INTERNATIONAL JOURNAL OF MODERN EDUCATION STUDIES
| Volume 3 - No 1 - June 2019 - ISSN 2618-6209 |

Research Article

An Application of Cognitive Diagnosis Modeling in TIMSS: A Comparison of Intuitive Definitions of Q-Matrices

Derya EVRAN[1]

International Journal of Modern Education Studies

June, 2019
Volume 3, No 1
Pages: 4-17
http://www.ijonmes.net
dergipark.gov.tr/ijonmes

Article Info:

Received : 01.03.2019
Revision 1 : 10.03.2019
Accepted : 11.03.2019
Online First Published : 12.03.2019
Published : 30.06.2019

Abstract:

Detection of students' ability levels is one of the common aims in educational studies. Cognitive Diagnosis Modeling approach has been used recently for the purpose of ability level detection by defined Q-matrices. To evaluate students' strengths and weaknesses, determine their mastery skills, and design instructions and interventions in learning process, Cognitive Diagnosis Modeling approach can be helpful. Cognitive Diagnosis Modeling is an alternative approach to Item Response Theory, and provides more information using multiple fine-grained skills in problem solving process rather than order students on a latent proficiency continuum This paper aims to use Cognitive Diagnosis Modeling (CDM) in order to investigate the definition of a Q-matrix across the cognitive skills of different years and countries in Trends in International Mathematics and Science Study (TIMSS). There is a subjective way in defining Q-matrices, an intuitive definition of Q-matrices, for this purpose, an application of building Q-matrices under specific Cognitive Diagnosis Models, from a set of expert proposed attributes is examined. The proposed attributes are used to build Q-matrices for TIMSS mathematics questions across its cycles, and across different nations.

Keywords: CDM, Q-matrix definition, cognitive assessment, TIMSS, attributes

Citation:
Evran, D. (2019). An application of cognitive diagnosis modeling in TIMSS: A Comparison of intuitive definitions of Q-Matrices. *International Journal of Modern Education Studies, 3*(1), 4-17.

[1] Harran University, deryaevran@harran.edu.tr, Orcid: 0000-0002-6960-342X

INTRODUCTION

Nowadays International assessments have become more important in recent decades as a consequence of globalization. To compare and evaluate the quality of education across countries, many cross-national student assessments have been conducted to compare learning outcomes across countries (Hambleton, 2005). One of the most prominent exams is the Trends in International Mathematics and Science Study (TIMSS) sponsored by the Institute of Educational Sciences (AFT, 1999; Olson, Martin, & Mullis, 2008; Mullis, Martin, Foy, & Arora, 2012). TIMSS has been conducted every four years since 1995. For its last two cycles, technical reports provided average tests scores and percentages to show achievement gaps using classical test theory (CTT), and item response theory (IRT) to estimate achievements on specific cognitive domains (Olson et al., 2008; Mullis et al, 2012).

In mathematics education, in addition to the comparison of achievement scores, students' mastery level of specific attributes has been evaluated by using the TIMSS (Lee, Park, Taylan, 2011). The Rule Space Model (RSM; Tatsuoka, 1983) applied to TIMSS 1999 mathematics items, and 30 attributes are defined to show the gap across gender and ethnicities (Birenbaum, Tatsuoka, & Yamada, 2004; Birenbaum, Nasser, & Tatsuoka, 2007). Similarly, Chen, Thompson, Gorin, and Tatsuoka (2008), Dogan and Tatsuoka (2008) also used the RSM approach in TIMSS 1999 mathematics items to report the classification rates of the students by gender and ethnicity. Furthermore, Lee, Park, and Taylan (2011) investigated TIMSS 2007 mathematics items using a cognitive diagnosis model (CDM), specifically the deterministic, inputs, noisy, "and" gate (DINA; Haertel, 1989; Junker & Sijtsma, 2001) model to explain the reasons for achievement gaps. In this study, 15 attributes were defined and mastery proportions were given for each attribute of the United States data.

To compare TIMSS scores across years and across nations by CDM models, an attribute pattern needs to be built for each assessment. The aims of this paper are (1) to fit the DINA and generalized DINA (G-DINA; de la Torre, 2011) models to TIMSS 2011 fourth grade mathematics items using the attributes that are proposed for TIMSS 2007 by Lee, Park, and Taylan (2011), and (2) to compare different Q-matrices and validate the most appropriate Q-matrix for the TIMSS 2011 data based on the G-DINA model fit results.

