2, 3, and 4 each measure 2 attributes, we perform full permutations of these 3 columns, resulting in 6 cases corresponding to 6 matrices. We then extract structured complete Q-matrices from each of these 6 matrices. Starting with R_1 (i.e., W), we generate two empty matrices $Q1$ and $Q2$. **Step 2**: First, the largest column in W is column 6, which we place in $Q1$ and delete from W. No maximum column remains among the remaining columns. Next, we find pillar columns in the remaining W. Column 1 is smaller than column 2, so we delete column 1 and continue comparing column 2 with subsequent columns. Since they are incomparable, column 2 is designated a pillar column and placed in $Q2$. After deleting column 2 from W, only 3 columns remain. We reset the column indices of W, with the original columns 3, 4, and 5 becoming new columns 1, 2, and 3. No maximum column exists in the remaining columns. We then find pillar columns: column 1 in W is incomparable with column 2, so we skip column 2. Since column 1 is smaller than column 3, we place column 3 in $Q2$ and delete columns 1 and 3 from W. Finally, only column 2 remains in W, which is also placed in $Q2$. Thus, $Q2$ ends up with 3 columns: original columns 2, 4, and 5. **Step 3**: $Q1$ contains only 1 column, which is therefore the pillar column. **Step 4**: Since this is a convergent structure with 3 branches, we randomly select 2 columns from $Q2$ and merge them with the 1 column in $Q1$ to form 3 matrices: $Q_{31} = (r_4, r_5, r_6)$, $Q_{32} = (r_2, r_4, r_6)$, and $Q_{33} = (r_2, r_5, r_6)$. **Step 5**: Verification confirms that Q_{31}, Q_{32}, and Q_{33} are all SSCQMs. The SSCQMs obtained from the other 5 matrices are identical to those from R_1. This algorithm yields polytomous SSCQMs consistent with existing results (Ding & Luo, 2014; Ding & Wang, 2014).

Noteworthy points: (1) **Permutation issue in Step 1**: Columns with equal column sums must be fully permuted; otherwise, only partial SSCQMs can be obtained. (2) **Linear structure issue**: Except for columns 1 and 2, all other columns in the linear structure's reachability matrix are placed in $Q1$, while column 2 is placed in $Q2$. Since the pillar column in $Q1$ is the all-1 column and the attribute hierarchy diagram has only 1 branch, columns in $Q2$ are not counted, making this pillar column the solution. (3) **Branch counting in Step 4**: Generally, if the attribute hierarchy diagram has clear branches, the number of branches can be calculated directly. If branches are complex, the number need not be calculated; instead, directly select 1 column, 2 columns, etc., from $Q2$ to merge with $Q1$ and verify the completeness of the resulting Q-matrix.

3 Design and Algorithm for Polytomous USCQM

Building upon SSCQM design, this study further explores design methods and algorithms for USCQM.

3.1 Design of Polytomous USCQM

For dichotomous scoring, unstructured complete Q-matrices lie between the identity matrix and reachability matrix (Kohn & Chiu, 2021), where the reachability matrix represents a structured complete Q-matrix and the identity matrix represents an unstructured complete Q-matrix (except for independent structures). Drawing on this dichotomous approach, we use the SSCQM as the upper bound and a sub-matrix of the identity matrix as the lower bound for polytomous scoring. Let the reachability matrix $R = (r_1, r_2, \ldots, r_K)$ (with columns arranged in order of attributes measured). If the columns of the upper-bound SSCQM correspond to column indices $(j_1, j_2, \ldots, j_t)$ in R, then the lower bound consists of columns with the same indices $(j_1, j_2, \ldots, j_t)$ from identity matrix E to form a sub-matrix of the identity matrix. Using convergent structure (a) as an example, let reachability matrix $R = (r_1, r_2, \ldots, r_6)$, with SSCQM Q_{C11} corresponding to columns (5,6) in R, i.e., $Q_{C11} = (r_5, r_6) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$. Taking the identity matrix of the same order $E = (e_1, e_2, \ldots, e_6)$, we form the identity sub-matrix $E_{11} = (e_5, e_6) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}$. We then identify all matrices I_1 between E_{11} and Q_{C11}, i.e., $E_{11} \le I_1 < Q_{C11}$, with each row of I_1 containing at least one "1" (to ensure each attribute is measured). For example, $I_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$. Similarly, Q_{C12} corresponds to columns (4,6) of the reachability matrix, with identity sub-matrix columns (4,6) being $E_{12} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$, where $E_{12} \le I_2 < Q_{C12}$, and each row of I_2 contains at least one "1". For example, $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}$. Compared to the 2 SSCQMs, there are 53 eligible USCQMs I_1 and I_2.

