



(ISSN: 2602-4047)

Uyumaz, G. (2023). The Impact of Q Matrix Misspecification on Item Parameters and Classification Accuracy in the DINA Model, *International Journal of Eurasian Education and Culture*, 8(23), 2238-2269.

DOI: <http://dx.doi.org/10.35826/ijoecc.769>

Article Type : Research Article

## THE IMPACT OF Q MATRIX MISSPECIFICATION ON ITEM PARAMETERS AND CLASSIFICATION ACCURACY IN THE DINA MODEL<sup>1</sup>

**Gizem UYUMAZ**

Assistant Professor, Giresun University, Giresun, Turkiye, [gizem.uyumaz@giresun.edu.tr](mailto:gizem.uyumaz@giresun.edu.tr)

ORCID: 0000-0003-0792-2289

Received: 15.03.2023

Accepted: 17.09.2023

Published: 01.10.2023

### ABSTRACT

To accurately reflect students' current proficiency levels through achievement tests, it is essential to make determinations based on the attribute levels measured by the items. The validity of classifying students based on their measured attributes is enhanced when the items in the achievement test have varying effects on classification, taking into consideration the attributes measured by each item. The accuracy of establishing item-attribute relationships, that is, the accuracy of the Q matrix, affects the accuracy of classification, the quality of inference, and the accuracy of decisions made about students. This research aims to investigate the impact of the Q matrix misspecification in the DINA model on the item parameters and the classification of individuals in simulated data sets. Within the scope of the study, examinations were conducted regarding misspecification rates of 5%, 7.5%, and 10%, and misspecification patterns of under-fitting, over-fitting, and balanced-misfit. Simulated data were generated and analyzed using R and Mplus software. Item parameters and classifications made with misspecified Q matrices were compared with the estimations made using the true Q matrix appropriate for the data set. Under the examined conditions of this study, it was found that each misspecification condition differentiated the classifications of individuals. In the under-fitting Q matrix conditions, the classification accuracy decreased as the misspecification rate increased. In the over-fitting conditions, the classification accuracy increased as the misspecification rate increased; all resulted in lower classification accuracy compared to those obtained using the true Q matrix. In balanced-misfit conditions, an increase in the misspecification rate decreased classification accuracy. Over-fitting and balanced-misfit conditions had a more substantial negative impact on classification accuracy compared to under-fitting Q matrices. Under-fitting conditions increased the slipping parameter. Over-fitting conditions increased the guessing parameter. Balanced-misfit conditions increased under-fitting items' slipping parameter, over-fitting items' guessing parameter, and balanced-misfit items' both slipping and guessing parameters. The increases in slipping parameters were much greater than the increases in guessing parameters.

**Keywords:** Over-fitting, under-fitting, balanced-misfit.

<sup>1</sup> A part of this article was created from the doctoral thesis prepared by the author.

## INTRODUCTION

In the field of education, at all levels and in every classroom setting, achievement tests are administered to assess students' learning progress. Achievement tests involve classifying student performance on an item level or the entire test. During these classifications, the attributes required to answer each item correctly are taken into account, and if the impact of each item on student classification in the test can be determined by considering the attributes necessary for its correct response, the validity of the classifications and the accuracy of the decisions made will improve. One effective approach to achieving this is by using Cognitive Diagnostic Models (CDMs).

In Cognitive Diagnostic Models, which are a type of specialized latent class model, an attribute profile is created for each student based on their responses to the test items, and groups of students with similar attribute profiles are identified. The number of possible groups (latent classes) is predetermined based on the total number of attributes and dependencies among attributes. Therefore, CDMs are considered constrained latent class models (Haertel, 1989; Jang, Kim, Vincett, Barron, & Russell, 2019; Rupp et al., 2010).

In Cognitive Diagnostic Models, when the attributes measured by each item are accurately identified and defined, students' performance on each attribute and the entire test can be determined in a more detailed, accurate, and valid manner. This enables the identification of students' strengths, weaknesses, and readiness levels for transitioning to the next educational stage, and facilitates guiding the educational process (de la Torre, 2009a; Henson & Templin, 2006; Leighton & Gierl, 2007; Rupp, 2007; Rupp & Templin, 2008a; 2008b; Rupp, Templin & Henson, 2010).

In Cognitive Diagnostic Models, the identification of attributes measured by each item is achieved through the Q matrix. The Q matrix is created by encoding the item-attribute relationships numerically (as a J×K matrix of 1s and 0s), where rows represent items and columns represent attributes (Rupp et al., 2010). For each item, if possessing a particular attribute (attribute k) is required to answer that item (item j) correctly, a "1" is placed in the intersection cell ( $q_{jk}$  cell) where the row corresponding to the item j and the column corresponding to attribute k meet. If not required, "0" is entered. The Q matrix reflects the structure of the administered achievement test and serves as the fundamental building block for the quality of decisions made using this achievement test (de la Torre, 2009a; Rupp & Templin, 2008a; Rupp, Templin & Henson, 2010). The theoretical and empirical soundness of the Q matrix determines the diagnostic power of the models, making it highly important for studies conducted with CDMs (Lee & Sawaki, 2009).

In Cognitive Diagnostic Models, the accuracy of decisions made relies on the items being well-crafted and aligned with criteria, and when the test's content validity is ensured. In such cases, the validity of individuals' classifications is contingent upon the test's construct validity. The Q matrix misspecification will result in erroneous individual profiles (de la Torre, 2009a; Im & Corter, 2011; Qin et al., 2015; Romero et al., 2014; Rupp & Templin, 2008a; 2008b; Rupp et al., 2010; Tatsuoka, 1983). The accuracy of parameter estimation in CDMs

relies on the accuracy of the Q matrix (Baker, 1993). Since the Q matrix is often determined through expert judgment, misclassifications can occur (DeCarlo, 2012). Ensuring the accurate determination of the classes in which students are placed holds utmost significance, as decisions concerning students are based on these classifications (Kunina-Habenicht, Rupp, & Wilhelm, 2012).

The aim of this study is to investigate the effect of the Q matrix misspecification on item parameters and the classification of individuals. In this research, the study was conducted with the DINA (The Deterministic-Input, Noisy 'And' Gate; de la Torre, 2009a; Haertel, 1989; Junker & Sijtsma, 2001) model.

### The DINA Model

The DINA model, where attributes and responses are binary scored, is a non-compensatory model. For a student to correctly answer an item, they must possess all attributes associated with that item. The absence of a required attribute in a student cannot be compensated by their possession of another attribute (de la Torre, 2009b; Henson & Templin, 2006; Rupp & Templin, 2008a; Rupp & Templin, 2008b; Rupp, Templin & Henson, 2010).

In the DINA model, two types of non-compensatory response probabilities are calculated for items: 'slipping ( $s_j$ )' and 'guessing ( $g_j$ ).' The probability of a student possessing all attributes associated with an item and yet responding incorrectly to that item is expressed by the slipping parameter, whereas the probability of a student not possessing at least one required attribute and still responding correctly to the item is represented by the guessing parameter (Rupp & Templin, 2008a; Rupp et al., 2010). The relevant formulas are presented as Equation 1 and Equation 2.

$$s_j = P(X_{ij} = 0 | \eta_{ij} = 1) \quad (1)$$

$$g_j = P(X_{ij} = 1 | \eta_{ij} = 0) \quad (2)$$

$s_j$ : The probability of a student possessing attributes related to an item, yet providing an incorrect response (false positive) to item j.

$g_j$ : The probability of a student not having attributes associated with an item, yet providing a correct response (true positive) to item j.

$X_{ij}$ : The score of student i on item j (0, 1), where ( $i = 1, \dots, I$  and  $j = 1, \dots, J$ ).

$\eta_{ij}$ : Latent determinantal response.

In the DINA model, the latent determinantal response is represented as shown in Equation 3:

$$\eta_{ij} = \prod_{k=1}^K \alpha_{ik}^{q_{jk}} \quad (3)$$

$\alpha_i = \{ \alpha_{ik} \}$ : Binary attribute vector indicating whether student  $i$  possess attribute  $k$  ( $k = 1, \dots, K$ )

$q_{jk}$ : Binary input in the Q matrix corresponding to item  $j$  and attribute  $k$  (1, 0)

The probability of a student possessing all associated attributes correctly answering an item is given by Equation 4 and Equation 5 (De la Torre, 2009a; Rupp et al., 2010).

$$P_j(\alpha_i) = P(X_{ij} = 1 | \eta_{ij}) = (1 - s_j)^{\eta_{ij}} g_j^{1-\eta_{ij}} \quad (4)$$

$$P(X_1 = 1 | \alpha_1 \alpha_2) = \frac{\exp(\lambda_{1,0} + \lambda_{1,(1)}\alpha_1 + \lambda_{1,1,(2)}\alpha_2 + \lambda_{1,2,(1,2)}\alpha_1\alpha_2)}{1 + \exp(\lambda_{1,0} + \lambda_{1,1,(1)}\alpha_1 + \lambda_{1,1,(2)}\alpha_2 + \lambda_{1,2,(1,2)}\alpha_1\alpha_2)} \quad (5)$$

$P_j(\alpha_i)$ : The probability of student  $i$ , possessing the attribute vector  $\alpha_i$ , answering item  $j$  correctly.

In the DINA model, when forming students' attribute profiles, individual parameters are initially computed. This process, which also includes item and structural parameters, is performed using the Maximum Likelihood approach. To differentiate the latent classes where students will be placed, a preliminary classification is computed, and this classification is used to compute item parameters. These computed parameters are taken as known item parameters, and a posterior latent class classification is computed. The calculation algorithm follows an iterative refinement process where item parameters are recalculated at each step. The estimations from the previous step are compared with the current estimations until the required convergence is achieved (De la Torre, 2009a; Rupp et al., 2010).

Building on previous research (Baker, 1993; de la Torre, 2008; Im & Corter, 2011; Kunina-Habenicht, Rupp, & Wilhelm, 2012; MacDonald, 2013; MacDonald & Kromrey, 2012; Rupp & Templin, 2008a) that explored various potential the Q matrix misspecification scenarios, this study had decided to investigate the impact of the misspecification rate (%5, %7.5, and %10) and the misspecification pattern (under-fitting, over-fitting, and balanced-misfit) on the Q matrix. The theoretical considerations regarding how the Q matrix misspecification affects item parameters and individuals' classifications are important in terms of the accuracy of decisions made through CDMs. This research aims to provide guidance to researchers and practitioners and contribute to the accuracy of decisions made using CDMs when the findings are applied. This study's significance lies in its potential to illuminate the way forward for researchers and practitioners and enhance the accuracy of decisions made through CDMs when the insights gained are put into practice.

## METHOD

Since this study examines the effect of the Q matrix misspecification on item parameters and individuals' classifications using simulated data in the DINA model, it falls under the category of a fundamental research.

### Data Generation

The number of attributes measured by the assessment tool, the sample size on which the measurement tool is applied, and the test length are crucial factors affecting classification consistency. Based on previous studies, the decision was made to have four attributes (Henson & Douglas, 2005; Kunina-Habenicht et al., 2012; Ömür-Sünbül & Kan, 2015; Rupp & Templin, 2008a), a sample size of 1000 (Baker, 1993; Cassuto, 1996; De la Torre et al., 2010; Kunina-Habenicht et al., 2012; MacDonald & Kromrey, 2011; Orlando & Marshall, 2002; Ömür-Sünbül & Kan, 2015; Tsutakawa & Johnson, 1990), and a test length of 30 items (Cassuto, 1996; Henson & Templin, 2006; Kunina-Habenicht et al., 2012; MacDonald, 2013; Ömür-Sünbül & Kan, 2015; Tatsuoka, 1990; Templin et al., 2009) for this study.

The decision regarding the number of items was also influenced by the potential attribute patterns. In the case of measuring four independent attributes, there are  $2^4$  possible attribute patterns, which results in 16 ( $2^4 = 16$ ) expected attribute classes. Correspondingly, there are 15 different items capable of measuring these 4 attributes ( $2^4 - 1$ , where  $2^4 - 1 = 15$ ). Each sub-attribute profile represents a multi-dimensional profile through differential combinations of 0s and 1s (Jang, Dunlop, Park, & van der Boom, 2015). The total number of possible items is one less than the expected number of attribute profiles because there won't be an item (represented by a row of 0000) that measures no attributes in the achievement test. In the administered achievement test, there can be multiple items measuring the same attributes, so the test may contain more than  $2^4 - 1$  items and doesn't necessarily have to encompass all possible attribute patterns (Rupp et al., 2010).

In this study, when constructing items and the Q matrix in terms of measurement sensitivity, care was taken to have approximately the same number of items associated with each attribute and to represent all possible attribute patterns (DiBello & Stout, 2007; Kunina-Habenicht et al., 2012; Rupp & Templin, 2008a). The first 15 items were created to represent the 15 different possible patterns, ensuring that each attribute was associated with a similar number of items. These initial 15 items were then duplicated to create the second half, resulting in a total of 30 items for the hypothetical measurement instrument under investigation (True Q matrix is presented in Appendix 1).