Background

To evaluate students' strengths and weaknesses, determine their mastery skills, and design instructions and interventions in learning process, CDM approach can be helpful. CDM is an alternative approach to IRT, and provides more information using multiple fine-grained skills in problem solving process rather than order students on a latent proficiency continuum (de la Torre, 2008). CDM is a latent variable model where skills are defined as attributes, and represented by the binary vector α to assess student mastery and non-mastery of the skills (de la Torre, 2011). Specific CDMs that are included in this paper are

DINA, and G-DINA. Both DINA and G-DINA models require a J x K Q-matrix; where J is the test length, K is the number of attributes (Tatsuoka, 1983). The element in row j and column k of the Q-matrix, qjk, is equal to 1 if the kth attribute is required to answer item j correctly; otherwise it is equal to 0 (de la Torre, 2011). Q-matrix is a cognitive design matrix that identifies the cognitive specification for each item (de la Torre, 2008).

The DINA Model

Let X_{ij} be the response of examinee i (i= 1, …, I) to item j (j= 1, …, J), and $\boldsymbol{\alpha}_i=\{\alpha_{ik}\}$ (k= 1, …, K) be the examinee's binary skills vector, where a 1 on the k^{th} element denotes the presence or mastery of skill k, and 0, the absence or non-mastery of the skill. The gate part of the DINA model creates two latent groups by comparing the examinee's skills vector and the Q-matrix (de la Torre, 2008).

$$\eta_{ij} = \prod_{k=1}^{K} \alpha_{ik}^{q_{jk}} , \qquad (1)$$

where $\eta_i=\{\eta_{ij}\}$ is a latent response vector where 1 indicates that examinee i possesses all the skills required for answering correctly item j, and 0 indicates that the examinee lacks at least one of the required skills. Therefore, the DINA model creates two groups for each item, one with the examinees who mastered all required attributes, and another with the examinees who lack at least one of the required attributes. Furthermore, the DINA model has slip and guessing parameters, which introduce the noise into the model. The slip parameter, $s_j = P\ (X_{ij} = 0|\eta_{ij} = 1)$, is defined as the probability that examinees who possess all the required skills for an item can slip and miss the item, while the guessing parameter, $g_j = P\ (X_{ij} = 1|\eta_{ij} = 0)$, is the probability that examinees who lack at least one of the required skills can guess and answer the item correctly. The probability of answering item j correctly by examinee i with the skills vector $\boldsymbol{\alpha}_i$ under the DINA model is given by,

$$P_j(\boldsymbol{\alpha}_i) = P\left(X_{ij} = 1\middle|\boldsymbol{\alpha}_i\right) = g_j^{1-\eta_{ij}}(1 - s_j)^{\eta_{ij}} , \qquad (2)$$

where answering an item correctly means that an examinee possessing all the necessary attributes, must not slip; or an examinee lacking at least one of the required skills must guess correctly (de la Torre, 2008).

The G-DINA Model

In the G-DINA model, the gate part creates $2^{K_j^*}$ latent groups by comparing the examinee's skills vector with the Q-matrix (de la Torre, 2011).

$$K_j^* = \sum_{k=1}^{K} q_{jk} , \qquad (3)$$

where K_j^* represents the number of required attributes for item j. The item response function of the G-DINA model is given by,

$$P(\boldsymbol{\alpha}_{lj}^{*}) = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk}\alpha_{lk} + \sum_{k'=k+1}^{K_j^*} \sum_{k=1}^{K_j^*-1} \delta_{jkk'}\alpha_{lk}\alpha_{lk'} \ldots + \delta_{j12..K_j^*} \prod_{k=1}^{K_j^*} \alpha_{lk}\,, \quad (4)$$

where
δ_{j0} is the intercept for item j;
δ_{jk} is the main effect due to α_k;
$\delta_{jkk'}$ is the interaction effect due α_k and α_{k}; and
$\delta_{j12\ldots}\, K_j^*$ is the interaction effect due to $\alpha_1, \ldots, \alpha\, K_j^*$,
where δ_0 represents the probability of a correct response when an examinee possesses none of the required attributes; δ_{ik} represents the change in the probability of a correct response when an examinee possesses a single attribute α_k; $\delta_{jkk'}$ is a first-order interaction effect which means the change in the probability of a correct response due to the mastery of both α_k and $\alpha_{k'}$ and $\delta_{j12\ldots}\, K_j^*$ is the change in the probability of a correct response due to the mastery of all the required attributes (de la Torre, 2011).

