ISSN: 1309 – 6575
Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi
Journal of Measurement and Evaluation in Education and Psychology
2018; 9(3);239-257

The Interaction Effect of the Correlation between Dimensions and Item Discrimination on Parameter Estimation*

Sakine GÖÇER ŞAHİN** Derya ÇAKICI ESER*** Selahattin GELBAL****

Abstract
There are some studies in the literature that have considered the impact of modeling multidimensional mixed structured tests as unidimensional. These studies have demonstrated that the error associated with the discrimination parameters increases as the correlation between dimensions increases. In this study, the interaction between items' angles on coordinate system and the correlations between dimensions was investigated when estimating multidimensional tests as unidimensional. Data were simulated based on two dimensional, and two-parameter compensatory MIRT model. Angles of items were determined as 0.15º; 0.30º; 0.45º; 0.60º and 0.75º respectively. The correlations between ability parameters were set to 0.15, 0.30, 0.45, 0.60 and 0.75 respectively, which are same with the angles of discrimination parameters. The ability distributions were generated from standard normal, positively and negatively skewed distributions. A total of 75 (5 x 5 x 3) conditions were studied: five different conditions for the correlation between dimensions; five different angles of items and three different ability distributions. For all conditions, the number of items was fixed at 25 and the sample size was fixed at n = 2,000. Item and ability parameter estimation were conducted using BILOG. For each condition, 100 replications were performed. The RMSE statistic was used to evaluate parameter estimation errors, when multidimensional response data were scaled using a unidimensional IRT model. Based on the findings, it can be concluded that the pattern of RMSE values especially for discrimination parameters are different from the existing studies in the literature in which multidimensional tests were estimated as unidimensional.

Key Words: Multidimensional data, unidimensional estimation, correlation, discrimination index.

INTRODUCTION

Unidimensionality, which is one of the most fundamental assumptions of modern measurement theories, refers to measuring a single trait through test. Unidimensionality is necessary for ranking individuals on a scale. On the other hand, unidimensionality assumption is not always met in practice since the measured traits may not be perfectly pure. Thus, the unidimensionality assumption and the item response theory (IRT) models relying on this assumption are criticized in various aspects.

The critics on unidimensionality assumption and structure of tests measuring multiple traits have encouraged researchers to develop and employ multidimensional measurement models. Therefore IRT, which has been used for unidimensional tests from its release until the late 1970s, has been extended to multidimensional tests and has started to be used with the test measuring multiple abilities under the name of multidimensional item response theory (MIRT) since the late 1970s and early 1980s (Ansley & Forsyth, 1985; Reckase, 2009).

Multidimensionality means that the test intends to measure multiple traits. Multidimensionality can be applied with different test structures. In this respect, multidimensional tests may have simple, approximate simple, complex, mixed and semi-mixed structures. A simple structured test consists of multiple subtests each of which measures a single trait, and each item in these subtests is related to a

*An early draft of this paper was presented at International Meeting of Psychometric Society (IMPS) in Beijing, China in 2015.
** Postdoctoral Researcher, University of Wisconsin-Madison, Madison, WI, USA, e-mail: sgocersahin@gmail.com, ORCID ID: orcid.org/0000-0002-6914-354X
***Assit. Prof. Dr., Kırıkkale University, Education Faculty, Kırıkkale, TURKEY, e-mail: deryacakicieser@gmail.com, ORCID ID: orcid.org/0000-0002-4152-6821
****Prof. Dr. Hacettepe University, Education Faculty, Ankara, TURKEY, e-mail: sgelbal@gmail.com, ORCID ID: orcid.org/0000-0001-5181-7262

Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, Cilt 9, Sayı 3, Sonbahar 2018, 239-257.
Journal of Measurement and Evaluation in Education and Psychology, Vol. 9, Issue 3, Autumn, 2018, 239-257.
Received: 07.03.2018
DOI:10.21031/epod.402992
Accepted: 05.07.2018

single trait. Tests with an approximately simple structure are also composed of subtests. Each subtest is approximately unidimensional, which means that there is a dimension that is measured recessively in addition to a dominant dimension (Zhang, 2005; Zhang, 2012). As for the tests with a complex structure, both the entire test and the items in the test are related to more than one ability. From a factor analytic perspective, in complex structured tests, items have factor loadings on multiple abilities (Bulut, 2013; Sheng & Wikle, 2007). Mixed structured tests include both simple and complex items. And the semi-mixed tests include both approximate simple and complex items (Zhang, 2012).