In this study, determining the Q matrix misspecification conditions was misspecified involved examining previous research (Baker, 1993; Kunina-Habenicht et al., 2012; MacDonald, 2013; MacDonald & Kromrey, 2012; Rupp & Templin, 2008a) and selecting scenarios that were likely to be encountered in real-life situations. To achieve this, it was decided to set the misspecification rate at 5%, 7.5%, and 10%, and the misspecification patterns as under-fitting, over-fitting, and balanced-misfit. The 5% misspecification rate implies that six cells in the Q matrix contain misclassifications; at 7.5%, there are nine misclassified cells; and at 10%, there are twelve misclassified cells. Under-fitting misspecification involves replacing relevant 1s with 0s in the cells of the Q matrix, whereas balanced-misfit means replacing half of the relevant 0s with 1s and half of the relevant 1s with 0s. When

constructing the Q matrix misspecification conditions, the items and attributes to be misclassified were randomly selected from the cells of the first 15 items.

In order to facilitate ease of use and provide succinct descriptions, the coding schemes used for the examined Q matrices are presented in Table 1, along with information on the average number of attributes per item, the average number of items per attribute, and the sum of 1s and 0s in the Q matrix.

**Table 1.** Information Regarding Q Matrices

Description	Condition	Average Number of Attributes per Item	Average Number of Items per Attribute	Total # of Change (0 to 1)	Total # of Change (1 to 0)
Data-generating Q-matrix	Q0	2.13	16.00	0	0
Q matrix with 5% under-fitting	Q1	1.93	14.50	0	6
Q matrix with 7.5% under-fitting	Q2	1.83	13.75	0	9
Q matrix with 10% under-fitting	Q3	1.73	13.00	0	12
Q matrix with 5% over-fitting	Q4	2.33	17.50	6	0
Q matrix with 7.5% over-fitting	Q5	2.43	18.25	9	0
Q matrix with 10% over-fitting	Q6	2.53	19.00	12	0
Q matrix with 5% balanced-misfit	Q7	2.13	16.00	3	3
Q matrix with 7.5% balanced-misfit	Q8	2.10	15.75	4	5
Q matrix with 10% balanced-misfit	Q9	2.13	16.00	6	6

The average number of attributes per item in Table 1 represents the ratio of the number of cells with a value of 1 in the Q matrix to the number of items (30). Similarly, the average number of items per attribute is the ratio of the number of cells with a value of 1 in the Q matrix to the number of attributes (4).

The data were generated using the "dina" package in the R software (Culpepper, 2015), based on the true Q matrix provided in Appendix 1 and in line with the appropriate DINA model by setting the g and s parameters to 0.00. Initially, respondent attribute profiles were created consisting of 1000 individuals with binary (0-1) responses, containing approximately an equal number of students in each of the 16 possible attribute profiles. Subsequently, response patterns that adhered to this respondent attribute profiles were generated. Each condition was replicated 100 times.

### The Data Analysis

The data analysis was conducted using R and Mplus software. In Mplus, the syntax designed by Templin and Hoffman (2013) for latent class models was adapted with specific modifications for DINA estimation. For each generated response pattern from the replications, both true classifications using the true Q matrix and misclassified classifications using each of the Q matrix misspecification was determined. In the estimations, 1000 iterations were performed, and the convergence criterion was set to 0.001.

Classification accuracy refers to the degree of concordance between the observed classification in the data, which represents the true latent classes, and the latent classes obtained from the analysis (Cui et al., 2012). In

the analyses with the true Q matrix, the latent classes defined during data generation were used as the true latent classes. These true latent classes were compared with the latent classes estimated using the MLE method from the analyses with the Q matrix misspecification. If the two latent class assignments were consistent for an individual, a value of "1" was assigned to indicate correct classification; otherwise, a value of "0" was assigned to represent misclassification. By averaging these 1-0 values across all individuals and all replications, overall correct classification rates were calculated for each condition.

**FINDINGS**

In this section, the findings obtained from the true Q matrix and the Q matrix misspecification for the simulated measurement tools containing 30 items are presented.

The item parameters estimated with under-fitting Q matrices are presented in Table 2.

**Table 2.** Under-Fitting Q Matrices Item Parameters

Item	Q0		Q1		Q2		Q3	
	g	s	g	s	g	s	g	s
1	0.012011	0.006900	0.000041	0.000039	0.000041	0.000040	0.000040	0.000040
2	0.011720	0.007101	0.000038	0.000042	0.000039	0.000039	0.000040	0.000040
3	0.012258	0.006797	0.000040	0.000039	0.000041	0.000039	0.000041	0.000039
4	0.012935	0.006413	0.000044	0.000037	0.000044	0.000037	0.000044	0.000036
5	0.006489	0.011160	0.000040	0.487064	0.000041	0.487105	0.000040	0.487059
6	0.006828	0.010827	0.000027	0.000072	0.000040	0.455277	0.000041	0.455312
7	0.006673	0.010969	0.000028	0.000075	0.000028	0.000078	0.000043	0.506586
8	0.005880	0.011855	0.000026	0.000086	0.000038	0.531560	0.000041	0.551707
9	0.006762	0.010803	0.000044	0.499083	0.000039	0.448053	0.000040	0.448044
10	0.006853	0.010896	0.000042	0.464016	0.000040	0.463959	0.000028	0.000073
11	0.003713	0.016189	0.000027	0.499901	0.000027	0.529256	0.000027	0.499957
12	0.002464	0.011658	0.000027	0.505639	0.000028	0.512943	0.000027	0.512834
13	0.003927	0.015646	0.000023	0.000146	0.000027	0.479313	0.000044	<b>0.743072</b>
14	0.003629	0.016529	0.000027	0.538618	0.000023	0.000163	0.000027	0.545403
15	0.002298	0.024557	0.000021	0.000319	0.000023	0.503996	0.000028	<b>0.771180</b>
16	0.012006	0.006933	0.000040	0.000040	0.000041	0.000040	0.000039	0.000041
17	0.006896	0.004577	0.000039	0.000041	0.000039	0.000041	0.000039	0.000042
18	0.012251	0.006751	0.000041	0.000040	0.000042	0.000038	0.000041	0.000040
19	0.012935	0.006411	0.000043	0.000037	0.000043	0.000037	0.000043	0.000037
20	0.006485	0.011221	0.000028	0.000078	0.000027	0.000077	0.000028	0.000078
21	0.006838	0.010740	0.000028	0.000073	0.000027	0.000072	0.000027	0.000073
22	0.006669	0.010957	0.000027	0.000076	0.000027	0.000076	0.000027	0.000076
23	0.005872	0.011904	0.000026	0.000087	0.000026	0.000087	0.000026	0.000086
24	0.006766	0.010860	0.000027	0.000074	0.000028	0.000075	0.000027	0.000075
25	0.006852	0.010736	0.000028	0.000074	0.000028	0.000074	0.000027	0.000070
26	0.003712	0.016268	0.000023	0.000152	0.000022	0.000158	0.000022	0.000158
27	0.003783	0.015930	0.000023	0.000151	0.000023	0.000151	0.000023	0.000148
28	0.003929	0.015551	0.000024	0.000146	0.000023	0.000147	0.000023	0.000147
29	0.003633	0.016427	0.000023	0.000161	0.000023	0.000161	0.000023	0.000163
30	0.002298	0.024666	0.000021	0.000333	0.000022	0.000329	0.000021	0.000318

s: slipping parameter, g: guessing parameter

When examining Table 2, it can be observed that the guessing parameters of the under-fitting items (highlighted in color) remain unchanged, while the slipping parameters and their standard error values increase.

The item parameters estimated with over-fitting Q matrices are presented in Table 3.

**Table 3.** Over-Fitting Q Matrices Item Parameters

Item	Q0		Q4		Q5		Q6	
	g	s	g	s	g	s	g	s
1	0.012011	0.006900	0.000041	0.000041	0.000041	0.000039	0.321910	0.000077
2	0.011720	0.007101	0.339005	0.000088	0.301799	0.000074	<b>0.413553</b>	0.000149
3	0.012258	0.006797	0.367551	0.000088	0.367456	0.000086	0.000041	0.000038
4	0.012935	0.006413	0.000043	0.000038	<b>0.471152</b>	0.000154	<b>0.466956</b>	0.000145
5	0.006489	0.011160	0.000027	0.000078	0.000026	0.000078	0.148145	0.000152
6	0.006828	0.010827	0.000027	0.000072	0.000028	0.000075	0.000028	0.000075
7	0.006673	0.010969	0.000027	0.000073	0.000027	0.000075	0.148693	0.000145
8	0.005880	0.011855	0.000026	0.000087	0.000026	0.000088	0.000026	0.000088
9	0.006762	0.010803	0.222890	0.000328	0.222893	0.000313	0.000028	0.000075
10	0.006853	0.010896	0.000028	0.000073	0.171486	0.000164	0.157968	0.000142
11	0.003713	0.016189	0.071455	0.000325	0.071428	0.000322	0.071465	0.000323
12	0.002464	0.011658	0.074628	0.000329	0.000023	0.000149	0.074649	0.000323
13	0.003927	0.015646	0.000023	0.000144	0.082119	0.000328	0.082126	0.000316
14	0.003629	0.016529	0.000023	0.000162	0.000023	0.000157	0.067149	0.000318
15	0.002298	0.024557	0.000021	0.000338	0.000021	0.000312	0.000021	0.000328
16	0.012006	0.006933	0.000040	0.000040	0.000041	0.000039	0.000046	0.000044
17	0.006896	0.004577	0.000046	0.000045	0.000045	0.000046	0.000046	0.000046
18	0.012251	0.006751	0.000048	0.000043	0.000047	0.000045	0.000041	0.000039
19	0.012935	0.006411	0.000043	0.000038	0.000051	0.000041	0.000051	0.000041
20	0.006485	0.011221	0.000027	0.000079	0.000027	0.000075	0.000026	0.000076
21	0.006838	0.010740	0.000027	0.000075	0.000028	0.000073	0.000027	0.000075
22	0.006669	0.010957	0.000027	0.000076	0.000027	0.000075	0.000027	0.000074
23	0.005872	0.011904	0.000026	0.000087	0.000027	0.000089	0.000026	0.000087
24	0.006766	0.010860	0.000028	0.000075	0.000027	0.000073	0.000027	0.000074
25	0.006852	0.010736	0.000028	0.000072	0.000027	0.000073	0.000027	0.000073
26	0.003712	0.016268	0.000023	0.000161	0.000023	0.000152	0.000023	0.000155
27	0.003783	0.015930	0.000023	0.000154	0.000023	0.000154	0.000023	0.000150
28	0.003929	0.015551	0.000023	0.000142	0.000024	0.000145	0.000024	0.000148
29	0.003633	0.016427	0.000023	0.000159	0.000023	0.000161	0.000023	0.000161
30	0.002298	0.024666	0.000022	0.000329	0.000022	0.000320	0.000021	0.000336

s: slipping parameter, g: guessing parameter

When examining Table 3, it can be observed that the slipping parameters of the over-fitting items (highlighted in color) remain unchanged, while the guessing parameters and their standard error values increase.

The item parameters estimated with balanced-misfit Q matrices are presented in Table 4. When Table 4 is examined, in the case of balanced-misfit; it can be observed that the guessing parameters and their standard error values for items with over-fitting have increased, the slipping parameters and their standard error values



for items with under-fitting have increased, and for items where both under-fitting and over-fitting occurred, both parameters and their standard error values have increased.