The G-DINA is a saturated model, and by applying constraints to the different link functions, specific reduced models can be obtained. For example the DINA model, the item response function is given,

$$P(\alpha_{lj}^{*}) = \begin{cases} g_j & \text{if} \quad \alpha_{lj}^{*} < \mathbf{1}_{K_j^*} \\ 1 - s_j & \text{otherwise,} \end{cases} \quad (6)$$

where $\mathbf{1}_{K_j^*}{}^*$ is a vector of ones and of length K_j^* (de la Torre, 2011). In the DINA model, except δ_{j0} and $\delta_{j12\ldots}\, K_j^*$ all parameters will be set to 0.

METHOD

Data

Data were taken from booklets 2 and 3 of TIMSS 2011 fourth grade mathematics assessment, which consist of 26 items with 12 multiple-choice items and 14 constructed response items. The data are recoded similar to Lee, Park, and Taylan (2011), as in constructed response items with polytomous responses were dichotomized as incorrect if they are wrong, partially true, unreached or omitted; or as correct if they are fully true.

Table 1
Attributes of TIMSS 2007 for fourth grade mathematics

Content Domain	**Attributes**
Number (N)	*Whole Numbers (4)* 1. Representing, comparing, and ordering whole numbers as well as demonstrating knowledge of place value. 2. Recognize multiples, computing with whole numbers using the four operations, and estimating computations. 3. Solve problems, including those set in real life contexts (for example, measurement and money problems). 4. Solve problems involving proportions. *Fractions and Decimals (2)* 5. Recognize, represent, and understand fractions and decimals as parts of a whole and their equivalents. 6. Solve problems involving simple fractions and decimals including their addition and subtraction. *Number Sentences with Whole Numbers (1)* 7. Find the missing number or operation and model simple situations in number sentence or expressions. *Patterns and Relationships (1)* 8. Describe relationships in patterns and their extensions; generate pairs of whole numbers by a given rule and identify a rule for every relationship given pairs of whole numbers.
Geometric Shapes & Measurement (GM)	*Lines and Angles (1)* 9. Measure, estimate, and understand properties of lines and angles and be able to draw them. *Two- and Three-dimensional Shapes (2)* 10. Classify, compare, and recognize geometric figures and shapes and their relationships and elementary properties. 11. Calculate and estimate perimeters, area, and volume.
Data & Display (DD)	*Location and Movement (1)* 12. Locate points in a coordinate to recognize and draw figures and their movement. 13. Read data from tables, pictographs, bar graphs, and pie charts. 14. Comparing and understanding how to use information from data. *Organizing and Representing (1)* 15. Understanding different representations and organizing data using tables, pictographs, and bar graphs.

To define the skills required in solving a particular item, 15 attributes are used in this study. Lee, Park, and Taylan (2011) developed those attributes based on the TIMSS 2007 Mathematics Framework (Mullis et al., 2005). The process of the attribute list is formed according to the TIMSS's specific content subdomain areas, which are number, geometric shapes and measures, and data display, also 38 objectives of TIMSS. Then three mathematics educators experienced in fourth-grade mathematics, and two domain-expert researchers defined the attributes for the TIMSS 2007. According to the attributes given in the Table 1, Q-matrix for the TIMSS 2007 is defined for 25 fourth grade mathematics items (Lee, Park, and Taylan, 2011) and given in Table 2.

Table 2
TIMSS 2007 fourth grade mathematics Q-matrix

		Attribute														
Item		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
1	M041052	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
2	M041056	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
3	M041069	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0
4	M041076	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
5	M041281	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0
6	M041164	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
7	M041146	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0
8	M041152	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0
9	M041258A	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
10	M041258B	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0
11	M041131	0	1	1	1	0	0	0	0	1	0	0	0	0	0	0
12	M041275	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1
13	M041186	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0
14	M041336	1	1	0	0	1	1	0	0	0	0	0	0	1	1	0
15	M031303	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
16	M031309	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
17	M031245	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
18	M031242A	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0
19	M031242B	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0
20	M031242C	0	1	1	0	0	0	0	1	0	0	0	0	0	1	0
21	M031247	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0
22	M031219	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0
23	M031173	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
24	M031085	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
25	M031172	1	1	0	0	0	0	0	0	0	0	0	0	1	0	1

The fundamental step of this study is to define the Q-matrix for TIMSS 2011 fourth grade mathematics questions from the proposed attributes for TIMSS 2007 by the experts. TIMSS mathematics results are stated to be comparable across participated countries and over years across cycles.