Test dimensionality should be carefully examined before implementation of the tests and analysis and interpretation of results. The implementation and interpretation stages of multidimensional analyses are more complicated than that of unidimensional structures. Stages of multidimensional analyses are more complicated than that of unidimensional structures. Due to convenience of implementing and interpreting the unidimensional IRT models, some researchers lean towards analyses in which multidimensional models are estimated as unidimensional. There are studies in the literature estimating multidimensional tests as unidimensional since 1980s (i.e., Ackerman, 1989; Ansley & Forsyth, 1985; Drasgow & Parsons, 1983; Harrison, 1986; Kirisci, Hsu, & Yu, 2001, Leucht & Miller; 1992; Reckase, Ackerman, & Carlson, 1988; Zhang, 2008; Zhang, 2012). Estimating multidimensional constructs as unidimensional is generally referred as model misspecification.

There are many studies in the literature about model misspecification. In a study carried out by Drasgow and Parsons (1983), impact of applying unidimensional IRT to multidimensional data on item and person parameters was analyzed using LOGIST program. In the study, conditions, in which medium level heterogenous items were used, fitted better to unidimensional model. In another study carried out by Ansley and Forsyth (1985), parameters acquired from unidimensional estimation of two-dimensional constructs were analyzed. According to the obtained findings, correlations between estimation values and true values of difficulty parameter were higher than the correlation between other parameters. Harrison (1986) analyzed robustness of IRT parameters based on hierarchical factor model under various conditions using LOGIST program. According to these results, it was observed that as the test length increased, estimated and observed values of discrimination index got closer to each other; indicating that LOGIST program created better values for unidimensional constructs; and *D* parameter acquired through this program was more robust to the violation of unidimensionality. With respect to the ability parameter, it has been observed that as the test length increased, and the strength of general factor increased, correlation between ability parameters acquired from unidimensional and multidimensional structures increased and RMSD values decreased. In a study carried out by Reckase, Ackerman, and Carlson (1988), a unidimensional test was attempted to be formed using multidimensional items. Two data sets were used in the study. In the first data set, 80 items were calibrated based on two-parameter logistic model (2 PL). First 20 items of these 80 items were formed to measure only θ_1; second 20 items were formed to measure θ_1 and θ_2 in an equal level; third 20 items were formed to measure only θ_2; and finally, a two-dimensional data set was created as angles of the fourth 20 items could distribute equally between 0 – 90º. According to the simulation results, it was observed that 20 items in the first three groups did not show too much deviation from unidimensionality, and the last 20 items showed better consistence with the multidimensional model. Additionally, it was observed that the whole test showed better fit with the multidimensional model. On the contrary, findings acquired from the real data set showed more different results from the simulation data, and a data set designed as two dimensional with 68 items showed better fit with unidimensional model. In the study carried out by Ackerman (1989), multidimensional data generated based on compensatory and non-compensatory models were calibrated using BILOG and LOGIST programs. According to the results observed using both programs, as the correlation between dimensions in the data generated based on non-compensatory model increased, the correlation of a_1 and a_2 parameters with the estimated a parameter approached to 0. It has been observed that although average absolute errors were a little higher for discrimination and difficulty parameters obtained from BILOG program, errors decreased as the correlation between dimensions increased. It was indicated that *D* parameter was more robust in both programs. Results acquired from non-compensatory model showed similarity with the compensatory model. In addition to this, average absolute errors obtained from BILOG program were lower than the errors obtained from LOGIST program. In a study carried out by Kirisci, Hsu, and Yu (2001), in cases that unidimensionality and normality assumptions were not met, estimations acquired from BILOG, MULTILOG, and XCALIBRE programs were compared. Test and individual parameters were estimated based on data including three dimensional structures where unidimensional and interdimensional correlation was 0.6 and ability distributions were normal, positively-skewed and platykurtic. RMSE values were used to evaluate the results. RMSE values on the basis of distributions, dimensions, and programs were compared via ANOVA. According to ANOVA results, main effect of distributions and its interaction with other variables were not significant. It was observed that main effect of the dimension was significant only for c_i parameter. In the study where Zhang (2008) analyzed unidimensional parameter estimations and deviations from unidimensionality, used the number of dimensions as four; the test length as 15, 30, and 60; the rate of number of items that load to other dimensions as 20%, 40%, and 60%; and the correlation between factors as 0.00, 0.40, and 0.80. According to the findings, it was observed that as the correlation between secondary dimensions and the dominant dimension increased, the structure did not deviate much from unidimensionality. It was indicated that as the correlation decreased and the rate of items loading to other dimensions increased, the structure diverged from approximate unidimensionality. Another factor affecting divergence from approximate unidimensionality was the test length. When interdimensional correlation was low, shorter tests produced better results compared to longer tests. One of the conditions examined in the studies mentioned above is the structure of the test (approximate simple or complex) while the other most-focused conditions are the skewness of distribution and correlation among the dimensions. In these studies, the general finding about effect of correlation is that when the correlation between dimensions increased, the estimation error was decreased. However, in a study conducted by Gocer Sahin, Walker, and Gelbal (2015), it was reported that contrary to the findings in the literature, especially errors of item parameters increased as the correlation among the dimensions increased and that the lowest level of errors occurred when the correlation was 0.45. In another study carried out by Gocer Sahin (2016), a multidimensional test with a semi-mixed structure was estimated as unidimensional, and the same unexpected pattern related to correlation and test parameters was obtained. A similar study carried out by Kahraman (2013) reported that errors of discrimination increased as the correlation increased when the second dimension of the multidimensional test was ignored and then estimated as unidimensional.