**Table 4.** Balanced-Misfit Q Matrices Item Parameters

Item	Q0		Q7		Q8		Q9	
	g	s	g	s	g	s	g	s
1	0.012011	0.006900	0.000041	0.000040	0.000040	0.000040	0.000040	0.000040
2	0.011720	0.007101	0.000040	0.000041	0.000040	0.000041	0.338976	0.000084
3	0.012258	0.006797	0.000040	0.000039	0.000041	0.000038	0.000040	0.000039
4	0.012935	0.006413	0.000043	0.000037	0.000043	0.000037	0.370377	0.000074
5	0.006489	0.011160	0.148088	0.000153	0.257507	0.742001	0.145145	0.000151
6	0.006828	0.010827	0.000028	0.000073	0.000027	0.000073	0.000028	0.000075
7	0.006673	0.010969	0.176365	0.492751	0.185206	0.512892	0.000041	0.469259
8	0.005880	0.011855	0.000040	0.531538	0.139193	0.529191	0.000041	0.551690
9	0.006762	0.010803	0.000027	0.000073	0.000028	0.000074	0.189636	0.505434
10	0.006853	0.010896	0.000028	0.000072	0.171466	0.000159	0.157952	0.000141
11	0.003713	0.016189	0.000023	0.000153	0.000028	0.500052	0.000026	0.439276
12	0.002464	0.011658	0.000027	0.505530	0.000023	0.000153	0.074657	0.000320
13	0.003927	0.015646	0.000023	0.000146	0.000023	0.000147	0.000023	0.000146
14	0.003629	0.016529	0.067176	0.000317	0.000022	0.000160	0.000027	0.456513
15	0.002298	0.024557	0.000021	0.000325	0.000021	0.000328	0.000023	0.504040
16	0.012006	0.006933	0.000040	0.000040	0.000041	0.000040	0.000040	0.000039
17	0.006896	0.004577	0.000039	0.000039	0.000040	0.000040	0.000046	0.000047
18	0.012251	0.006751	0.000041	0.000039	0.000040	0.000039	0.000041	0.000038
19	0.012935	0.006411	0.000044	0.000036	0.000045	0.000038	0.000051	0.000042
20	0.006485	0.011221	0.000027	0.000079	0.000027	0.000077	0.000027	0.000075
21	0.006838	0.010740	0.000028	0.000073	0.000027	0.000074	0.000028	0.000074
22	0.006669	0.010957	0.000027	0.000074	0.000028	0.000076	0.000027	0.000075
23	0.005872	0.011904	0.000026	0.000089	0.000026	0.000085	0.000026	0.000087
24	0.006766	0.010860	0.000027	0.000076	0.000028	0.000077	0.000027	0.000075
25	0.006852	0.010736	0.000028	0.000073	0.000027	0.000074	0.000028	0.000072
26	0.003712	0.016268	0.000023	0.000154	0.000023	0.000156	0.000023	0.000158
27	0.003783	0.015930	0.000023	0.000151	0.000023	0.000152	0.000023	0.000151
28	0.003929	0.015551	0.000022	0.000144	0.000023	0.000141	0.000024	0.000142
29	0.003633	0.016427	0.000023	0.000159	0.000022	0.000161	0.000023	0.000156
30	0.002298	0.024666	0.000021	0.000323	0.000021	0.000327	0.000021	0.000315

s: slipping parameter, g: guessing parameter

The classification accuracy values obtained by comparing the attribute profiles of each individual derived from the true Q matrix condition with the attribute profiles obtained using the Q matrix misspecification, as well as the goodness-of-fit indices of the estimates, are presented in Table 5.

When examining Table 5, it can be observed that under conditions where the Q matrix is under-specified (Q1, Q2, and Q3), as the misspecification rate increases, the classification accuracy decreases. Conversely, under conditions where the Q matrix is over-specified (Q4, Q5, and Q6), an increase in the misspecification rate leads to an increase in classification accuracy. In conditions where the Q matrix is balanced-misfit (Q7, Q8, and Q9), as the misspecification rate increases, the classification accuracy decreases. Given that values of entropy close to 1

and lower values of goodness-of-fit indices indicate better fit, it can be seen that consistently with the classification accuracy findings, under conditions where the Q matrix is under-specified and balanced-misfit, an increase in the misspecification rate results in worse goodness-of-fit indices. Conversely, under conditions where the Q matrix is over-specified, an increase in the misspecification rate leads to an improvement in the goodness-of-fit indices.

**Table 5.** Classification Accuracies and Model Fit Indices

Condition	Classification Accuracy	LL	AIC	BIC	Entropy
Q1	0.941	-11647.863	23447.727	23820.716	0.997
Q2	0.871	-11844.227	23838.454	24206.535	0.971
Q3	0.741	-12024.674	24199.349	24567.431	0.972
Q4	0.667	-11172.531	22497.062	22870.052	0.980
Q5	0.691	-11884.880	23919.759	24287.841	0.966
Q6	0.736	-12068.489	24288.978	24661.968	0.895
Q7	0.871	-12315.613	24779.226	25142.400	0.939
Q8	0.731	-11711.180	23572.363	23940.444	0.974
Q9	0.626	-11576.394	23302.788	23670.869	0.978

LL = Log-Likelihood; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion

## CONCLUSION and DISCUSSION

In this study, the impact of Q1-Q9 matrix misspecification conditions on item parameters and the classification of individuals was examined within the DINA model. In each examined condition, the Q matrix misspecification affects both item parameters and the classification of individuals

In the DINA model, the lower the slipping and guessing parameters (close to zero), it indicates that the model is well-defined, responses are not random, and the items fit the data well. Item-specific model fit is achieved (De la Torre, 2009a; Rupp and Templin, 2008a). When the slipping and guessing parameter values are higher than zero, it suggests that the attributes defined in the Q matrix are insufficient to explain participants' response patterns, and that answering items involves a different strategy (De la Torre and Douglas, 2004). As the parameter values decrease, the probability of individuals who have mastered the related attributes answering the item correctly increases, while the probability of individuals who have not mastered the attributes answering the item incorrectly also increases. (De la Torre, 2009a). In this study, in parallel with similar research (Baker, 1993; De la Torre and Douglas, 2004; Im and Corter, 2011; Kunina-Habenicht et al., 2012; MacDonald, 2013; Rupp and Templin, 2008a), it has been found that the Q matrix misspecification affects parameter estimations.

Under-fitting Q matrix conditions did not affect the estimation of the guessing parameter of the item but increased the slipping parameter. Since the response pattern contains more incorrect responses than expected, the slipping parameter is estimated at a higher value than its true value (Rupp and Templin, 2008a). This situation implies that most of the individuals taking the test, although they seem to have mastered all the attributes measured by the relevant item, answer the item incorrectly. This is because at least one attribute measured by this item has been defined as if it is not measured by this item. And since individuals have not mastered this attribute, they answer the item incorrectly. However, in contrast to the expected situation based on their

performance on other items, individuals who are determined to have mastered the attributes defined by the relevant item, when the item is answered incorrectly, lead to an increase in the item's slipping parameter.

Over-fitting Q matrix conditions did not affect the estimation of the slipping parameter of the item but increased the guessing parameter. When an item is erroneously defined to require an additional attribute to be answered correctly, the item will be defined as more difficult than it actually is. Even if individuals have not mastered this additional attribute erroneously related to the item, they will be able to answer the item correctly. Since the response pattern contains more correct responses than expected, the guessing parameter is estimated at a higher value than its true value (Rupp and Templin, 2008a).

Balanced-misfit Q matrix conditions increased under-fitting items' slipping parameter, over-fitting items' guessing parameter, and balanced-misfit items' both slipping and guessing parameters. The increases in slipping parameters were much greater than the increases in guessing parameters. Similarly, when the same item had under-fitting in an attribute and over-fitting in another attribute, the slipping parameter of the item increased more than the guessing parameter.

Baker (1993) and MacDonald (2013) reported in their studies that the misspecification rates affected item parameter estimates, with estimates becoming less accurate as the misspecification rate increased. However, in this study, under different conditions, when the same misspecification was made for the same attribute, the parameter estimates almost remained the same. When multiple misspecifications were made for more than one attribute in an item, the increase in parameter estimates also escalated. For instance, when multiple over-fitting were made for a single item, the number of individuals who would answer the item correctly would be greater than in the case where only one over-fitting was made. This resulted in the guessing parameter for that attribute increasing more than in the case of only one misspecification.

In under-fitting Q matrix conditions, individuals' responses tend to contain more errors than they actually are due to the presence of items with incorrect specifications, making the task relatively easier. Conversely, in over-fitting Q matrix conditions, individuals' responses tend to contain more correct responses than they actually are due to the presence of items with incorrect specifications, making the task relatively harder (Rupp and Templin, 2008a). In balanced-misfit Q matrix conditions, the variation in the average difficulty of the test is related to randomly selected items with over-fitting or under-fitting specifications. Within the scope of this study, under the examined conditions, classification accuracies were found to be higher in cases with under-fitting specifications compared to cases with over-fitting and balanced-misfit specifications. This finding is consistent with the findings related to the DINA model from the study conducted by Miao, Miller, and Ren (2017).

When examining previous research, it has been found that as the Q matrix misspecification rate increases, the amount of error in the estimates also increases (Baker, 1993; Başokçu, 2014; MacDonald, 2013). Similarly, within the scope of this study, it has been observed that as the Q matrix misspecification rate increases, the classification accuracy decreases. This finding is consistent with the study conducted by Miao, Miller, and Ren (2017), where

various conditions were investigated including different test lengths and correlations between attributes within the DINA model. In the context of this study, similar to the Q matrix under-fitting conditions, it has been found that in conditions with Q matrix balanced-misfit specifications, as the misspecification rate increases, the classification accuracy decreases. Conversely, in conditions with the Q matrix over-fitting, as the misspecification rate increases, the classification accuracy decreases, all significantly lower compared to cases using the true Q matrix.

In general, within the DINA model, the over-fitting Q matrix and those with balanced-misfit Q matrix specifications have a more significant negative impact on classification accuracy compared to the over-fitting Q matrix conditions. Additionally, it's not just the increase in the misspecification rate, but also the variation in the misspecification pattern that affects classification accuracy. These results are in parallel with the findings of Miao, Miller, and Ren's (2017) study.

### **SUGGESTIONS**

The findings of this study suggest that the accuracy of the item parameters and classifying individuals within the DINA model is influenced by the Q matrix misspecification. In their study, Miao, Miller, and Ren (2017) suggested that in cases where there is uncertainty about whether a particular item measures a specific attribute, over-fitting might be preferable to under-fitting. In practical applications, adhering to test development procedures and utilizing Q matrix validation methods is of utmost importance. It should be remembered that a comprehensive Q matrix validation process requires both statistical expertise and domain knowledge (de la Torre, 2008; Leighton, 2004; Tatsuoka, Corter, & Tatsuoka, 2004). Therefore, to enhance the item parameters and classification accuracy predictions, practitioners are recommended to seek expert opinions to verify whether misspecifications are present in the relationships between items and attributes after the Q matrix has been established. Subsequently, they could employ methods proposed by de la Torre (2008) and de la Torre and Chiu (2016) to validate the Q matrix.

### **ETHICAL TEXT**

In this article, the journal writing rules, publication principles, research and publication ethics, and journal ethical rules were followed. The responsibility belongs to the author for any violations that may arise regarding the article. Since this study is a simulation study, it does not require an ethics committee decision regarding the administration of the data collection tool.

**Author(s) Contribution Rate:** In this study, the contribution rate of the sole author was 100%.

I would like to thank Professor Omay Cokluk-Bokeoglu and The Scientific and Technological Research Council of Turkiye, Science Fellowship and Grant Programmes Directorate, 2211 - National PhD Scholarship Program.

## REFERENCES

- Baker, F. (1993) Sensitivity of the linear logistic test model to misspecification of the weight matrix. *Applied Psychological Measurement*, 17(3), 201-210. <https://doi.org/10.1177/014662169301700301>
- Basokcu, T. O. (2014). Classification accuracy effects of Q-matrix validation and sample size in DINA and G-DINA models. *Journal of Education and Practice*, 5(6), 220-230. <https://doi.org/10.7176/JEP>
- Cassuto, N. (1996). *The performance of the linear logistic test model under different testing conditions* [Unpublished doctoral dissertation]. University of Minnesota.
- Cui, Y., Gierl, M. J., & Chang, H. H. (2012). Estimating classification consistency and accuracy for cognitive diagnostic assessment. *Journal of Educational Measurement*, 49(1), 19-38. <https://doi.org/10.1111/j.1745-3984.2011.00158.x>
- Culpepper, S. A. (2015). Bayesian estimation of the DINA model with gibbs sampling. *Journal of Educational and Behavioral Statistics*, 40(5), 454–476. <https://doi.org/10.3102/1076998615595403>
- DeCarlo, L. T. (2012). Recognizing uncertainty in the Q-matrix via a Bayesian extension of the DINA model. *Applied Psychological Measurement*, 36(6), 447-468. <https://doi.org/10.1177/0146621612449069>
- 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. <https://doi.org/10.1111/j.1745-3984.2008.00069.x>
- De la Torre, J. (2009a). DINA model and parameter estimation: A didactic. *Journal of Educational and Behavioral Statistics*, 34(1), 115–130. <https://doi.org/10.3102/1076998607309474>
- De la Torre, J. (2009b). A cognitive diagnosis model for cognitively-based multiple-choice options. *Applied Psychological Measurement*, 33(3), 163–183. <https://doi.org/10.1177/0146621608320523>
- de la Torre, J., & Chiu, C. Y. (2016). A general method of empirical Q-matrix validation. *Psychometrika*, 81(2), 253-273. <https://doi.org/10.1007/s11336-015-9467-8>
- de la Torre, J. Hong, Y., & Deng W. (2010). Factors affecting the item parameter estimation and classification accuracy of the DINA model. *Journal of Educational Measurement*, 47(2), 227-249. <https://doi.org/10.1111/j.1745-3984.2010.00110.x>
- DiBello, L. V., & Stout, W. (2007). Guest editors' introduction and overview: IRT-based cognitive diagnostic models and related methods. *Journal of Educational Measurement*, 44(4), 285-291. <http://www.jstor.org/stable/20461864>
- Haertel, E. H. (1989). Using restricted latent class models to map the skill structure of achievement items. *Journal of Educational Measurement*, 26(4), 333–352. <https://doi.org/10.1111/j.1745-3984.1989.tb00336.x>
- Henson, R., & Douglas, J. (2005) Test construction for cognitive diagnosis. *Applied Psychological Measurement*, 29(4), 262-277. <https://doi.org/10.1177/0146621604272623>
- Henson, R. A., & Templin, J. L. (2006). *Implications of Q-matrix misspecification in cognitive diagnosis*. Manuscript submitted for publication.
-