Table 3

TIMSS 2011 fourth grade mathematics Q-matrix 1

		Attribute														
Item		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
1	M051305	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0
2	M051091	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0
3	M051001	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0
4	M051007	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0
5	M051203	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
6	M051601	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0
7	M051064A	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1
8	M051064B	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
9	M051015	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0
10	M051123	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
11	M051109	1	1	0	0	0	0	0	0	0	0	0	0	1	0	1
12	M051117	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0
13	M041010	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	M041098	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0
15	M041064	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
16	M041003	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0
17	M041104	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
18	M041299	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
19	M041329	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
20	M041143	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
21	M041158	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0
22	M041328	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0
23	M041155	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0
24	M041284	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0
25	M041335	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0
26	M041184	1	0	1	0	0	0	0	0	0	0	0	0	1	0	1

Furthermore, since the same mathematics framework, and the same specific content subdomain areas, with 38 objectives are defined for TIMSS mathematics questions for all cycles (Olson et al., 2009; Mullis et al., 2012), the use of the identical attributes is expected to yield valid results over time. Thus, the defined 15 mathematics attributes for the TIMSS 2007 is used across countries, also across cycles of TIMSS. For this purpose, the defined 15 mathematics attributes is used in order to build the Q-matrices of American and Turkish samples in the TIMSS 2011 mathematics assessment. According to the attributes assigned to the TIMSS 2007 items by Lee, Park, and Taylan (2011), booklets are chosen from TIMSS 2011 and the first Q-matrix for TIMSS 2011 dataset is defined in Table 3.

Table 4

TIMSS 2011 fourth grade mathematics Q-matrix 2

		Attribute														
Item		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
1	M051305	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0
2	M051091	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
3	M051001	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0
4	M051007	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0
5	M051203	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
6	M051601	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0
7	M051064A	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
8	M051064B	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
9	M051015	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0
10	M051123	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
11	M051109	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0
12	M051117	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0
13	M041010	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	M041098	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
15	M041064	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
16	M041003	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0
17	M041104	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
18	M041299	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
19	M041329	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
20	M041143	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
21	M041158	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0
22	M041328	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0
23	M041155	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0
24	M041284	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0
25	M041335	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
26	M041184	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1

The second Q-matrix for TIMSS 2011 dataset is defined according to the solution of the 26 specific items and given at Table 4. Some of the required attributes that are defined for several items on the first Q-matrix that may not be required are removed and thus an under-specified Q-matrix is created. For example, an examinee requires to posses attributes 2, 4, and 5 according to the Q-matrix 1, and attribute 5 only according to Q-matrix 2 to answer the item 2 correct.

The third Q-matrix for TIMSS 2011 dataset is defined according to the solution of the 26 specific items again and given at Table 5. However, some of the required attributes for several items on the first Q-matrix that may not be included are added and an over-specified Q-matrix is created compared to the first Q-matrix. For example, an examinee requires to posses attribute 2 only according to the Q-matrix 1, and attributes 2, and 7 according to Q-matrix 3 to answer the item 5 correct.

Table 5

TIMSS 2011 fourth grade mathematics Q-matrix 3

		Attribute														
Item		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
1	M051305	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0
2	M051091	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0
3	M051001	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0
4	M051007	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0
5	M051203	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
6	M051601	1	1	0	1	0	0	0	1	0	0	0	0	0	0	0
7	M051064A	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1
8	M051064B	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
9	M051015	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0
10	M051123	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0
11	M051109	1	1	0	0	0	0	0	0	0	0	0	0	1	0	1
12	M051117	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0
13	M041010	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	M041098	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0
15	M041064	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0
16	M041003	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0
17	M041104	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
18	M041299	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
19	M041329	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
20	M041143	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
21	M041158	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0
22	M041328	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0
23	M041155	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0
24	M041284	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0
25	M041335	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0
26	M041184	1	0	1	0	0	0	0	0	0	0	0	0	1	0	1

In the last step, the fourth Q-matrix for TIMSS 2011 dataset is defined according to the results of the first three Q-matrices and given at Table 6. The decision is made according to the attribute classification results from Q-matrix 1. Some of the required attributes for several items on the first Q-matrix that may not be required are removed and some of the required attributes for several items on the first Q-matrix that may not be included are added.