Although there are studies in the literature showed that as the correlation between dimensions increased the estimation errors decreased, in the recent studies an opposite pattern was observed. This may be because of the test structure. In the previous studies, the tests had approximately simple structured items which most of items loaded one factor dominantly and recessively loaded on the second dimension. However, in the recent studies, test structure had mixed format which some items loaded dominantly on one factor some loaded on both dimension. Thus, one factor that makes this study different than others is the test structure. Although the results in the studies conducted by Kahraman (2013), Gocer Sahin, Walker, and Gelbal (2015), Gocer Sahin (2016) appear to be promising, they have not explained the possible reasons behind that results. So, in this study, the focus was on the interaction between correlation and items.

Purpose of the Study

In the recent studies related to the estimations of semi-mixed structured multidimensional tests as unidimensional, we think that increase in errors associated with item parameters because of the increase in correlation between the dimensions may stem from the interaction between the items' angles and the correlation. This study was carried out in order to test whether this hypothesis was true. Therefore, this study aims to answer following questions:

1. How much error is included in parameter estimation when a two-dimensional test is treated as unidimensional?
2. Is there a pattern for error associated with ability parameters in the case of misspecification of two-dimensional tests as unidimensional?

3. How the ability estimations are affected by the interaction among different ability distributions, correlation between dimensions and angles of items on the x-axis?