- Im, S., & Corter, J. (2011). Statistical consequences of attribute misspecification in the rule space method. *Educational and Psychological Measurement*, 71(4), 712–731. <https://doi.org/10.1177/0013164410384855>
- Jang, E. E., Dunlop, M., Park, G., & van der Boom, E. H. (2015). How do young students with different profiles of reading skill mastery, perceived ability, and goal orientation respond to holistic diagnostic feedback? *Language Testing*, 32(3), 359–383. <https://doi.org/10.1177/0265532215570924>
- Jang, E. E., Kim, H., Vincett, M., Barron, C., & Russell, B. (2019). *Improving IELTS reading test score interpretations and utilisation through cognitive diagnosis model-based skill profiling* (IELTS Research Reports Online Series No. 2). British Council, Cambridge Assessment English and IDP: IELTS Australia. <https://www.ielts.org/research/research-reports/onlineseries-2019-2>
- Junker, B. W., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement*, 25(3), 258–272. <https://doi.org/10.1177/01466210122032064>
- Kunina-Habenicht, O., Rupp, A. A., & Wilhelm, O. (2009). A practical illustration of multidimensional diagnostic skills profiling: comparing results from confirmatory factor analysis and diagnostic classification models. *Studies in Educational Evaluation*, 35(2), 64–70. <https://doi.org/10.1016/j.stueduc.2009.10.003>
- Lee, Y. W., & Sawaki, Y. (2009). Cognitive diagnostic approaches to language assessment: An overview. *Language Assessment Quarterly*, 6(3), 172–189. <https://doi.org/10.1016/j.stueduc.2009.10.003>
- Leighton, J. (2004). Avoid misconception, misuse, and missed opportunities: The collection of verbal reports in educational achievement testing. *Educational Measurement: Issues and Practice*, 23(2), 6–15. <https://doi.org/10.1111/j.1745-3992.2004.tb00164.x>
- Leighton, J. P., & M. J. Gierl (2007). *Cognitive diagnostic assessment for education: Theory and practices*. Cambridge University Press.
- MacDonald, G. T. (2013). *The performance of the linear logistic test model when the Q-matrix is misspecified: a simulation study* [Unpublished doctoral dissertation]. University of South Florida.
- MacDonald, G., & Kromrey, J. (2011). *Linear logistic test model: Using SAS® to simulate the decomposition of item difficulty by algorithm, sample size, cognitive component and time to convergence*. Proceedings of the American Statistical Association’s Joint Statistical Meeting, Social Statistics Section, Miami, FL.
- MacDonald, G., & Kromrey, J. (2012). *The effects of Q-matrix misspecification when employing Proc NL MIXED: A simulation study*. Paper presented at the annual SESUG conference, Durham, NC.
- Miao, G. A. O., Miller, M. D., & Ren, L. I. U. (2017). The impact of Q-matrix misspecification and model misuse on classification accuracy in the generalized DINA model. *Journal of Measurement and Evaluation in Education and Psychology*, 8(4), 391-403. <https://doi.org/10.21031/epod.332712>
- Orlando, M., & Marshall, G. N. (2002). Differential item functioning in a spanish translation of the PTSD checklist: detection and evaluation of impact. *Psychological Assessment*, 14(1), 50-9. <https://doi.org/10.1037/1040-3590.14.1.50>
-

- Ömür-Sünbül, S. Ö., & Kan, A. (2015). Bilişsel tanı modellerinde parametre kestirimini ve sınıflama tutarlılığını etkileyen faktörlerin incelenmesi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 31(4), 778-795. <https://doi.org/10.16986/HUJE.2015014663>
- Qin, C., Zhang, L., Qiu, D., Huang, L., Geng, T., Jiang, H., Ren, Q. & Zhou, J. (2015). Model identification and Q-matrix incremental inference in cognitive diagnosis. *Knowledge-Based Systems*, 86, 66-76. <https://doi.org/10.1016/j.knosys.2015.05.024>
- Romero, S. J., Ordoñez, X. G., Ponsoda, V., & Revuelta, J. (2014). Detection of Q-matrix misspecification using two criteria for validation of cognitive structures under the least squares distance model. *Psicologica: International Journal of Methodology and Experimental Psychology*, 35(1), 149-169.
- Rupp, A. A. (2007). The answer is in the question: A guide for describing and investigating the conceptual foundations and statistical properties of cognitive psychometric models. *International Journal of Testing*, 7(2), 95–125. <https://doi.org/10.1080/15305050701193454>
- Rupp, A., & Templin, J. (2008a). The effects of Q-matrix misspecification on parameter estimates and classification accuracy in the DINA model. *Education and Psychological Measurement*, 68(1), 78–96. <https://doi.org/10.1177/0013164407301545>
- Rupp, A. A., & Templin, J. L. (2008b). Unique characteristics of diagnostic classification models: A comprehensive review of the current state of the art. *Measurement: Interdisciplinary Research and Perspectives*, 6(4), 219–262. <https://doi.org/10.1080/15366360802490866>
- Rupp, A. A., Templin, J., & Henson, R. A. (2010). *Diagnostic measurement: Theory, methods, and applications: Methodology in the social sciences*. Guilford Press.
- Templin, J., & Hoffman, L. (2013). Obtaining diagnostic classification model estimates using Mplus. *Educational measurement: Issues and practice*, 32(2), 37-50. <https://doi.org/10.1111/emip.12010>
- Tatsuoka, K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement* 20, 345–354. <https://www.jstor.org/stable/1434951>
- Tatsuoka, K. (1990). Toward an integration of item-response theory and cognitive error diagnosis. In N. Frederiksen, R. Glaser, A. Lesgold, & Safto, M. (Eds.). *Monitoring skills and knowledge acquisition* (pp. 453-488). Erlbaum.
- Tatsuoka, K. K., Corter, J., & Tatsuoka, C. (2004). Patterns of diagnosed mathematical content and process skills in TIMMS-R across a sample of 20 countries. *American Educational Research Journal*, 41, 901–926. <https://doi.org/10.3102/00028312041004901>

**APPENDIX****Appendix 1.** True Q matrix

Item	Attribute				Item	Attribute			
	1	2	3	4		1	2	3	4
1	1	0	0	0	16	1	0	0	0
2	0	1	0	0	17	0	1	0	0
3	0	0	1	0	18	0	0	1	0
4	0	0	0	1	19	0	0	0	1
5	1	1	0	0	20	1	1	0	0
6	1	0	1	0	21	1	0	1	0
7	1	0	0	1	22	1	0	0	1
8	0	1	1	0	23	0	1	1	0
9	0	1	0	1	24	0	1	0	1
10	0	0	1	1	25	0	0	1	1
11	1	1	1	0	26	1	1	1	0
12	1	1	0	1	27	1	1	0	1
13	1	0	1	1	28	1	0	1	1
14	0	1	1	1	29	0	1	1	1
15	1	1	1	1	30	1	1	1	1



## DINA MODELDE Q MATRİSİNİN HATALI BELİRLENMESİNİN MADDE PARAMETRELERİNE VE SINIFLAMA DOĞRULUĞUNA ETKİSİ<sup>2</sup>

### Öz

Öğrencilere durum belirleme amacıyla uygulanan başarı testleriyle, mevcut durumlarının tam olarak yansıtılabilmesi için, testte yer alan maddeler ile ölçülen beceriler düzeyinde belirlemeler yapılması gerekir. Öğrencilerin ölçülen beceriler bakımından sınıflandırılma geçerliği, test maddelerinin, o maddeyle ölçülen beceriler dikkate alınarak, sınıflandırmada farklı etkilere sahip olmasıyla artırılır. Başarı testindeki madde-beceri ilişkilerinin hatasız kurulması yani Q matrisinin hatasızlığı, sınıflama doğruluğunu, çıkarım kalitesini ve bireyler hakkında verilen kararların doğruluğunu etkiler. Bu araştırmada, bilişsel tanı modellerinden The Deterministic-Input, Noisy “And” Gate (DINA) modelde, Q matrisinin hatalı belirlenmesinin, 30 maddelik test uzunluğuna sahip simülatif olarak oluşturulan veri setlerinde madde parametrelerine ve bireylerin sınıflandırılmasına etkisinin incelenmesi amaçlanmıştır. Çalışmada, hatalı belirleme oranının %5, %7.5 ve %10 ve hatalı belirleme düzeninin eksik, fazla ve dengeli olduğu hatalı Q matrislerine ilişkin incelemeler yapılmıştır. Sümülatif verilerin üretilmesinde ve analizinde R ve Mplus yazılımları kullanılmıştır. Hatalı belirlenen Q matrisleri ile yapılan madde parametresi kestirimleri ve sınıflamalar, veri setine uygun hatasız Q matrisi ile yapılan kestirimlerle karşılaştırılmıştır. Bu çalışma kapsamında incelenen koşullar kapsamında, Q matrisinin hatalı belirlenmesinin bireylerin sınıflandırmalarını farklılaştırdığı bulunmuştur. Q matrisinin eksik belirlendiği koşullarda, hatalı belirleme oranı arttıkça sınıflama doğruluğunun azaldığı; Q matrisinin fazla belirlendiği koşullarda, hatalı belirleme oranı arttıkça tümü hatasız Q matrisi ile yapılanlardan çok düşük olmak üzere sınıflama doğruluğunun arttığı ve Q matrisinin dengeli hatalı belirlendiği koşullarda, hatalı belirleme oranı arttıkça sınıflama doğruluğunun azaldığı bulunmuştur. Fazla belirleme yapılan ve dengeli hatalı belirleme yapılan Q matrisleri, eksik belirleme yapılan Q matrislerine göre sınıflama doğruluğu üzerinde daha ciddi olumsuz etkiye sahiptir. Eksik belirlemeler maddenin kaydırma parametresi; fazla belirlemeler tahmin parametresi değerlerini artırmıştır. Dengeli hatalı belirlemeler ise maddede eksik belirleme yapıldıysa kaydırma, fazla belirleme yapıldıysa tahmin, hem eksik hem fazla belirleme yapıldıysa iki parametreyi de artırmıştır. Eksik belirleme yapılan maddelerin kaydırma parametrelerindeki yükselmeler, fazla belirleme yapılan maddelerin tahmin parametrelerindeki yükselmelerden oldukça fazladır.

**Anahtar kelimeler:** Eksik tanımlama, fazla tanımlama, dengeli hatalı tanımlama.

---

<sup>2</sup> Bu makalenin bir kısmı, yazar tarafından hazırlanan doktora tezinden oluşturulmuştur.

## GİRİŞ

Eğitim-öğretimde her alanda ve her sınıf düzeyinde, öğrencilerin öğrenmelerine ilişkin durum belirleme amacıyla başarı testleri uygulanmaktadır. Başarı testleri ile madde düzeyinde ya da testin tamamı için öğrenci performansına ilişkin bir sınıflama yapılır. Bu sınıflamalar yapılırken, maddenin doğru yanıtlanması için hangi becerilere sahip olunması gerektiği için katılarak, testteki her bir maddenin öğrencilerin sınıflandırılmasında farklı etkiye sahip olması sağlanabilirse, sınıflamaların geçerliği ve verilecek kararların doğruluğu artacaktır. Bunu sağlamanın bir yolu Bilişsel Tanı Modellerinin (BTM) kullanmasıdır.

Tamamı özel bir gizil sınıf modeli türü olan BTM’lerde, her öğrencinin test maddelerine verdikleri yanıtlara göre beceri profili oluşturulur ve eş beceri profilli öğrenci grupları belirlenir. Ortaya çıkabilecek grupların (gizil sınıfların) sayısı; beceri sayısı ve becerilerin birbirine bağımlılıklarının toplam sayısı ile ilişkili olarak önceden belirlidir. Bu nedenle, BTM’ler kısıtlı gizil sınıf modelleridir (Haertel,1989; Jang, Kim, Vincett, Barron ve Russell, 2019; Rupp, Templin ve Henson, 2010).BTM’lerde, madde ile ölçülen beceriler hatasız belirlendiği ve tanımlandığında, öğrencilerin her bir becerideki ve testin tamamındaki performansları daha ayrıntılı, doğru ve geçerliği yüksek bir şekilde belirlenebilir. Öğrencilerin bir sonraki eğitim aşamasına geçişte güçlü ve zayıf yönleri ve hazırbulunuşluk düzeyi belirlenebilir, eğitim sürecine yön verilebilir (de la Torre, 2009a; Henson ve Templin, 2006; Leighton ve Gierl, 2007; Rupp, 2007; Rupp ve Templin, 2008a; 2008b; Rupp, Templin ve Henson, 2010).