Given the defined four Q-matrices for TIMSS 2011 and one Q-matrix for TIMSS 2007, the datasets of TIMSS 2007 and TIMSS 2011 are fitted to (1) the DINA and the G-DINA model respectively, and (2) four Q-matrices for TIMSS 2011 are compared using the G-DINA model. The computer program Ox (Doornik, 2002) was used for analysis.

Table 6

TIMSS 2011 fourth grade mathematics Q-matrix 4

		Attribute														
Item		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
1	M051305	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0
2	M051091	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
3	M051001	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0
4	M051007	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0
5	M051203	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
6	M051601	1	1	0	1	0	0	0	1	0	0	0	0	0	0	0
7	M051064A	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
8	M051064B	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
9	M051015	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0
10	M051123	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0
11	M051109	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0
12	M051117	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0
13	M041010	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	M041098	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
15	M041064	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0
16	M041003	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0
17	M041104	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
18	M041299	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
19	M041329	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
20	M041143	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
21	M041158	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0
22	M041328	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0
23	M041155	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0
24	M041284	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0
25	M041335	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
26	M041184	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1

RESULTS

The model fit results for the DINA and G-DINA models are evaluated by -2loglikelihood (-2LL), AIC and BIC statistics. The results are given in Table 7 and Table 8.

Table 7

Model fit of the DINA and the G-DINA model

		2LL	AIC	BIC
2007	DINA	22383.0145	88017.0145	242682.0981
	G-DINA	20804.8657	86862.8657	242527.0961
2011	DINA	24747.1937	90385.1937	247921.9752
	G-DINA	23304.2153	89242.2153	247499.0223

According to the results on Table 7, for both TIMSS 2007 and TIMSS 2011 the G-DINA model fit is better than that of the DINA model. The differences between the model fits are significant based on the BIC results across the DINA and the G-DINA models. For TIMMS 2007 the difference is 155.002 (242682.0981 - 242527.0961) and for TIMSS 2011 the difference is 422.9529 (247499.0223 - 247921.9752).

Table 8

Model fit of the G-DINA model based on different Q-matrices

		2LL	AIC	BIC
2011	Q-matrix 1	23304.2153	89242.2153	247499.0223
	Q-matrix 2	23302.3245	89216.3245	247487.0719
	Q-matrix 3	23304.4782	89240.4782	247492.1330
	Q-matrix 4	23301.1198	89123.1198	247101.5169

According to the results on Table 8, the differences across four Q-matrices are significant except for the differences between Q-matrix 1 and Q-matrix 3. The best model fit is obtained with the last Q-matrix when the required attributes are modified after considering the Q-matrix 1 results. The BIC differences are 11.9504, 6.8893, and 397.5054 between Q-matrix 1 to 2, 1 to 3 and 1 to 4 respectively.

DISCUSSION AND CONCLUSION

To evaluate the students' mastery skills and gain information about their ability levels, CDM models provide better estimation and detailed information than the traditional methods. This paper investigated that the model-data fit of the DINA and the G-DINA models in the United States sample for the fourth grade mathematics items in TIMSS 2007 and TIMSS 2011, and G-DINA model fit of the TIMSS 2011 with four Q-matrices defined.

To design a Q-matrix is one of the biggest challenges in the use of CDM models (de la Torre, 2009). Usually some content experts, teachers and researchers need to work together and decide for the attributes, and build Q-matrices for each specific test. In this paper, the use of the same attributes is investigated across years with an internationally administered large-scale test by comparing model fit of different Q-matrices under the G-DINA model. Furthermore, the process of building Q-matrices should be investigated by different empirical and methodological approaches to find the more appropriate Q-matrix.

This study investigates the fit of the DINA and the G-DINA models in the United States and Turkish samples for the fourth grade mathematics items in TIMSS 2007 and TIMSS 2011. Also the G-DINA model interpretations are interpreted to define mastery levels of Turkish students in TIMSS 2011 mathematics assessment.