METHOD

In this study, simulated data sets were used to perform research purpose. Simulation models should be based on realistic situations (Davey, Nering, & Thompson, 1997). In this study the minimum number of items in the large-scale tests was considered test length. In large scale tests for example, in high school entrance exams, each sub test includes 20 questions. So, two dimensional tests with 25 items and with a semi-mixed structure were simulated. According to Hambleton (1989), a large (around 1,000) sample is required to obtain accurate item-parameter estimates in IRT (Hambleton, 1989) for accurate estimates of ability parameter, upon which some high-stakes decisions are made. To eliminate the sample size effect, an enough number of examinees were simulated. In the whole design, the sample size was fixed to be 2,000. The independent variables of the study are correlation among dimensions, items' angle with x-axis, and distribution of ability parameters.

In this respect, the correlation among the ability parameters in the two-dimensional tests is manipulated in an order from the lowest relation to the highest relation (ρ=0.15; ρ=0.30; ρ=0.45; ρ=0.60; ρ=0.75). There are some findings in the literature showing that the shape of distributions affects the parameter estimation in BILOG (Abdel-Fattah, 1994; Kim & Lee, 2014; Kirisci, Hsu, & Yu, 2001; Seong, 1990; Toland, 2008; Yen, 1987). Although it is known that the ability distribution has impacts on the parameter estimation, its impact on semi-mixed structured tests is not known yet. So, in this study ability distribution was one of the independent variables. Since the standard normal distribution is used by default as the initial (prior) ability distribution for calibrating item parameters in BILOG, standard normal distributions were added to the design as a baseline condition. For standard normal distributions, underlying ability distributions for both dimensions were simulated as standard normal $N(0, 1)$. For positive and negative skewed distributions, the values in the Fleisman's (1978) study were used. For positively skewed distributions and negatively skewed distributions skewness and kurtosis were (1.75, 3.75) and (-1.75, 3.75), respectively. For each condition, 100 replications were performed.

In MIRT, items can be represented by item vectors on Cartesian coordinate system. Each item vector is on a line that crosses the origin. The direction of the vector is defined as the vector's angle with positive θ_1 axis. The direction of an i item is calculated through the following equation (Reckase, 2009):

$$\alpha_i = arccos \frac{\alpha_{i1}}{\sqrt{\alpha_{i1}^2+\alpha_{i2}^2}} \tag{1}$$

In Equation 1, a_i refers to the discrimination of item i. Items that are closer to θ_1 axis primarily measure the θ_1 ability while items that are closer to θ_2 axis primarily measure the θ_2 ability. Items have an angle of 45^o with both ability axes equally measure both of the abilities (Ackerman, 1994; Ackerman, Gierl, & Walker, 2003). Accordingly, in this study, the angles of item vectors with x axis are manipulated as 15^o, 30^o, 45^o, 60^o, and 75^o, which are the same numerical values as the correlations. In such a design, the items with angles of 15^o and 30^o measure the θ_1 ability, the items with angles of 45^o measure both θ_1 and θ_2, and the items with angles of 60^o and 75^o primarily measure the θ_2 ability. Ability parameters were acquired from three different distributions, which were standard normal, positive skewed and negative skewed distribution. In this arrangement, the ability distributions had three conditions, items' angles with x axis had five conditions, and correlations among dimensions had five conditions; which resulted in a total of 75 conditions (3 x 5 x 5). Data were generated through the SAS software on the basis of compensatory two parameter logistic model with the following equation (2) (Reckase, 2009):

$$\mathrm{P(U_{ij}} = 1 \mid \theta_i, a_i, d_i = \frac{e^{a_i\theta'_j+d_i}}{1+e^{a_i\theta'_j+d_i}} \tag{2}$$

where P is the conditional probability that examinee j's response, U_{ij}, to item i is correct, θ_j is the ability vector, a_i is the discrimination parameter vector, and d_i represents scalar difficulty of item i.

Item and ability parameter estimation were conducted using BILOG.