3.2 Algorithm for Polytomous USCQM

Based on the above design principle, the polytomous USCQM algorithm is as follows:

Step 1: Input the structured simplest complete Q-matrix $Q = (r_{j1}, r_{j2}, \dots, r_{jt})$ and its corresponding identity sub-matrix $E' = (e_{j1}, e_{j2}, \dots, e_{jt})$. Compute $Q - E'$ to obtain matrix $Q_0 = (q_{ij})$.

Step 2: Randomly replace one or more $q_{ij} = 1$ in Q_0 with $q_{ij} = 0$.

Step 3: Add the matrix from Step 2 to the identity sub-matrix E' to obtain matrix Q'.

Step 4: Determine whether the Boolean union of columns in matrix Q' forms an all-1 column. If yes, output matrix Q'; otherwise, discard it.

4.1 Polytomous Cognitive Diagnosis Models

Most polytomous cognitive diagnosis models extend from dichotomous cognitive diagnosis models. Gao et al. (2021) detailed three types of polytomous cognitive diagnosis models: (1) adjacent category models; (2) continuation ratio models; and (3) cumulative probability models. Since completeness verification involves cognitive diagnosis models, we introduce only the GPDINA model (Chen & de la Torre, 2018) and RP-DINA model (Cai et al., 2017) relevant to this study.

The GPDINA model integrates the GDINA model (the generalized DINA model) (de la Torre, 2011) with the GPDM model (the general polytomous diagnosis model) (Chen & de la Torre, 2018):

$$P(Y_j = t|\alpha^*_{ij}) = P^*(Y_j \geq t|\alpha^*_{ij}) - P^*(Y_j \geq t+1|\alpha^*_{ij}) \tag{3}$$

$$P^*(Y_j \geq t|\alpha^*_{ij}) = \begin{cases} 1, & t = 0 \\ 0, & t = m_j \\ \mathcal{F}[P(Y_j = t|\alpha^*_{ij})] = \beta_{jto} + \sum \beta_{jtk}\alpha_{ijk} + \dots, & \alpha_{ij} \text{ pattern} \end{cases}$$

where $P^*(Y_j \geq t|\alpha^*_{ij})$ represents the probability of scoring greater than or equal to t points for $\alpha^*_{ij} = \{\alpha_{ij1}, \dots, \alpha_{ijK^*_j}\}$ based on item j's attribute reduction, $P(Y_j = t|\alpha^*_{ij})$ represents the probability of scoring exactly t points for α^*_{ij}, $\mathcal{F}[\cdot]$ is the link function, and $\beta_{jto}, \beta_{jtk}, \dots$ are intercept, main effect, and interaction parameters. Under the Logit link function, the GPDINA model corresponds to the PLCDM model; under the identity link function, it simplifies to PDINA, PACDM, PDINO models, etc.

PDINA Model:

$$P(Y_j = t|\alpha^*_{ij}) = \beta_{jto} + \beta_{jt12\dots K^*_{jt}} \prod \alpha_{ik} \tag{6}$$

That is:

$$P(Y_{ij} = t|\alpha^*_{ij}) = P^*(Y_{ij} > t|\alpha^*_{ij}) - P^*(Y_{ij} > t + 1|\alpha^*_{ij}) \quad (7)$$

where $P^*(Y_{ij} = t|\alpha^*_{ij}) = (1 - s_{jt})^{\eta_{ij}} g_{jt}^{1-\eta_{ijt}}$, with $\eta_{ij} = \prod \alpha_{ik}$. The relationship between the two models' parameters is: $\beta_{jto} = g_{jt}$, $\beta_{jt12\ldots K^*_{jt}} = 1 - g_{jt} - s_{jt}$.