BTM’lerde madde ile ölçülen becerilerin tanımlanması Q matrisi ile gerçekleştirilir. Q matrisi, satırları maddeleri ve sütunları becerileri ifade edecek şekilde, madde-beceri ilişkilerinin sayısal olarak (1 ve 0’lardan oluşan  $J \times K$  matrisi) kodlanması ile oluşturulur (Rupp ve diğerleri, 2010). Her madde için, ilgili maddeye ( $j$  maddesi) doğru yanıt verilebilmesi için o beceriye ( $k$  becerisi) sahip olmak gerekiyorsa, matriste ilgili maddenin satırıyla becerinin sütununun kesişim hüccesine ( $j$ ’inci satır ve  $k$ ’inci sütunun kesiştiği  $q_{jk}$  hüccesi) “1”; gerekmiyorsa “0” yazılır. Q matrisi, kullanılan başarı testinin yapısını yansıtır ve bu başarı testi kullanılarak alınacak kararların kalitesinin temel yapıtaşdır (de la Torre, 2009a; Rupp ve Templin, 2008a; Rupp Templin ve Henson, 2010). Q matrisinin teorik ve ampirik sağlamlığı, modellerin tanısallık gücünü belirlediğinden, Q matrisi BTM’ler ile yürütülen çalışmalar için oldukça önemlidir (Lee ve Sawaki, 2009).

BTM’lerde alınacak kararların doğruluğu, maddelerin nitelikli ve kriterlere uygun yazıldığı ve testin kapsam geçerliğinin sağlandığı durumlarda, bireylerin sınıflandırılmasının geçerliğine bağlıdır. Q matrisinin hatalı oluşturulması da, birey profilinin yanlış olmasına neden olmaktadır (de la Torre, 2009a; Im ve Corter, 2011; Qin, Zhang, Qiu, Huang, Geng, Jiang, Ren ve Zhou, 2015; Romero, Ordoñez, Ponsoda ve Revuelta, 2014; Rupp ve Templin, 2008a; 2008b; Rupp ve diğerleri, 2010; Tatsuoka, 1983). BTM’lerde parametre tahminlerinin doğruluğu Q matrisinin hatasızlığına bağlıdır (Baker, 1993). Q matrisi genellikle uzman görüşüyle belirlendiği için hatalı tanımlamalar yapılabilir (DeCarlo, 2012). Öğrencilerin yer aldığı sınıfların doğru belirlenmesi, bu sınıflandırmalar dikkate alınarak öğrenciler hakkında bazı kararlar alındığından son derece önemlidir (Kunina-Habenicht, Rupp ve Wilhelm, 2012).

Q matrisinin hatalı belirlenmesinin madde parametrelerine ve bireylerin sınıflandırılmasına etkisini incelemeyi amaçlayan bu araştırmada, BTM'lerden DINA (The Deterministic-Input, Noisy "And" Gate; de la Torre, 2009; Haertel, 1989; Junker ve Sijtsma, 2001) model ile çalışılmıştır.

### DINA Model

Becerilerin ve yanıtların ikili puanlandığı DINA model, telafi edici olmayan bir modeldir. Öğrencinin bir maddeyi doğru yanıtlanabilmesi için, o madde ile ilişkili olan bütün becerilere sahip olması gerekir. Öğrencinin gerekli bir beceriye sahip olmaması durumunda, bu eksiklik diğer bir beceriye sahip olması ile telafi edilemez (de la Torre, 2009b; Henson ve Templin, 2006; Rupp ve Templin, 2008a; Rupp ve Templin, 2008b; Rupp Templin ve Henson, 2010).

DINA modelde maddeler için "kaydırma (slip/s<sub>j</sub>)" ve "tahmin (guess/g<sub>j</sub>)" isimli iki tipik olmayan yanıt olasılığı hesaplanır. Öğrencinin maddeyle ilişkili tüm becerilere sahip olup maddeye yanlış yanıt verme olasılığı kaydırma parametresi ile, en az bir beceriye sahip olmayıp maddeye doğru yanıt verme olasılığı tahmin parametresi ile ifade edilir (Rupp ve Templin, 2008a; Rupp ve diğerleri, 2010). İlgili formüller Eşitlik 1 ve 2'deki gibi gösterilir.

$$s_j = P(X_{ij} = 0 | \eta_{ij} = 1) \quad (1)$$

$$g_j = P(X_{ij} = 1 | \eta_{ij} = 0) \quad (2)$$

$s_j$  : Maddeyle ilişkili becerilere sahip öğrencinin  $j$  maddesine yanlış yanıt verme (yanlış pozitif) olasılığı,

$g_j$  : Maddeyle ilişkili becerilere sahip olmayan öğrencinin  $j$  maddesine doğru yanıt verme (doğru pozitif) olasılığı

$X_{ij}$  :  $i$  öğrencisinin  $j$  maddesine ilişkin puanı (0,1), ( $i= 1, \dots, I$  ve  $j= 1, \dots, J$ ).

$\eta_{ij}$  : Belirleyici gizil yanıt

DINA modelde gizil yanıt Eşitlik 3'teki gibi gösterilir.

$$\eta_{ij} = \prod_{k=1}^K \alpha_{ik}^{q_{jk}} \quad (3)$$

$\alpha_i = \{ \alpha_{ik} \}$  :  $i$  öğrencisinin  $k$  becerisine sahip olup olmadığını gösteren ikili beceri vektörü ( $k= 1, \dots, K$ )

$q_{jk}$  :  $j$  maddesi ve  $k$  becerisine ilişkin Q matrisindeki ikili girdi (1,0)

İlişkili tüm becerilere sahip bir öğrencinin, maddeyi doğru yanıtlama olasılığı Eşitlik 4 ve Eşitlik 5'te verilmiştir (De la Torre, 2009a; Rupp ve diğerleri, 2010).

$$P_j(\alpha_i) = P(X_{ij} = 1 | \eta_{ij}) = (1 - s_j)^{\eta_{ij}} g_j^{1-\eta_{ij}} \quad (4)$$

$$P(X_1 = 1 | \alpha_1 \alpha_2) = \frac{\exp(\lambda_{1,0} + \lambda_{1,(1)}\alpha_1 + \lambda_{1,1,(2)}\alpha_2 + \lambda_{1,2,(1,2)}\alpha_1\alpha_2)}{1 + \exp(\lambda_{1,0} + \lambda_{1,1,(1)}\alpha_1 + \lambda_{1,1,(2)}\alpha_2 + \lambda_{1,2,(1,2)}\alpha_1\alpha_2)} \quad (5)$$

$P_j(\alpha_i)$ :  $\alpha_i$  beceriler vektörüne sahip  $i$  öğrencisinin  $j$  maddesini doğru yanıtlama olasılığı

DINA modelde, öğrencilerin beceri profilleri oluşturulurken önce birey parametreleri hesaplanır. Madde ve yapısal parametrelerin de hesaplandığı bu süreç, Maksimum Olasılık yaklaşımı ile gerçekleştirilir. Öğrencilerin yerleştirileceği gizil sınıfların ayrışımı için, ön bir ayırım hesaplanır ve bu ayırım kullanılarak madde parametreleri hesaplanır. Hesaplanan bu parametreler bilinen madde parametreleri olarak ele alınıp sonsal gizil sınıf ayrımı hesaplanır. Hesaplama algoritması, madde parametrelerinin her aşamada tekrar hesaplandığı yinelemeli bir iyileştirme süreci takip eder. Gerekli yakınsama sağlanana kadar bir önceki adımdaki kestirimler, güncel kestirimlerle karşılaştırılır (De la Torre, 2009a; Rupp ve diğerleri, 2010).

Önceki araştırmalarda ortaya konan (Baker, 1993; de la Torre, 2008; Im ve Corter, 2011; Kunina-Habenicht, Rupp ve Wilhelm, 2012; MacDonald, 2013; MacDonald ve Kromrey, 2012; Rupp ve Templin, 2008a) Q matrisi oluşturulurken yapılması olası farklı hatalı belirleme durumları incelendikten sonra bu çalışmada, Q matrisine ilişkin hatalı belirleme oranının (%5, %7.5 ve %10) ve hatalı belirleme düzeninin (eksik, fazla ve dengeli) olduğu durumların parametre kestirimine ve sınıflama doğruluğuna etkisinin incelenmesine karar verilmiştir. Q matrisinin hatalı belirlenmesinin madde parametrelerini ve bireylerin sınıflandırılmasını nasıl etkilediğine ilişkin teorik belirlemeler, BTM'ler yoluyla alınacak kararların doğruluğu açısından önemlidir. Bu araştırma ile yapılan tespitlerin araştırmacılara ve uygulayıcılara yol gösterilebilecek olması ve yapılan belirlemeler uygulamaya yansıtıldığında BTM'ler ile alınan kararların doğruluğuna katkı sağlayacak olması bu çalışmanın önemini yansıtmaktadır.

## YÖNTEM

DINA modelde Q matrisin hatalı belirlenmesinin madde parametrelerinin kestirim iyiliğine ve bireylerin sınıflandırılmasına etkisi simülatif veriler kullanılarak incelendiğinden bu çalışma temel araştırma modelindedir.

## Verilerin Üretilmesi

Ölçme aracı ile ölçülen beceri sayısı, ölçme aracının uygulandığı örneklem büyüklüğü ve testin uzunluğu sınıflama tutarlılığı açısından önemli faktörlerdir. Yapılan çalışmalar incelenerek; beceri sayısının dört (Henson ve Douglas, 2005; Kunina-Habenicht ve diğerleri, 2012; Ömür-Sünbül ve Kan, 2015; Rupp ve Templin, 2008a), örneklem büyüklüğünün 1000 (Baker, 1993; Cassuto, 1996; De la Torre, Hong ve Deng, 2010; Kunina-Habenicht ve diğerleri, 2012; MacDonald ve Kromrey, 2011; Orlando ve Marshall, 2002; Ömür-Sünbül ve Kan, 2015; Tsutakawa ve Johnson, 1990;), test uzunluğunun 30 madde (Cassuto, 1996; Henson ve Templin, 2006; Kunina-Habenicht ve diğerleri, 2012; MacDonald, 2013; Ömür-Sünbül ve Kan, 2015; Tatsuoka, 1990; Templin ve diğerleri, 2009) olmasına karar verilmiştir.

Madde sayısına karar verirken olası beceri sınıfları da dikkate alınmıştır. Birbirinden bağımsız dört becerinin ölçüldüğü durumda olası  $2^4$  sayıda yani  $2^4=16$  beklenen beceri sınıfı, bu becerileri ölçebilen  $2^4-1$  sayıda yani  $2^4-1=15$  farklı madde sözkonusudur. Her bir alt beceri profili, 0 ve 1'in diferansiyel kombinasyonları ile çok boyutlu bir profili temsil eder (Jang, Dunlop, Park ve van der Boom, 2015). Olası toplam madde sayısı, beklenen beceri profillerinin sayısından bir eksiktir çünkü başarı testinde hiçbir becerinin ölçülmediği madde (0000 ile gösterilen bir satır) yer almayacaktır. Kullanılan başarı testinde, aynı becerileri ölçen birden fazla madde yer alabileceğinden, test  $2^4-1$ 'den daha fazla madde içerebileceği gibi, mutlaka olası tüm beceri örüntülerini içermek zorunda da değildir (Rupp ve diğerleri, 2010).

Bu çalışmada ölçme duyarlılığı açısından, maddeler ve Q matrisi oluşturulurken, her bir beceri ile ilişkili madde sayısının yaklaşık olarak aynı olmasına ve olası tüm beceri profillerinin temsil edilmesine (DiBello ve Stout, 2007; Kunina-Habenicht ve diğerleri, 2012; Rupp ve Templin, 2008a) dikkat edilerek, 15 farklı örüntüyü temsil eden ilk 15 madde oluşturulmuş ve bunlar aynen tekrar edilerek toplam 30 maddelik hipotetik ölçme aracı incelenmiştir.

Çalışmada Q matrisin hatalı belirlendiği koşulların belirlenmesinde önceki çalışmalar (Baker, 1993; Kunina-Habenicht ve diğerleri, 2012; MacDonald, 2013; MacDonald ve Kromrey, 2012; Rupp ve Templin, 2008a) incelenmiş ve gerçek yaşam durumlarında karşılaşılma olasılığı yüksek olan koşulları yansıtabilecek şekilde, hatalı belirleme oranının %5, %7.5 ve %10 olmasına ve hatalı belirleme düzeninin ise eksik, fazla ve dengeli olarak iki şekilde oluşturulmasına karar verilmiştir. Hatalı belirleme oranının %5 olması Q matrisinde yer alan altı hücrede, %7.5 olması dokuz, %10 olması 12 hücrede hatalı belirleme yapıldığı anlamına gelmektedir. Hatalı belirleme düzeninin eksik olarak ifade edilmesi, Q matrisinde yer alan hücrelerdeki ilgili orandaki 1'lerin 0'lar ile değiştirilmesi, dengeli olması ise ilgili oranın yarısı kadar sayıda 0'ın 1 ile, yarısı kadar 1'in ise 0 ile değiştirilmesi anlamına gelir. Hatalı Q matrisleri oluşturulurken, hatalı tanımlama yapılacak madde ve beceriler ilk 15 madde içindeki hücrelerden seçkisiz olarak belirlenmiştir. Çalışmada incelenen Q matrislerine ilişkin, kullanım kolaylığı sağlaması açısından yapılan kodlamalar, bunların kısa açıklamaları, madde başına ortalama beceri sayısı, beceri başına ortalama madde sayısı ve Q matriste yer alan 1 ve 0'ların toplam değişimi ile ilgili bilgiler Tablo 1'de sunulmuştur.