The results reveal that the fit of the G-DINA models are better than the DINA models under all circumstances. According to the literature, CDMs have better model fit to data than traditional statistical models and IRT models, and also they are more useful to obtain detailed diagnostic information in student attainment. The advantages of CDMs are to be able to define student attainment in attribute mastery level, and to outline information for student instruction in specific subjects even in very small sample sizes. On the other hand, studying with CDMs has drawback of relying on the Q-matrix. Therefore, the Q-matrix needs to be validated by the combined findings from different approaches, together with relative theories, and content expert opinions before making final conclusions (De la Torre, 2008; Tatsuoka, 2009). The use of the CDMs in large-scale surveys is demonstrated in this study with TIMSS fourth grade mathematics items. The Q-matrices are created by the experienced content specialists, and tested by using the DINA and the G-DINA models. Thus, the first limitation of this study is the use of the particular attributes under the particular models only. In further studies, alternative attribute sets and Q-matrices can be proposed by the other researchers and experienced content specialists, also other statistical methods can be used to test the validation of the Q-matrices. The process of building Q-matrices should be investigated by different approaches to find the more appropriate Q-matrices.

In future research, the validation of the intuitive Q-matrices should be investigated by using methodological approaches. The major limitation of expert-base Q-matrices, being subjective, should be supported by the computer-based definitions of Q-matrices.

REFERENCES

American Federation of Teachers (AFT). (1999). What TIMSS tells us about mathematics achievement, curriculum, and instruction. *In Educational Issues Policy Brief, 10*, 2-10.Washington, DC: AFT Educational Issues Department.

Birenbaum, M., Nasser, F., & Tatsuoka, C. (2007). Effects of gender and ethnicity on fourth graders' knowledge states in mathematics. *International Journal of Mathematical Education in Science and Technology*, *38*, 301-319.

Birenbaum, M., Tatsuoka, C., & Yamada, T. (2004). Diagnostic assessment in TIMSS-R: Between countries and within-country comparisons of eighth graders' mathematics performance. *Studies in Educational Evaluation, 30*, 151-173.

Chen, Y., Thompson, M., Gorin, J. S., & Tatsuoka K. K. (2008). Cross-Cultural validity of the TIMSS-1999 mathematics test: Verification of a Cognitive Model. *International Journal of Testing, 8*, 251-271.

De la Torre, J. (2008). An empirically-based method of Q-matrix validation for the DINA model: Development and applications. *Journal of Educational Measurement, 45,* 343-362.

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

De la Torre, J. (2011). The generalized DINA model framework. *Psychometrika, 76*, 179-199.

Dogan, E., & Tatsuoka, K. K. (2008). An international comparison using a diagnostic testing model: Turkish students' profile of mathematical skills on TIMSS-R. *Educational Studies in Mathematics, 68*, 263-272.

Doornik, J. A. (2002). Object-Oriented Matrix Programming Using Ox (Version 3.1) [Computer software]. London: Timberlake Consultants Press.

Foy, P., & Olson, J. F. (2009). *TIMSS 2007 user guide for the international database*. Chestnut Hill, MA: IEA.

Haertel, E. H. (1989). Using restricted latent class models to map the skill structure of achievement items. *Journal of Educational Measurement, 26*, 333-352.

Hambleton, R. K. (2005). *Issues, designs, and technical guidelines for adapting tests in multiple languages*. In R. K. Hambleton, P. Merenda, & C. D. Spielberger (Eds.). Hillsdale, NJ: Lawrence Erlbaum.

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, Park, & Taylan (2011). Cognitive diagnostic modeling of attribute mastery in Massachusetts, Minnesota, and the U.S. national sample using the TIMSS 2007. *International Journal of Testing, 11*, 144-177.

Mullis, I.V.S., Martin, M.O., Foy, P., & Arora, A. (2012). *TIMSS 2011 International Results in Mathematics.* Chestnut Hill, MA: IEA.

Olson, J. F., Martin,M. O., & Mullis, I. V. S. (2009). *TIMSS 2007 technical report*. Chestnut Hill, MA: IEA.

Tatsuoka, K. K. (2009). *Cognitive assessment: An introduction to the rule space method.* Routledge, NY.

Tatsuoka, K. K. (1983). Rule-space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement, 20*, 345-354.