Since PDINA constructs ideal scores using only 0 and 1, affecting classification precision, Cai et al. (2017) modified PDINA's scoring method and proposed the RP-DINA model.

RPDINA Model:

$$P^*(Y_{ij} = t|\alpha_i) = (1 - s_{jt})^{\delta_{ijt}} g_{jt}^{1-\delta_{ijt}} \quad (8)$$

where $\delta_{ijt} = \begin{cases} 1, & \text{if } \eta_{ij} \geq t \\ 0, & \text{if } \eta_{ij} < t \end{cases}$, and $\eta_{ij} = f_{ix}[\cdots \times m_j]$, with f_{ix} being the integer function and m_j the maximum score for item j. Other symbols in the RPDINA model are defined consistently with the PDINA model. Research shows that due to the scoring modification in the RPDINA model, its classification accuracy is higher than that of the PDINA model.

PACDM Model:

$$P(Y_j = t|\alpha^*_{ij}) = \beta_{jto} + \sum \beta_{jtk}\alpha_{ik} \quad (10)$$

PDINO Model:

$$P(Y_j = t|\alpha^*_{ij}) = \beta_{jto} + \beta_{jtk}\alpha_{ik} \quad (11)$$

4.2 Validation of SCQM Combined with Polytomous Cognitive Diagnosis Models

Furthermore, based on the one-to-one correspondence between ideal response patterns and knowledge states, we validate whether the constructed SCQM maintains its completeness when combined with cognitive diagnosis models.

(1) Scoring Method Specification: In polytomous scoring, the completeness of the Q-matrix depends on the scoring method and maximum score value, which affect the quality of diagnostic information from the Q-matrix (Ding & Wang et al., 2014). If using the scoring method and maximum score proposed by Cai et al. (2017), the reachability matrix is not a complete Q-matrix. Since complete Q-matrices can improve diagnostic classification accuracy, the cognitive diagnosis models in our validation adopt the scoring method in formula (1) and maximum score method in formula (2).

(2) Q-matrix Completeness Verification: Kohn and Chiu (2021) proposed that when expected observed response patterns are equal, the corresponding KSs are identical. The expectation for examinee α on item q_j is:

$$S_j(\alpha) = E(Y_j|\alpha) = \sum tP(Y_j = t|\alpha) \tag{12}$$

If knowledge states α and α' have equal expected observed response pattern vectors, then $\alpha = \alpha'$, i.e., $S(\alpha) = S(\alpha') \Rightarrow \alpha = \alpha'$, making Q a complete Q-matrix, where $S(\alpha) = \{S_1(\alpha), S_2(\alpha), \dots, S_J(\alpha)\}$ and $S(\alpha') = \{S_1(\alpha'), S_2(\alpha'), \dots, S_J(\alpha')\}$.

We validate the completeness of the SCQM proposed in this paper by combining it with cognitive diagnosis models. Specifically, we select the linear Q_L from Section 2.1 as the SSCQM and convergent I_1 from Section 3.1 as the USCQM.

For the PDINA model, since ideal scoring still follows a dichotomous approach where interaction exists only when examinees master all attributes in an item, examinees who fail to master at least one attribute cannot be identified. Therefore, Q_L and I_1 are not complete Q-matrices. The SSCQM for PDINA is the reachability matrix, and there are multiple USCQMs, one of which is the identity matrix.

For the RPDINA model, the results are shown in the tables below.