**Tablo 1.** Q Matrislerine İlişkin Bilgiler

Açıklama	Koşul	Madde Başına Ortalama Beceri Sayısı	Beceri Başına Ortalama Madde Sayısı	Toplam Değişim (0'ın 1'e)	Toplam Değişim (1'in 0'a)
Hatasız Q matrisi	Q0	2.13	16.00	0	0
%5 eksik tanımlama yapılan Q matrisi	Q1	1.93	14.50	0	6
%7.5 eksik tanımlama yapılan Q matrisi	Q2	1.83	13.75	0	9
%10 eksik tanımlama yapılan Q matrisi	Q3	1.73	13.00	0	12
%5 fazla hatalı tanımlanan Q matrisi	Q4	2.33	17.50	6	0
%7.5 fazla hatalı tanımlanan Q matrisi	Q5	2.43	18.25	9	0
%10 fazla hatalı tanımlanan Q matrisi	Q6	2.53	19.00	12	0
%5 dengeli hatalı tanımlanan Q matrisi	Q7	2.13	16.00	3	3
%7.5 dengeli hatalı tanımlanan Q matrisi	Q8	2.10	15.75	4	5
%10 dengeli hatalı tanımlanan Q matrisi	Q9	2.13	16.00	6	6

Tablo 1’de yer alan madde başına ortalama beceri sayısı, Q matrisindeki 1 olan hücre sayısının madde sayısına (30) oranıdır. Beceri başına ortalama madde sayısı ise Q matrisindeki 1 olan hücre sayısının beceri sayısına (4) oranıdır.

Veriler R yazılımında yer alan “dina” paketi (Culpepper, 2015) kullanılarak, Ek 1’de sunulan hatasız Q matrisine ve DINA modele uygun olacak şekilde, g ve s parametreleri 0.00’a sabitlenerek üretilmiştir. İlk olarak 1000 kişilik, 0-1 verilerden oluşan ve 16 farklı olası beceri profilinde yaklaşık eşit sayıda öğrenci yer alan yanıtlayıcı beceri profili oluşturulmuştur. Daha sonra bu yanıtlayıcı beceri profiline uygun yanıt örüntüleri oluşturulmuştur. Her bir durum için 100 replikasyon yapılmıştır.

### Verilerin Analizi

Analizler R ve Mplus yazılımları kullanılarak yapılmıştır. Mplus’ta DINA tahminleri için, Templin ve Hoffman (2013) tarafından gizli sınıf modeli için oluşturulan sözdizimi, belirli kısıtlamalar kullanılarak değiştirilmiştir. Replikasyonlarla oluşturulan yanıt örüntülerinin her biri için, hem hatasız Q matrisi ile hatasız sınıflamalar, hem de her bir hatalı Q matrisi kullanılarak hatalı sınıflamalar belirlenmiştir. Kestirimlerde 1000 iterasyon yapılmış ve yakınsama kriteri 0.001 olarak seçilmiştir.

Sınıflama doğruluğu, gözlenen verilerdeki sınıflama durumu yani gerçek gizil sınıflar ile analiz sonucunda elde edilen gizli sınıfların uyuma derecesidir (Cui ve diğerleri, 2012). Hatasız Q matrisi ile yapılan analizlerden elde edilen yani verilerin üretilmesi esnasında tanımlanan gizli sınıflar gerçek gizil sınıflar olarak kullanılmıştır. Gerçek gizil sınıflar ile, hatalı Q matrisleri ile yapılan analizler sonucunda MLE yöntemi kullanılarak tahmin edilen gizil sınıflar her bir birey için karşılaştırılmıştır. Bu iki gizil sınıf tutarlı ise bireyin doğru sınıflandırıldığını temsil etmek üzere bu bireye “1” değeri atanmış tersi durumda ise yanlış sınıflandırılmayı temsil eden “0” değeri atanmıştır. Tüm bireyler ve tüm replikasyonlar üzerinden bu 1-0 değerlerinin ortalaması alınarak, her koşul için genel doğru sınıflandırma oranları hesaplanmıştır.

### BULGULAR

Bu bölümde simülatif olarak üretilen ve 30 madde içeren ölçme araçlarına ilişkin hatasız Q matrisi ve hatalı Q matrisleriyle elde edilen bulgular sunulmuştur.

Q matrisinin hatasız olduğu durum ile Q matrisinin eksik belirlendiği durumlarda kestirilen madde parametreleri Tablo 2’de sunulmuştur.

Tablo 2 incelendiğinde, eksik belirleme yapılan maddelerin (renklendirilerek gösterilmiştir) tahmin parametrelerinin ve bunlara ilişkin standart hata değerlerinin değişmediği, kaydırma parametrelerinin ve bunlara ilişkin standart hata değerlerinin arttığı görülmektedir.

**Tablo 2.** Eksik Belirlenen Q Matrisleriyle Kestirilen Madde Parametreleri

Madde	Q0		Q1		Q2		Q3	
	g	s	g	s	g	s	g	s
1	0.012011	0.006900	0.000041	0.000039	0.000041	0.000040	0.000040	0.000040
2	0.011720	0.007101	0.000038	0.000042	0.000039	0.000039	0.000040	0.000040
3	0.012258	0.006797	0.000040	0.000039	0.000041	0.000039	0.000041	0.000039
4	0.012935	0.006413	0.000044	0.000037	0.000044	0.000037	0.000044	0.000036
5	0.006489	0.011160	0.000040	0.487064	0.000041	0.487105	0.000040	0.487059
6	0.006828	0.010827	0.000027	0.000072	0.000040	0.455277	0.000041	0.455312
7	0.006673	0.010969	0.000028	0.000075	0.000028	0.000078	0.000043	0.506586
8	0.005880	0.011855	0.000026	0.000086	0.000038	0.531560	0.000041	0.551707
9	0.006762	0.010803	0.000044	0.499083	0.000039	0.448053	0.000040	0.448044
10	0.006853	0.010896	0.000042	0.464016	0.000040	0.463959	0.000028	0.000073
11	0.003713	0.016189	0.000027	0.499901	0.000027	0.529256	0.000027	0.499957
12	0.002464	0.011658	0.000027	0.505639	0.000028	0.512943	0.000027	0.512834
13	0.003927	0.015646	0.000023	0.000146	0.000027	0.479313	0.000044	<b>0.743072</b>
14	0.003629	0.016529	0.000027	0.538618	0.000023	0.000163	0.000027	0.545403
15	0.002298	0.024557	0.000021	0.000319	0.000023	0.503996	0.000028	<b>0.771180</b>
16	0.012006	0.006933	0.000040	0.000040	0.000041	0.000040	0.000039	0.000041
17	0.006896	0.004577	0.000039	0.000041	0.000039	0.000041	0.000039	0.000042
18	0.012251	0.006751	0.000041	0.000040	0.000042	0.000038	0.000041	0.000040
19	0.012935	0.006411	0.000043	0.000037	0.000043	0.000037	0.000043	0.000037
20	0.006485	0.011221	0.000028	0.000078	0.000027	0.000077	0.000028	0.000078
21	0.006838	0.010740	0.000028	0.000073	0.000027	0.000072	0.000027	0.000073
22	0.006669	0.010957	0.000027	0.000076	0.000027	0.000076	0.000027	0.000076
23	0.005872	0.011904	0.000026	0.000087	0.000026	0.000087	0.000026	0.000086
24	0.006766	0.010860	0.000027	0.000074	0.000028	0.000075	0.000027	0.000075
25	0.006852	0.010736	0.000028	0.000074	0.000028	0.000074	0.000027	0.000070
26	0.003712	0.016268	0.000023	0.000152	0.000022	0.000158	0.000022	0.000158
27	0.003783	0.015930	0.000023	0.000151	0.000023	0.000151	0.000023	0.000148
28	0.003929	0.015551	0.000024	0.000146	0.000023	0.000147	0.000023	0.000147
29	0.003633	0.016427	0.000023	0.000161	0.000023	0.000161	0.000023	0.000163
30	0.002298	0.024666	0.000021	0.000333	0.000022	0.000329	0.000021	0.000318

g: tahmin parametresi, s: kaydırma parametresi

Q matrisinin hatasız olduğu durum ile Q matrisinin fazla belirlendiği durumlarda kestirilen madde parametreleri Tablo 3’te sunulmuştur.

Tablo 3 incelendiğinde, fazla belirleme yapılan maddelerin (renklendirilerek gösterilmiştir) kaydırma parametrelerinin ve bunlara ilişkin standart hata değerlerinin değişmediği, tahmin parametrelerinin ve bunlara ilişkin standart hata değerlerinin arttığı görülmektedir.

Q matrisinin hatasız olduğu durum ile Q matrisinin dengeli hatalı belirlendiği durumlarda kestirilen madde parametreleri Tablo 4’te sunulmuştur.

**Tablo 3.** Fazla Belirlenen Q Matrisleriyle Kestirilen Madde Parametreleri

Madde	Q0		Q4		Q5		Q6	
	g	s	g	s	g	s	g	s
1	0.012011	0.006900	0.000041	0.000041	0.000041	0.000039	0.321910	0.000077
2	0.011720	0.007101	0.339005	0.000088	0.301799	0.000074	<b>0.413553</b>	0.000149
3	0.012258	0.006797	0.367551	0.000088	0.367456	0.000086	0.000041	0.000038
4	0.012935	0.006413	0.000043	0.000038	<b>0.471152</b>	0.000154	<b>0.466956</b>	0.000145
5	0.006489	0.011160	0.000027	0.000078	0.000026	0.000078	0.148145	0.000152
6	0.006828	0.010827	0.000027	0.000072	0.000028	0.000075	0.000028	0.000075
7	0.006673	0.010969	0.000027	0.000073	0.000027	0.000075	0.148693	0.000145
8	0.005880	0.011855	0.000026	0.000087	0.000026	0.000088	0.000026	0.000088
9	0.006762	0.010803	<b>0.222890</b>	0.000328	<b>0.222893</b>	0.000313	0.000028	0.000075
10	0.006853	0.010896	0.000028	0.000073	0.171486	0.000164	0.157968	0.000142
11	0.003713	0.016189	0.071455	0.000325	0.071428	0.000322	0.071465	0.000323
12	0.002464	0.011658	0.074628	0.000329	0.000023	0.000149	0.074649	0.000323
13	0.003927	0.015646	0.000023	0.000144	<b>0.082119</b>	0.000328	0.082126	0.000316
14	0.003629	0.016529	0.000023	0.000162	0.000023	0.000157	0.067149	0.000318
15	0.002298	0.024557	0.000021	0.000338	0.000021	0.000312	0.000021	0.000328
16	0.012006	0.006933	0.000040	0.000040	0.000041	0.000039	0.000046	0.000044
17	0.006896	0.004577	0.000046	0.000045	0.000045	0.000046	0.000046	0.000046
18	0.012251	0.006751	0.000048	0.000043	0.000047	0.000045	0.000041	0.000039
19	0.012935	0.006411	0.000043	0.000038	0.000051	0.000041	0.000051	0.000041
20	0.006485	0.011221	0.000027	0.000079	0.000027	0.000075	0.000026	0.000076
21	0.006838	0.010740	0.000027	0.000075	0.000028	0.000073	0.000027	0.000075
22	0.006669	0.010957	0.000027	0.000076	0.000027	0.000075	0.000027	0.000074
23	0.005872	0.011904	0.000026	0.000087	0.000027	0.000089	0.000026	0.000087
24	0.006766	0.010860	0.000028	0.000075	0.000027	0.000073	0.000027	0.000074
25	0.006852	0.010736	0.000028	0.000072	0.000027	0.000073	0.000027	0.000073
26	0.003712	0.016268	0.000023	0.000161	0.000023	0.000152	0.000023	0.000155
27	0.003783	0.015930	0.000023	0.000154	0.000023	0.000154	0.000023	0.000150
28	0.003929	0.015551	0.000023	0.000142	0.000024	0.000145	0.000024	0.000148
29	0.003633	0.016427	0.000023	0.000159	0.000023	0.000161	0.000023	0.000161
30	0.002298	0.024666	0.000022	0.000329	0.000022	0.000320	0.000021	0.000336

g: tahmin parametresi, s: kaydırma parametresi

Tablo 4 incelendiğinde, dengeli hatalı belirleme durumunda; fazla belirleme yapılan maddelerin tahmin parametrelerinin ve bunlara ilişkin standart hata değerlerinin arttığı, eksik belirleme yapılan maddelerin kaydırma parametrelerinin ve bunlara ilişkin standart hata değerlerinin arttığı, bir beceride eksik bir beceride fazla belirleme yapılan maddelerin her iki parametre ve buna ilişkin standart hata değerlerinin arttığı görülmektedir.