In order to have a baseline condition for comparison purposes, a unidimensional data set was also simulated. To generate unidimensional data, multidimensional test parameters were utilized. MDISC (maximum discrimination index) and *D* were used as the discrimination and difficulty parameters for unidimensional tests, respectively. *MDISC* is the overall discriminating power of an item which shares the same interpretation as the discrimination parameter in the unidimensional models (Reckase & McKinley, 1991).

$$\mathrm{MDISC}_i = \sqrt{\sum_{\mathrm{k=1}}^{\mathrm{m}} a_{ik}^2} \quad (3)$$

where *m* refers to the number of ability dimensions the a_{ik} variable refers to the discrimination value that belongs to each dimension. The difficulty level of an item is defined as (Reckase, 2009):

$$D_i = \frac{-d_i}{MDISC} \quad (4)$$

In Equation 4, d_i is intercept term. The value of D_i has the same interpretation as the *b* parameter in the unidimensional IRT. The number of items was fixed at 25 and the sample size was fixed at *n* = 2,000 for the simulated unidimensional test data as well. The RMSE values obtained from the unidimensional tests were used as the baseline criterion to evaluate the magnitude of the errors that were obtained from the multidimensional data.

$$RMSE = \sqrt{\frac{\sum_r^n (\hat{X}_{ir} - X_i)^2}{n}} \quad (5)$$

In Equation 5, *i* and *r* represent items (or examinees) and replications, respectively, *n* is the total number of replications, and $\hat{X}_{ir}$ is the estimate of parameter X_i (a_1, a_2 *and* a_{avg} *(the average of* a_1 *and* a_2*)*, *D*, θ_1, θ_2, *and* θ_{avg} *(the average of* θ_1*, and* θ_2*)* or *MDISC*). RMSE (Root Mean Square Error) statistics in the equation (5) were used to evaluate the errors associated with the estimated parameters. This equation is used to calculate the error in ability parameters, and this formula was also adapted to item parameters.

In the findings part, ANOVA was conducted to determine the impact of different correlations, distributions, and angles given in Table 1-7. Although the homogeneity of variances for some data was not met, ANOVA was continued in order to provide consistency in all results. With the aim of comparing the results, Bonferroni’s method was used for post hoc comparisons.

RESULTS

a_1 Parameter:

The RMSE values obtained for the a_1 parameter are displayed in Table 1. When the distribution of errors pertaining to the a_1 parameter along the change of the correlations are examined by keeping the item’s angle constant, it was observed that the errors decreased as the correlation among the dimensions increased under the conditions with the angles smaller than 45^o. Under the conditions where angles were higher than 45^o, the errors increased as the correlation among the dimensions increased. The only condition that did not conform to the pattern related to correlation and angle was when the distributions were standard normal, and the angle was 45^o.

When the distributions were standard normal, and the item’s angle was 45^o, then the errors had a hyperbolic curve. In this respect, when the correlation was kept constant, the errors decreased until the angle reached to 45^o whereas the errors increased after 45^o. An evaluation according to the distributions showed that the skewness of the distributions affected the a_1 parameter. Especially when the items’ angles were higher than 45^o (when the angles are 60^o and 75^o), the RMSE values obtained under the conditions of standard normal distributions were higher than the error values obtained under the conditions of skewed distributions. Under other conditions apart from this, the RMSE values obtained in skewed distributions were bigger than the error values obtained in standard normal distributions. It should also be added that the direction of the skewness had no effect on the a_1 parameter. The important point here is whether the distribution is skewed or standard normal; it is not the direction of the skewness. A comparison of the RMSE values obtained through the estimation of multidimensional data as unidimensional revealed that the errors closest to the criterion values were observed under the conditions where angles were 45^o.