Table 2 $S(\alpha)$ for linear structure α on Q_L

$\alpha(\in KS)$	$S(\alpha)$
(0,0,0,0,0,0)	$1\times(g_1-g_2)+2\times(g_2-g_3)+3\times(g_3-g_4)+4\times(g_4-g_5)+5\times(g_5-g_6)+6\times g_6$
(1,0,0,0,0,0)	$1\times(1-s_1-g_2)+2\times(g_2-g_3)+3\times(g_3-g_4)+4\times(g_4-g_5)+5\times(g_5-g_6)+6\times g_6$
(1,1,0,0,0,0)	$1\times(s_2-s_1)+2\times(1-s_2-g_3)+3\times(g_3-g_4)+4\times(g_4-g_5)+5\times(g_5-g_6)+6\times g_6$
(1,1,1,0,0,0)	$1\times(s_2-s_1)+2\times(s_3-s_2)+3\times(1-s_3-g_4)+4\times(g_4-g_5)+5\times(g_5-g_6)+6\times g_6$
(1,1,1,1,0,0)	$1\times(s_2-s_1)+2\times(s_3-s_2)+3\times(s_4-s_3)+4\times(1-s_4-g_5)+5\times(g_5-g_6)+6\times g_6$
(1,1,1,1,1,0)	$1\times(s_2-s_1)+2\times(s_3-s_2)+3\times(s_4-s_3)+4\times(s_5-s_4)+5\times(1-s_5-g_6)+6\times g_6$
(1,1,1,1,1,1)	$1\times(s_2-s_1)+2\times(s_3-s_2)+3\times(s_4-s_3)+4\times(s_5-s_4)+5\times(s_6-s_5)+6\times(1-s_6)$

Table 2 shows that different $S(\alpha)$ correspond to different α, indicating that with the scoring method provided in this paper, Q_L is a structured complete Q-matrix for the RPDINA model. Since it contains only 1 column, Q_L is an SSCQM.

Table 3 $S(\alpha)$ for convergent structure (a) α on I_1

$\alpha(\in KS)$	$q_1=(101010)'$ $S_1(\alpha)$	$q_2=(010101)'$ $S_2(\alpha)$
1. (0,0,0,0,0,0)	$1\times(g_{21}-g_{22})+2\times(g_{22}-g_{23})+3\times g_{23}$	$1\times(g_{11}-g_{12})+2\times(g_{12}-g_{13})+3\times g_{13}$
2. (1,0,0,0,0,0)	$1\times(g_{21}-g_{22})+2\times(g_{22}-g_{23})+3\times g_{23}$	$1\times(1-s_{11}-g_{12})+2\times(g_{12}-g_{13})+3\times g_{13}$
3. (1,1,0,0,0,0)	$1\times(1-s_{21}-g_{22})+2\times(g_{22}-g_{23})+3\times g_{23}$	$1\times(1-s_{11}-g_{12})+2\times(g_{12}-g_{13})+3\times g_{13}$
4. (1,1,1,0,0,0)	$1\times(1-s_{21}-g_{22})+2\times(g_{22}-g_{23})+3\times g_{23}$	$1\times(s_{12}-s_{11})+2\times(1-s_{12}-g_{13})+3\times g_{13}$
5. (1,1,1,1,0,0)	$1\times(s_{12}-s_{11})+2\times(1-s_{12}-g_{13})+3\times g_{13}$	$1\times(s_{12}-s_{11})+2\times(1-s_{12}-g_{13})+3\times g_{13}$
6. (1,1,1,0,1,0)	$1\times(s_{22}-s_{21})+2\times(1-s_{22}-g_{23})+3\times g_{23}$	$1\times(s_{12}-s_{11})+2\times(1-s_{12}-g_{13})+3\times g_{13}$
7. (1,1,1,1,1,0)	$1\times(s_{12}-s_{11})+2\times(s_{13}-s_{12})+3\times(1-s_{13})$	$1\times(1-s_{21}-g_{22})+2\times(g_{22}-g_{23})+3\times g_{23}$
8. (1,1,1,1,1,1)	$1\times(s_{12}-s_{11})+2\times(s_{13}-s_{12})+3\times(1-s_{13})$	$1\times(s_{22}-s_{21})+2\times(s_{23}-s_{22})+3\times(1-s_{23})$

Based on Table 3's results, since each α has a distinct $S(\alpha)$ on I_1, and deleting any item would make at least two α values have equal $S(\alpha)$, I_1 is a USCQM. Similarly, we can prove that I_2 is also a USCQM.