Her bir bireyin, hatasız durumdaki beceri profili ile hatalı Q matrisi kullanılarak yapılan kestirimlerden elde edilen beceri profilinin karşılaştırılması ile elde edilen sınıflama doğruluğu değerleri ve kestirimlerin uyum iyiliği indeksleri Tablo 5’de sunulmuştur.



**Tablo 4.** Dengeli Hatalı Belirlenen Q Matrisleriyle Kestirilen Madde Parametreleri

Madde	Q0		Q7		Q8		Q9	
	g	s	g	s	g	s	g	s
1	0.012011	0.006900	0.000041	0.000040	0.000040	0.000040	0.000040	0.000040
2	0.011720	0.007101	0.000040	0.000041	0.000040	0.000041	0.338976	0.000084
3	0.012258	0.006797	0.000040	0.000039	0.000041	0.000038	0.000040	0.000039
4	0.012935	0.006413	0.000043	0.000037	0.000043	0.000037	0.370377	0.000074
5	0.006489	0.011160	0.148088	0.000153	0.257507	0.742001	0.145145	0.000151
6	0.006828	0.010827	0.000028	0.000073	0.000027	0.000073	0.000028	0.000075
7	0.006673	0.010969	0.176365	0.492751	0.185206	0.512892	0.000041	0.469259
8	0.005880	0.011855	0.000040	0.531538	0.139193	0.529191	0.000041	0.551690
9	0.006762	0.010803	0.000027	0.000073	0.000028	0.000074	0.189636	0.505434
10	0.006853	0.010896	0.000028	0.000072	0.171466	0.000159	0.157952	0.000141
11	0.003713	0.016189	0.000023	0.000153	0.000028	0.500052	0.000026	0.439276
12	0.002464	0.011658	0.000027	0.505530	0.000023	0.000153	0.074657	0.000320
13	0.003927	0.015646	0.000023	0.000146	0.000023	0.000147	0.000023	0.000146
14	0.003629	0.016529	0.067176	0.000317	0.000022	0.000160	0.000027	0.456513
15	0.002298	0.024557	0.000021	0.000325	0.000021	0.000328	0.000023	0.504040
16	0.012006	0.006933	0.000040	0.000040	0.000041	0.000040	0.000040	0.000039
17	0.006896	0.004577	0.000039	0.000039	0.000040	0.000040	0.000046	0.000047
18	0.012251	0.006751	0.000041	0.000039	0.000040	0.000039	0.000041	0.000038
19	0.012935	0.006411	0.000044	0.000036	0.000045	0.000038	0.000051	0.000042
20	0.006485	0.011221	0.000027	0.000079	0.000027	0.000077	0.000027	0.000075
21	0.006838	0.010740	0.000028	0.000073	0.000027	0.000074	0.000028	0.000074
22	0.006669	0.010957	0.000027	0.000074	0.000028	0.000076	0.000027	0.000075
23	0.005872	0.011904	0.000026	0.000089	0.000026	0.000085	0.000026	0.000087
24	0.006766	0.010860	0.000027	0.000076	0.000028	0.000077	0.000027	0.000075
25	0.006852	0.010736	0.000028	0.000073	0.000027	0.000074	0.000028	0.000072
26	0.003712	0.016268	0.000023	0.000154	0.000023	0.000156	0.000023	0.000158
27	0.003783	0.015930	0.000023	0.000151	0.000023	0.000152	0.000023	0.000151
28	0.003929	0.015551	0.000022	0.000144	0.000023	0.000141	0.000024	0.000142
29	0.003633	0.016427	0.000023	0.000159	0.000022	0.000161	0.000023	0.000156
30	0.002298	0.024666	0.000021	0.000323	0.000021	0.000327	0.000021	0.000315

g: tahmin parametresi, s: kaydırma parametresi

**Tablo 5.** Kestirilen Sınıflama Doğrulukları ve Model Uyum İndeksleri

Koşul	Sınıflama Doğruluğu	LL	AIC	BIC	Entropi
Q1	0.941	-11647.863	23447.727	23820.716	0.997
Q2	0.871	-11844.227	23838.454	24206.535	0.971
Q3	0.741	-12024.674	24199.349	24567.431	0.972
Q4	0.667	-11172.531	22497.062	22870.052	0.980
Q5	0.691	-11884.880	23919.759	24287.841	0.966
Q6	0.736	-12068.489	24288.978	24661.968	0.895
Q7	0.871	-12315.613	24779.226	25142.400	0.939
Q8	0.731	-11711.180	23572.363	23940.444	0.974
Q9	0.626	-11576.394	23302.788	23670.869	0.978

LL = log-likelihood, AIC = Akaike bilgi kriteri, BIC = Bayesçi bilgi kriteri

Tablo 5 incelendiğinde, Q matrisinin eksik belirlendiği koşullarda (Q1, Q2 ve Q3), hatalı belirleme oranı arttıkça sınıflama doğruluğunun azaldığı, Q matrisinin fazla belirlendiği koşullarda (Q4, Q5 ve Q6), hatalı belirleme oranı arttıkça sınıflama doğruluğunun arttığı ve Q matrisinin dengeli hatalı belirlendiği koşullarda (Q7, Q8 ve Q9), hatalı

belirleme oranı arttıkça sınıflama doğruluğunun azaldığı görülmektedir. Entropinin 1'e yakın değerleri ve uyum iyiliği indekslerinin küçük değerleri iyi uyumun göstergesi olduğundan, sınıflama doğruluğu bulgularına tutarlı olarak Q matrisinin eksik ve dengeli hatalı belirlendiği koşullarda hatalı belirleme oranı arttıkça uyum iyiliği indekslerinin kötüleştiği, Q matrisinin fazla belirlendiği koşullarda ise hatalı belirleme oranı arttıkça uyum iyiliği indekslerinin iyileştiği bulunmuştur.

### **TARTIŞMA ve SONUÇ**

Bu çalışmada DINA modelde Q1-Q9 hatalı belirleme koşullarının madde parametrelerine ve bireylerin sınıflandırılmasına etkisi incelenmiştir. İncelenen her koşulda, Q matrisin hatalı belirlenmesi madde parametrelerini ve bireylerin sınıflandırmalarını etkilemektedir.

DINA modelde kaydırma ve tahmin parametreleri ne kadar düşük olursa (sıfıra yakın), bu durum modelin iyi tanımlandığının, yanıtların rasgele olmadığını ve maddenin veriye iyi uyum gösterdiğinin göstergesidir. Maddeye özgü model uyumu sağlanmış olur (De la Torre, 2009a ; Rupp ve Templin, 2008a). Kaydırma ve tahmin parametresi değerlerinin sıfırdan yüksek olması, Q matrisinde tanımlanan becerilerin katılımcıların yanıt örüntüsünü açıklamada yetersiz kaldığını ve maddelerin yanıtlanmasının farklı bir strateji içerdiğini göstermektedir (De la Torre ve Douglas, 2004). Parametreler ne kadar düşük değerlere sahipse, ilgili becerilere sahip bireylerin maddeye doğru yanıt verme olasılığı, sahip olmayan bireylerin de maddeye yanlış yanıt verme olasılığı o kadar artmaktadır (De la Torre, 2009a). Bu çalışmada, benzer araştırmalarla (Baker, 1993; De la Torre ve Douglas, 2004; Im ve Corter, 2011; Kunina-Habenicht ve diğerleri, 2012; MacDonald, 2013; Rupp ve Templin, 2008a) paralel olarak, Q matrisinin hatalı belirlenmesinin parametre kestirimlerini etkilediği bulunmuştur.

Hatasız Q matrisinde yapılan eksik belirlemeler, maddenin tahmin parametresi kestirimini etkilemezken, kaydırma parametresi değerini arttırmıştır. Yanıt örüntüsü beklenenden daha fazla yanlış cevap içerdiğinden, kaydırma parametresi gerçek değerinden yüksek değerde kestirilmektedir (Rupp ve Templin, 2008a). Bu durum, testi alan bireylerin çoğunun, ilgili madde ile ölçülüyor görünen bütün becerilere sahip olmalarına rağmen maddeye yanlış yanıt verdiği anlamına gelir. Çünkü maddeyle ölçülen en az bir beceri, bu madde tarafından ölçülmüyormuş gibi tanımlanmıştır. Ve bireyler bu beceriye sahip olmadıklarından maddeyi yanlış yanıtlamıştır. Ancak diğer maddelerdeki performanslarına göre, ilgili maddeyle ölçüldüğü tanımlanan becerilere sahip olduğu belirlenen bireylerin, bu maddeyi beklenen durumun aksine yanlış yanıtlaması, maddenin kaydırma parametresinin yükselmesine neden olmaktadır.

Hatasız Q matrisinde yapılan fazla belirlemeler, maddenin kaydırma parametresi kestirimini etkilemezken, tahmin parametresi değerini artırmıştır. Ölçme aracında yer alan bir madde, doğru yanıtlanması için bir beceri daha gerektirecek şekilde hatalı tanımlandığında, madde olduğundan daha zor olarak tanımlanmış olacaktır. Bireyler, madde ile ölçüldüğü hatalı olarak tanımlanan bu ek beceriye sahip olmasalar bile maddeyi doğru yanıtlayabileceklerdir. Yanıt örüntüsü beklenenden daha fazla doğru cevap içerdiğinden, tahmin parametresi gerçek değerinden yüksek değerde kestirilmektedir (Rupp ve Templin, 2008a).

Hatasız Q matrisinde yapılan dengeli hatalı belirlemeler, fazla belirleme yapılan maddelerde tahmin, eksik belirleme yapılan maddelerde kaydırma parametrelerinin artmasına, bir beceride eksik bir beceride fazla belirleme yapılan maddelerin hem tahmin hem de kaydırma parametresinin artmasına sebep olmaktadır. Eksik belirleme yapılan maddelerin kaydırma parametrelerindeki yükselmeler, fazla belirleme yapılan maddelerin tahmin parametrelerindeki yükselmelerden oldukça fazladır. Benzer biçimde aynı maddede hem eksik hem de fazla belirleme yapıldığında maddenin kaydırma parametresi, tahmin parametresinden daha fazla yükselmektedir.

Baker (1993) ve MacDonald (2013) tarafından yapılan çalışmalarda, Q matrisindeki hatalı belirleme oranlarının madde parametresi kestirimlerini etkilediği, hatalı belirleme oranı arttıkça kestirimlerin hata miktarının da arttığı raporlanmıştır. Bu çalışmada farklı koşullar kapsamında, aynı beceriye ilişkin aynı hatalı belirleme yapıldığında parametre kestirimleri hemen hemen aynı değeri almaktadır. Bir maddede birden fazla becerinin ölçülme durumuna ilişkin hatalı tanımlama yapıldığında, parametre kestirimlerindeki artış da fazlaşmaktadır. Örneğin, bir madde için birden fazla sayıda fazla tanımlama yapıldığında, maddeye doğru yanıt verecek kişi sayısı bir beceride fazla tanımlama yapıldığındaki durumdan daha fazla olmaktadır. Bu da tahmin parametresinin bir beceride fazla tanımlama yapılan durumdan daha fazla yükselmesine neden olmaktadır.

Eksik belirleme yapılan Q matrislerinde, hatalı belirleme yapılan maddeler olduğundan daha kolay hale geldiğinden bireylerin yanıtları olduğundan daha fazla yanlış içerirken, fazla belirleme yapılan Q matrislerinde, hatalı belirleme yapılan maddeler olduğundan daha zor hale geldiğinden bireylerin yanıtları olduğundan daha fazla doğru içerir (Rupp ve Templin, 2008a). Dengeli hatalı belirleme yapılan Q matrislerinde ise testin ortalama zorluğundaki farklılaşma rasgele seçilen ve eksik ya da fazla belirleme yapılan maddelerle ilişkilidir. Bu çalışma kapsamında incelenen koşullarda eksik belirleme yapılan koşullarda, fazla belirleme yapılan ve dengeli hatalı belirleme yapılan koşullara kıyasla sınıflama doğrulukları daha yüksek bulunmuştur. Bu bulgu Miao, Miller ve Ren (2017) tarafından yapılan çalışmanın DINA modele ilişkin bulgularıyla benzerlik göstermektedir.

Önceki araştırmalar incelendiğinde Q matrisinde yapılan hatalı belirleme oranı arttıkça kestirimlerin hata miktarının arttığı sonucuna ulaşılmıştır (Baker, 1993; Başokçu, 2014; MacDonald, 2013). Bu çalışma kapsamında yapılan incelemelerde de, benzer şekilde Q matrisinin eksik belirlendiği koşullarda hatalı belirleme oranı arttıkça sınıflama doğruluğunun azaldığı sonucuna ulaşılmıştır. Bu bulgu Miao, Miller ve Ren (2017) tarafından yapılan çalışmanın DINA modele ilişkin farklı test uzunlukları ve beceriler arası korelasyonların da incelendiği tüm koşullarla tutarlılık göstermektedir. Bu çalışma kapsamında yapılan incelemelerde, eksik belirleme koşullarına benzer şekilde, dengeli hatalı belirleme yapılan koşullarda da hatalı belirleme oranı arttıkça sınıflama doğruluğunun azalırken, fazla belirleme yapılan koşullarda ise hatalı belirleme oranı arttıkça, sınıflama doğruluğunun tümü hatasız Q matrisi ile yapılanlardan çok düşük olmak üzere arttığı sonucuna ulaşılmıştır.