***a_2* Parameter:**

The RMSE values obtained for the a_2 parameter are presented in Table 2. Evaluation of a_2 parameter showed an opposite pattern with a_1 parameter. When the angle was kept constant, errors pertaining to the a_2 parameter increased as the correlation increased in the conditions with the angles smaller than 45^o. In the conditions with the angles higher than 45^o, the errors decreased as the correlation increased. An evaluation based on the distributions showed that the same symmetric pattern between a_1 and a_2 also occurred. Specifically, when the items' angles were smaller than 45^o (when the angles are 15^o and 30^o), the RMSE values obtained under the conditions of standard normal distribution were higher than the error values obtained under the conditions of other skewed distribution. In the cases that angles were 45^o or above, the RMSE values obtained under the conditions of standard normal distribution were lower than the RMSE values obtained under the conditions of skewed distribution. When all these values are compared with the criterion RMSE values, it is observed that in the condition where angle is 45^o, the errors related to a_2 parameter were generally lower than the criteria values.

The comparison of the error sizes pertaining to the a_1 and a_2 parameters revealed that in some cases, the errors of a_1 were higher and in other cases, the errors of a_2 were higher. The patterns obtained were generally symmetrical. It is observed that the average error within each condition for both parameters were close to each other.

Table 1. RMSE Values for a_1 Parameter

		Correlation of Between Abilities														
Angles	**Results of Unidimensional data**	**ρ_1=0.15**			**ρ_1=0.30**			**ρ_1=0.45**			**ρ_1=0.60**			**ρ_1=0.75**		
		SND*	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****
15^0	0.058	0.421	0.473	0.472	0.407	0.466	0.465	0.397	0.458	0.456	0.387	0.453	0.444	0.379	0.443	0.435
30^0	0.080	0.307	0.359	0.378	0.273	0.336	0.357	0.238	0.317	0.338	0.211	0.292	0.317	0.184	0.274	0.294
45^0	0.121	0.151	0.245	0.242	0.109	0.219	0.224	0.083	0.206	0.204	0.088	0.193	0.190	0.113	0.183	0.182
60^0	0.101	0.179	0.136	0.151	0.224	0.160	0.173	0.267	0.185	0.196	0.310	0.213	0.225	0.354	0.245	0.251
75^0	0.125	0.533	0.455	0.444	0.564	0.473	0.460	0.595	0.493	0.478	0.626	0.517	0.501	0.655	0.542	0.521

*SND: Standard Normal Distribution, **PSD: Positive Skewed Distribution, ***NSD: Negative Skewed Distribution

Table 2. RMSE Values for a_2 Parameter

		Correlation of Between Abilities														
Angles	**Results of Unidimensional data**	**ρ_1=0.15**			**ρ_1=0.30**			**ρ_1=0.45**			**ρ_1=0.60**			**ρ_1=0.75**		
		SND*	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****
15^0	0.058	0.471	0.434	0.427	0.479	0.436	0.431	0.486	0.441	0.437	0.496	0.442	0.448	0.504	0.449	0.455
30^0	0.080	0.183	0.163	0.173	0.206	0.177	0.181	0.234	0.191	0.192	0.260	0.215	0.207	0.293	0.235	0.227
45^0	0.121	0.146	0.238	0.239	0.107	0.214	0.221	0.080	0.200	0.202	0.086	0.186	0.188	0.115	0.177	0.182
60^0	0.101	0.368	0.457	0.451	0.320	0.434	0.427	0.277	0.410	0.402	0.235	0.384	0.374	0.194	0.357	0.350
75^0	0.125	0.576	0.679	0.692	0.545	0.663	0.678	0.514	0.645	0.660	0.484	0.624	0.639	0.455	0.601	0.620

*SND: Standard Normal Distribution, **PSD: Positive Skewed Distribution, ***NSD: Negative Skewed Distribution