For the PACDM model, Q_L and I_1 remain complete matrices. For the PDINO model, Q_L and I_1 are not complete matrices. The examples suggest that as long as main effects exist in the cognitive diagnosis model, the completeness of SCQM remains unchanged.

5 Simulation Study

From the IRP perspective, a single SCQM can completely distinguish KSs, just like the reachability matrix. However, in practical applications, since a single SCQM has fewer columns, measurement errors may result in classification accuracy (PMR or MMR) that is not necessarily higher (and may even be lower) than that of the reachability matrix. To investigate the classification effectiveness of polytomous SSCQMs and USCQMs, simulation experiments primarily examine how many SSCQMs/USCQMs are needed to achieve classification precision equivalent to the reachability matrix when the total number of columns does not exceed that of the reachability matrix. To avoid conceptual confusion, structured complete Q-matrices in the following discussion do not include the reachability matrix.

To better correspond to summative and formative assessments in practice, this study constructs both long and short tests, examining the classification effects of three matrix types (containing SSCQM, USCQM, and reachability matrix) across attribute hierarchy structures, item parameters, and number of attributes. All studies use Python programs for simulation and analysis.

5.1 Study 1: Examining the Effects of Attribute Hierarchy Structure and Item Parameters

5.1.1 Experimental Design Number of attributes $K = 6$, number of examinees $N = 2000$. Since the SSCQM for independent structures is the reachability matrix itself and no unstructured Q-matrix exists, this study only involves linear, convergent (a)(b)(c), divergent, and unstructured attribute hierarchy structures. The research proceeds along two dimensions: First, comparing classification accuracy between tests containing one or multiple SSCQMs/USCQMs (with total columns not exceeding the reachability matrix) and tests containing the reachability matrix; Second, based on multiple SSCQMs/USCQMs, adding columns from SSCQM/USCQM to match the reachability matrix's column count, then comparing their diagnostic classification effects.

5.1.2 Monte Carlo Simulation **(1) Simulation of True Attribute Mastery Patterns**: Examinees are evenly distributed across KSs for each attribute hierarchy structure; any remainder is randomly assigned.

(2) Simulation of Test Q-matrix and Item Parameters: Long tests contain 40 items. Tests respectively include 1 reachability matrix, 1 or multiple SSCQMs and USCQMs of different attribute hierarchy structures (with total columns not exceeding the reachability matrix), with remaining items selected from structured/unstructured incomplete Q-matrices.

Short tests contain 6 items (equal to the reachability matrix's column count). Tests consist of 1 reachability matrix, the maximum number of SSCQMs and USCQMs (with remaining items selected from SSCQM/USCQM to match the reachability matrix's column count).

For item j, parameters $s_j = (s_{j1}, s_{j2}, \dots, s_{jm_j})$ and $g_j = (g_{j1}, g_{j2}, \dots, g_{jm_j})$ satisfy $s_{j1} < s_{j2} < \cdots < s_{jm_j}$ and $g_{j1} > g_{j2} > \cdots > g_{jm_j}$. For long tests, parameters s_{jt}, g_{jt} $(0 \le t \le m_j)$ follow three distributions: $U(0, 0.15)$, $U(0, 0.25)$, and $U(0, 0.35)$. Since short tests have fewer items, they primarily use medium-to-high quality items, with parameters s_{jt}, g_{jt} following $U(0, 0.05)$, $U(0, 0.15)$, and $U(0, 0.25)$.

(3) Simulation of Examinee Response Patterns: Based on examinees' true values and the test Q-matrix, we obtain the ideal score matrix using formulas (1) and (2), then simulate examinee scores according to the RPDINA model.

(4) Estimation of Examinee KS: We use the RP-DINA model and Maximum A Posteriori (MAP) estimation to estimate examinees' KSs.