Genel olarak, DINA modelde fazla belirleme yapılan ve dengeli hatalı belirleme yapılan Q matrisleri, eksik belirleme yapılan Q matrislerine göre sınıflama doğruluğu üzerinde daha ciddi olumsuz etkiye sahiptir. Ayrıca

yalnızca hatalı belirleme oranının artması değil, hatalı belirleme düzeninin farklılaşması da sınıflama doğruluğu üzerinde etkilidir. Bu sonuçlar Miao, Miller ve Ren'in (2017) çalışmasıyla paralellik göstermektedir.

## ÖNERİLER

Bu çalışmanın bulguları doğrultusunda ulaşılan sonuçlar, DINA modelde madde parametrelerinin ve bireylerin sınıflanma doğruluğunun, Q matrisinin hatalı belirlenmesinden etkilendiğini göstermektedir. Miao, Miller ve Ren (2017) yaptıkları çalışmada uygulayıcılara, bir maddenin bir özelliği ölçüp ölçmediğinin belirlenmesinde sıkıntı yaşanan durumlarda fazla tanımlamanın, eksik tanımlamadan daha iyi olabileceğini önermiştir. Uygulamada test geliştirme adımlarına uyulması ve Q matrisi doğrulama yöntemlerinin kullanılması oldukça önemlidir. Eksiksiz bir Q matrisi doğrulama sürecinin hem istatistiksel hem de temel bilgilerin ve alan uzmanlığının kullanılmasını gerektirdiği (de la Torre, 2008; Leighton, 2004; Tatsuoka, Corter ve Tatsuoka, 2004) unutulmamalıdır. Bu nedenle uygulayıcılara madde parametrelerinin ve bireylerin sınıflandırılmasına ilişkin yapılan kestirimlerin doğruluğunu artırmak amacıyla, Q matrisi oluşturulduktan sonra, madde-beceri ilişkilerinin tanımlanması ile ilgili hatalı belirleme yapılıp yapılmadığını kontrol etmek için uzman görüşü almaları, sonrasında de la Torre (2008) ve de la Torre ve Chiu (2016) tarafından önerilen yöntemler kullanarak Q matrisini doğrulanmaları önerilmektedir.

## Etik Metni

Bu makalede dergi yazım kurallarına, yayın ilkelerine, araştırma ve yayın etiğine ve dergi etik kurallarına uyulmuştur. Makaleyle ilgili meydana gelebilecek ihlallerin sorumluluğu yazara aittir. Bu çalışma simülasyon çalışması olduğundan veri toplama aracının uygulanmasına ilişkin etik kurul kararı gerektirmemektedir.

**Yazarın Katkı Oranı Beyanı:** Bu çalışmada yazarın katkı oranı %100'dür.

Katkılarından dolayı Prof. Dr. Ömay Çokluk-Bökeoğlu'na ve Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK) Bilim İnsanı Destek Programları Başkanlığı (BİDEB) tarafından yürütülen 2211 - Yurt İçi Doktora Burs Programı'na çok teşekkür ederim.

## KAYNAKÇA

- Baker, F. (1993) Sensitivity of the linear logistic test model to misspecification of the weight matrix. *Applied Psychological Measurement*, 17(3), 201-210. <https://doi.org/10.1177/014662169301700301>
- Basokcu, T. O. (2014). Classification accuracy effects of Q-matrix validation and sample size in DINA and G-DINA models. *Journal of Education and Practice*, 5(6), 220-230. <https://doi.org/10.7176/JEP>
- Cassuto, N. (1996). *The performance of the linear logistic test model under different testing conditions* [Unpublished doctoral dissertation]. University of Minnesota.
- Cui, Y., Gierl, M. J., & Chang, H. H. (2012). Estimating classification consistency and accuracy for cognitive diagnostic assessment. *Journal of Educational Measurement*, 49(1), 19-38. <https://doi.org/10.1111/j.1745-3984.2011.00158.x>

- Culpepper, S. A. (2015). Bayesian estimation of the DINA model with gibbs sampling. *Journal of Educational and Behavioral Statistics*, 40(5), 454–476. <https://doi.org/10.3102/1076998615595403>
- DeCarlo, L. T. (2012). Recognizing uncertainty in the Q-matrix via a Bayesian extension of the DINA model. *Applied Psychological Measurement*, 36(6), 447-468. <https://doi.org/10.1177/0146621612449069>
- 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. <https://doi.org/10.1111/j.1745-3984.2008.00069.x>
- De la Torre, J. (2009a). DINA model and parameter estimation: A didactic. *Journal of Educational and Behavioral Statistics*, 34(1), 115–130. <https://doi.org/10.3102/1076998607309474>
- De la Torre, J. (2009b). A cognitive diagnosis model for cognitively-based multiple-choice options. *Applied Psychological Measurement*, 33(3), 163–183. <https://doi.org/10.1177/0146621608320523>
- de la Torre, J., & Chiu, C. Y. (2016). A general method of empirical Q-matrix validation. *Psychometrika*, 81(2), 253-273. <https://doi.org/10.1007/s11336-015-9467-8>
- de la Torre, J. Hong, Y., & Deng W. (2010). Factors affecting the item parameter estimation and classification accuracy of the DINA model. *Journal of Educational Measurement*, 47(2), 227-249. <https://doi.org/10.1111/j.1745-3984.2010.00110.x>
- DiBello, L. V., & Stout, W. (2007). Guest editors' introduction and overview: IRT-based cognitive diagnostic models and related methods. *Journal of Educational Measurement*, 44(4), 285-291. <http://www.jstor.org/stable/20461864>
- Haertel, E. H. (1989). Using restricted latent class models to map the skill structure of achievement items. *Journal of Educational Measurement*, 26(4), 333–352. <https://doi.org/10.1111/j.1745-3984.1989.tb00336.x>
- Henson, R., & Douglas, J. (2005) Test construction for cognitive diagnosis. *Applied Psychological Measurement*, 29(4), 262-277. <https://doi.org/10.1177/0146621604272623>
- Henson, R. A., & Templin, J. L. (2006). *Implications of Q-matrix misspecification in cognitive diagnosis*. Manuscript submitted for publication.
- Im, S., & Corter, J. (2011). Statistical consequences of attribute misspecification in the rule space method. *Educational and Psychological Measurement*, 71(4), 712–731. <https://doi.org/10.1177/0013164410384855>
- Jang, E. E., Dunlop, M., Park, G., & van der Boom, E. H. (2015). How do young students with different profiles of reading skill mastery, perceived ability, and goal orientation respond to holistic diagnostic feedback? *Language Testing*, 32(3), 359–383. <https://doi.org/10.1177/0265532215570924>
- Jang, E. E., Kim, H., Vincett, M., Barron, C., & Russell, B. (2019). *Improving IELTS reading test score interpretations and utilisation through cognitive diagnosis model-based skill profiling* (IELTS Research Reports Online Series No. 2). British Council, Cambridge Assessment English and IDP: IELTS Australia. <https://www.ielts.org/research/research-reports/onlineseries-2019-2>
-

- Junker, B. W., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement*, 25(3), 258–272. <https://doi.org/10.1177/01466210122032064>
- Kunina-Habenicht, O., Rupp, A. A., & Wilhelm, O. (2009). A practical illustration of multidimensional diagnostic skills profiling: comparing results from confirmatory factor analysis and diagnostic classification models. *Studies in Educational Evaluation*, 35(2), 64–70. <https://doi.org/10.1016/j.stueduc.2009.10.003>
- Lee, Y. W., & Sawaki, Y. (2009). Cognitive diagnostic approaches to language assessment: An overview. *Language Assessment Quarterly*, 6(3), 172–189. <https://doi.org/10.1016/j.stueduc.2009.10.003>
- Leighton, J. (2004). Avoid misconception, misuse, and missed opportunities: The collection of verbal reports in educational achievement testing. *Educational Measurement: Issues and Practice*, 23(2), 6–15. <https://doi.org/10.1111/j.1745-3992.2004.tb00164.x>
- Leighton, J. P., & M. J. Gierl (2007). *Cognitive diagnostic assessment for education: Theory and practices*. Cambridge University Press.
- MacDonald, G. T. (2013). *The performance of the linear logistic test model when the Q-matrix is misspecified: a simulation study* [Unpublished doctoral dissertation]. University of South Florida.
- MacDonald, G., & Kromrey, J. (2011). *Linear logistic test model: Using SAS® to simulate the decomposition of item difficulty by algorithm, sample size, cognitive component and time to convergence*. Proceedings of the American Statistical Association's Joint Statistical Meeting, Social Statistics Section, Miami, FL.
- MacDonald, G., & Kromrey, J. (2012). *The effects of Q-matrix misspecification when employing Proc NL MIXED: A simulation study*. Paper presented at the annual SESUG conference, Durham, NC.
- Miao, G. A. O., Miller, M. D., & Ren, L. I. U. (2017). The impact of Q-matrix misspecification and model misuse on classification accuracy in the generalized DINA model. *Journal of Measurement and Evaluation in Education and Psychology*, 8(4), 391-403. <https://doi.org/10.21031/epod.332712>
- Orlando, M., & Marshall, G. N. (2002). Differential item functioning in a spanish translation of the PTSD checklist: detection and evaluation of impact. *Psychological Assessment*, 14(1), 50-9. <https://doi.org/10.1037/1040-3590.14.1.50>
- Ömür-Sünbül, S. Ö., & Kan, A. (2015). Bilişsel tanı modellerinde parametre kestirimini ve sınıflama tutarlılığını etkileyen faktörlerin incelenmesi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 31(4), 778-795. <https://doi.org/10.16986/HUJE.2015014663>
- Qin, C., Zhang, L., Qiu, D., Huang, L., Geng, T., Jiang, H., Ren, Q. & Zhou, J. (2015). Model identification and Q-matrix incremental inference in cognitive diagnosis. *Knowledge-Based Systems*, 86, 66-76. <https://doi.org/10.1016/j.knosys.2015.05.024>
- Romero, S. J., Ordoñez, X. G., Ponsoda, V., & Revuelta, J. (2014). Detection of Q-matrix misspecification using two criteria for validation of cognitive structures under the least squares distance model. *Psicologica: International Journal of Methodology and Experimental Psychology*, 35(1), 149-169.
-

- Rupp, A. A. (2007). The answer is in the question: A guide for describing and investigating the conceptual foundations and statistical properties of cognitive psychometric models. *International Journal of Testing*, 7(2), 95–125. <https://doi.org/10.1080/15305050701193454>
- Rupp, A., & Templin, J. (2008a). The effects of Q-matrix misspecification on parameter estimates and classification accuracy in the DINA model. *Education and Psychological Measurement*, 68(1), 78–96. <https://doi.org/10.1177/0013164407301545>
- Rupp, A. A., & Templin, J. L. (2008b). Unique characteristics of diagnostic classification models: A comprehensive review of the current state of the art. *Measurement: Interdisciplinary Research and Perspectives*, 6(4), 219–262. <https://doi.org/10.1080/15366360802490866>
- Rupp, A. A., Templin, J., & Henson, R. A. (2010). *Diagnostic measurement: Theory, methods, and applications: Methodology in the social sciences*. Guilford Press.
- Templin, J., & Hoffman, L. (2013). Obtaining diagnostic classification model estimates using Mplus. *Educational measurement: Issues and practice*, 32(2), 37-50. <https://doi.org/10.1111/emip.12010>
- Tatsuoka, K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement* 20, 345–354. <https://www.jstor.org/stable/1434951>
- Tatsuoka, K. (1990). Toward an integration of item-response theory and cognitive error diagnosis. In N. Frederiksen, R. Glaser, A. Lesgold, & Safto, M. (Eds.). *Monitoring skills and knowledge acquisition* (pp. 453-488). Erlbaum.
- Tatsuoka, K. K., Corter, J., & Tatsuoka, C. (2004). Patterns of diagnosed mathematical content and process skills in TIMMS-R across a sample of 20 countries. *American Educational Research Journal*, 41, 901–926. <https://doi.org/10.3102/00028312041004901>

**EKLER****Ek 1. Hatasız Q Matrisi**

Madde	Beceri				Madde	Beceri			
	1	2	3	4		1	2	3	4
1	1	0	0	0	16	1	0	0	0
2	0	1	0	0	17	0	1	0	0
3	0	0	1	0	18	0	0	1	0
4	0	0	0	1	19	0	0	0	1
5	1	1	0	0	20	1	1	0	0
6	1	0	1	0	21	1	0	1	0
7	1	0	0	1	22	1	0	0	1
8	0	1	1	0	23	0	1	1	0
9	0	1	0	1	24	0	1	0	1
10	0	0	1	1	25	0	0	1	1
11	1	1	1	0	26	1	1	1	0
12	1	1	0	1	27	1	1	0	1
13	1	0	1	1	28	1	0	1	1
14	0	1	1	1	29	0	1	1	1
15	1	1	1	1	30	1	1	1	1