(5) Evaluation Metrics: We employ Pattern Match Ratio (PMR) and Marginal Match Ratio (MMR) as evaluation metrics for classification ability.

$$PMR = \frac{\sum N_{i-correct}}{N}$$

$$MMR = \frac{\sum N_{ik-correct}}{N \times K}$$

where N is the total number of examinees, $N_{i-correct}$ indicates whether examinee i's attribute mastery pattern is correctly classified (1 if correct, 0 otherwise); K is the number of attributes, and $N_{ik-correct}$ indicates whether examinee i's attribute k is correctly classified (1 if correct, 0 otherwise).

Simulation experiments are repeated 100 times (since many USCQMs are available, they are randomly selected), and the average values of PMR and MMR are computed.

5.1.3 Research Results (1) Long Test Results

Based on three different item parameter settings, Tables 4-6 present the mean PMR and MMR for three test types (containing 1 or multiple SSCQMs, USCQMs, and reachability matrix) across linear, convergent, divergent, and unstructured attribute hierarchy structures.

Table 4 shows simulation results for 6 attributes, 2000 examinees, and s, g following $U(0, 0.15)$ for long tests: (a) Following the order of attribute hierarchy structures, classification accuracy decreases for all three test types, with PMR decreasing by less than 3% and MMR decreasing by no more than 0.5%. (b) For each attribute hierarchy structure, more SSCQMs/USCQMs yield higher classification accuracy, with MMR above 0.99 and PMR above 0.95. (c) When multiple SSCQMs/USCQMs have fewer columns than the reachability matrix (e.g., linear structure with 3 columns, convergent structures (a)(b) with 4 columns, convergent structure (c) with 3 columns, divergent structure with 3 columns, and unstructured with 5 columns), their classification accuracy exceeds that of tests containing the reachability matrix (6 columns). (d) When column counts equal the reachability matrix, following the order of attribute hierarchy structures, SSCQM tests show gradually lower classification accuracy than USCQM tests, with reachability matrix tests showing the lowest accuracy, though differences among the three are minimal (PMR differences ≤ 0.05, MMR differences even smaller). Tables 5-6 indicate similar conclusions under different s, g parameter levels.

Table 4 Classification Accuracy of Three Long Tests (6 Attributes, $s, g \sim U(0, 0.15)$)

Attribute Hierarchy Structure	Multiplicity of Simplest Complete Q Columns (Total Columns)	1+ SSCQM(s)	1+ USCQM(s)	Reachability Matrix
Linear	1+(6)	0.9611/0.3318	0.9611/0.3318	0.9845/1.0389
Convergent (a)	2+(6)	0.9823/0.6636	0.9823/0.6636	0.9845/1.0389
Convergent (b)	2+(6)	0.9823/0.6636	0.9823/0.6636	0.9845/1.0389
Convergent (c)	3+(6)	0.9845/0.9954	0.9845/0.9954	0.9845/1.0389
Divergent	3+(6)	0.9845/0.9954	0.9845/0.9954	0.9845/1.0389
Unstructured	5+(6)	0.9845/1.0389	0.9845/1.0389	0.9845/1.0389

Tables 4-6 also show that for the same attribute hierarchy structure, as s, g parameter levels increase, PMR decreases by 2%-6% and MMR decreases by 0.3%-1%; under the same parameters, following the order of attribute hierarchy structures, classification accuracy decreases by 1%-5% for PMR and 0.2%-1% for MMR. Overall results indicate that attribute hierarchy structure and item parameters have similar effects on the classification accuracy of polytomous SSCQMs and USCQMs.

(2) Short Test Results

Based on three different item parameter settings and following the order of linear, convergent, divergent, and unstructured attribute hierarchy structures, Figure 2 presents the mean PMR for three short test types (multiple SSCQMs, multiple USCQMs with total columns equal to the reachability matrix, and the reachability matrix).

Figure 2 reveals: (a) As attribute hierarchy structures change and parameters increase, classification accuracy decreases for all three test types, with differences gradually increasing, particularly when the attribute hierarchy structure is unstructured and parameter values are 0.25, where the reachability matrix's accuracy differs most from other tests. (b) Except for unstructured, when s and g are identical and attribute hierarchy structures are the same, differences among the three test types' accuracy do not exceed 0.1. (c) In most cases, multiple SSCQMs achieve higher accuracy than the reachability matrix, though for convergent structure (c) and unstructured with larger parameters, the opposite occurs. (d) Multiple USCQMs almost always show the lowest accuracy. In fact, multiple SSCQMs and the reachability matrix are structured complete Q-matrices, while multiple USCQMs are unstructured complete Q-matrices. Overall, for short tests with equal column counts, structured complete Q-matrices yield higher classification accuracy than unstructured complete Q-matrices.

5.2 Study 2: Examining the Effect of Attribute Count Across Different Hierarchy Structures

Generally, classification accuracy decreases as the number of attributes increases. The extent to which attribute count affects SCQM is the focus of Study 2.

5.2.1 Monte Carlo Simulation

(1) Building on Study 1's simulation conditions, we examine four attribute hierarchy structures with 5-8 attributes.

(2) Item selection follows Study 1. Test parameters are fixed: long test parameters s_{jt} and g_{jt} follow $U(0, 0.35)$; short test parameters s_{jt} and g_{jt} follow $U(0, 0.05)$.

(3) True examinee values, scoring, and evaluation metrics follow Study 1.

5.2.2 Research Results (1) Long Test Results

Tables 7-9 present simulation results for 2000 examinees with $s_{jt}, g_{jt} \sim U(0, 0.35)$ across 5-8 attributes: (a) Overall, as attribute count increases and attribute hierarchy structures change, classification accuracy decreases for all three test types, but PMR remains above 0.88 and MMR above 0.97. (b) With fixed attribute count, following the order of attribute hierarchy structures, PMR decreases by 2%-8% and MMR by 0.5%-2%; with fixed attribute hierarchy structure, as attribute count increases, PMR decreases by 0.1%-5% and MMR by 0.1%-1%. (c) With equal column counts, the three test types show comparable accuracy with minimal differences (PMR differences between 0.001-0.05, MMR differences even smaller), with reachability matrix tests performing worst. Tests with 5-7 attributes perform better with SSCQMs, while 8-attribute tests perform better with USCQMs. (d) For some tests with multiple SSCQMs, when column counts reach certain thresholds (even if fewer than the reachability matrix's columns), their accuracy exceeds that of the reachability matrix. Similar patterns occur for tests with multiple USCQMs.

Tables 7-9 demonstrate that increasing attribute count affects accuracy similarly to changes in attribute hierarchy structure.

(2) Short Test Results

Based on different attribute counts and hierarchy structures, Figure 3 presents mean PMR for three short test types (multiple SSCQMs, multiple USCQMs, and reachability matrix) with column counts equal to the reachability matrix. With good item quality, as attribute count increases, multiple SSCQMs achieve the highest accuracy, while the other two matrices show varying relative performance, though all differences do not exceed 0.05.

Based on long and short test simulation results: The three test types show minimal accuracy differences (around 0.05). For long tests, following the order of attribute hierarchy structures (from linear to unstructured), when attribute count is 5-7 or item parameters are small, tests with multiple SSCQMs outperform those with multiple USCQMs; when attribute count is 8 or item parameters increase, the opposite holds. For short tests, tests with multiple SSCQMs outperform those with multiple USCQMs.

6 Empirical Data Analysis

To further validate SCQM performance, we analyze empirical data from Kang (2013). The data were collected from 8 primary schools in Jinhua and Wenzhou, Zhejiang Province, comprising 1300 examinees with 1240 valid cases. The test content involved fifth-grade travel word problems, with 17 items (including the reachability matrix) covering 8 attributes (see Table 10): basic arithmetic operations, general travel problem formulas, level complexity, complex travel problem formulas, identifying implicit conditions, relationship representation, schema representation, and formal algebraic strategies. The attribute hierarchy relationship is shown in Figure 4, representing a composite of linear, convergent, and unstructured components.

Figure 4 Cognitive Attribute Hierarchy for Primary School Travel Word Problems

Using the algorithm, we obtain the SSCQM: $Q = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

This study examines four tests: the original test containing the reachability matrix (denoted R) and three tests containing 1 SSCQM with different item counts (denoted $Q1, Q2, Q3$, none containing the reachability matrix), with item counts of 14, 15, and 16, respectively. Based on the attribute hierarchy structure, the expansion algorithm yields 39 knowledge states. Data analysis shows that tests $Q1, Q2, Q3$, and R can identify 24, 25, 21, and 22 knowledge state types, respectively. Tests $Q1, Q2$, and $Q3$ share 1150, 1124, and 1175 identifications with test R, respectively, accounting for 93%, 90%, and 95% of the total sample. Thus, SSCQM-based examinee identification shows over 90% overlap with reachability matrix-based identification, using fewer items. Tests $Q1$ and $Q2$ identify more knowledge state types than test R, enabling finer classification. The three tests' estimated attribute mastery proportions differ from test R's by no more than 7% (see Table 11).

Table 10 Test Items and Measured Attributes

Item	Attributes
...	...

Table 11 Differences in Attribute Mastery Proportions Between Tests Q1, Q2, Q3 and Test R (%)

Attribute	Q1-R	Q2-R	Q3-R
...	...	...	...

7 Discussion and Conclusions

7.1 Discussion

7.1.1 Relationship Between Polytomous Scoring Methods and Complete Q-Matrices Generally, for polytomous scoring, different scoring methods and maximum score values yield different complete Q-matrices. This study proposes a polytomous SCQM design method based on scoring method (1) and maximum score formula (2). Under this scoring method and maximum score, the reachability matrix is a complete Q-matrix. If the scoring method and maximum score change, the reachability matrix may no longer be complete. For example, Cai et al. (2017) proposed reducing maximum scores and using integer-based scoring, which makes some response patterns that would differ under formulas (1) and (2) become identical, thus altering the reachability matrix's completeness. Using a 4-attribute linear structure as an example, with reachability matrix items $q_1 = (1000)^T$, $q_2 = (1100)^T$, $q_3 = (1110)^T$, and $q_4 = (1111)^T$, and item scores set at 0.5, 1, 2, and 3 points respectively, examinees $\alpha_1 = (0000)$ and $\alpha_2 = (1000)$ have identical ideal response patterns (0000) on the reachability matrix under this algorithm, making them indistinguishable. Under formulas (1) and (2), the reachability matrix remains complete. In fact, if each attribute has equal scoring weight, when each item's maximum score is greater than or equal to the total number of attributes measured, the reachability matrix remains complete, and the completeness of SCQMs extracted from it is preserved. This favorable property expands SCQM's applicability. If attribute scoring weights differ, this conclusion may not hold.

7.1.2 Test Input-Output Ratio Tests incur time and economic costs. Therefore, improving test efficiency and maximizing examinee differentiation with minimal items represents a persistent goal in test design. The input-output ratio is defined as classification accuracy (MMR and PMR) divided by the number of items and estimation time, with larger values indicating higher accuracy per unit time and per item—i.e., higher input-output ratio. This metric is particularly evident in short tests. Across four attribute hierarchy structures, SSCQM/USCQM input-output ratios exceed those of the reachability matrix. The linear structure's simplest complete Q-matrix uses the fewest items, thus requiring less time for both item development and examinee estimation, yielding the highest input-output ratio. For example, in short tests with 6 attributes, the linear structure SSCQM (only 1 column) has an input-output ratio of $0.9611/0.3318 \times 1 = 2.90$, while the reachability matrix's ratio is $0.9845/1.0389 \times 6 = 0.16$. Convergent structure SSCQM/USCQM input-output ratios are 0.93 and 1.08, respectively, compared to approximately 0.17 for the reachability matrix. SSCQMs/USCQMs are highly efficient tests. However, SSCQM quantities are limited, so USCQMs must be considered.

Note: Figure translations are in progress. See original paper for figures.