a_{avg} Parameter

The RMSE values obtained for the a_{avg} parameter can be seen in Table 3. Under the conditions with standard normal distribution, the highest errors were obtained when the correlation among the dimensions was 0.15, and the lowest errors were obtained when the correlation was 0.45 for the average of *a* parameters. No regular pattern was found under the conditions with standard normal distribution. When the errors are examined for the correlations by keeping the angles fixed, it can be suggested that the errors of a_{avg} yielded a hyperbolic curve for to the correlation between the dimensions. The RMSE values obtained under the conditions with standard normal distribution were generally lower than the values obtained under the conditions with skewed distribution. Under the conditions with skewed distribution, the errors decreased as the correlation among the dimensions increased. When the distributions were skewed, the highest errors were found at 45^{o}, and the lowest errors were found at 15^{o}. The errors closest to the criterion values under the conditions with skewed distribution were obtained when the correlation was 0.75. The sizes of the errors pertaining to the a_{avg} parameter were between the a_1 and a_2 parameters. A comparison of all the obtained values with the criterion RMSE values showed that the errors, which were obtained when the correlation among the dimensions was 0.45 and the distribution was standard normal, were generally lower than the criterion values.

MDISC Parameter:

The RMSE values obtained for the *MDISC* parameter are presented in Table 4. It is observed that the *MDISC* parameter which corresponds to the discrimination parameter in the unidimensional IRT included more errors than all other discrimination parameters. The error values decreased as the correlation increased. In general, the errors increased as the angles increased. Under each condition of distribution, the lowest errors were obtained when the correlation was 0.75. The RMSE values obtained under the conditions of standard normal distribution were lower than the error values obtained under the conditions of skewed distribution. Whether the distribution is right or left skewed is not very influential on the RMSE. Accordingly, the effective condition for the RMSE is whether the distribution is standard normal or not. In general, it can be suggested that, the errors pertaining to the *MDISC* were quite higher than the criterion values.

Table 3. RMSE Values for a_{avg} Parameter

		Correlation of Between Abilities														
		ρ_1=0.15			ρ_1=0.30			ρ_1=0.45			ρ_1=0.60			ρ_1=0.75		
Angles	**Results of Unidimensional data**	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****
15^0	0.058	0.081	0.116	0.098	0.070	0.105	0.089	0.066	0.096	0.080	0.068	0.086	0.077	0.076	0.080	0.074
30^0	0.080	0.103	0.156	0.183	0.072	0.138	0.164	0.051	0.125	0.149	0.055	0.112	0.137	0.082	0.109	0.126
45^0	0.121	0.139	0.236	0.234	0.094	0.210	0.216	0.062	0.196	0.196	0.069	0.182	0.182	0.101	0.172	0.174
60^0	0.101	0.113	0.207	0.205	0.073	0.190	0.188	0.056	0.174	0.170	0.068	0.160	0.156	0.102	0.151	0.148
75^0	0.125	0.061	0.173	0.184	0.058	0.164	0.176	0.071	0.158	0.167	0.092	0.154	0.159	0.116	0.152	0.155

*SND: Standard Normal Distribution, **PSD: Positive Skewed Distribution, ***NSD: Negative Skewed Distribution

Table 4. RMSE Values for *MDISC* Parameter

		Correlation of Between Abilities														
		ρ_1=0.15			ρ_1=0.30			ρ_1=0.45			ρ_1=0.60			ρ_1=0.75		
Angles	**Results of Unidimensional data**	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****
15^0	0.058	0.463	0.515	0.513	0.448	0.508	0.506	0.438	0.500	0.497	0.427	0.494	0.484	0.419	0.484	0.475
30^0	0.080	0.466	0.514	0.539	0.427	0.489	0.516	0.387	0.467	0.495	0.354	0.437	0.471	0.315	0.413	0.444
45^0	0.121	0.551	0.636	0.636	0.449	0.601	0.611	0.443	0.576	0.580	0.391	0.547	0.551	0.345	0.510	0.523
60^0	0.101	0.551	0.636	0.633	0.501	0.611	0.609	0.457	0.585	0.582	0.414	0.557	0.552	0.370	0.526	0.527
75^0	0.125	0.627	0.729	0.743	0.596	0.713	0.728	0.565	0.695	0.711	0.535	0.673	0.689	0.505	0.650	0.670

*SND: Standard Normal Distribution, **PSD: Positive Skewed Distribution, ***NSD: Negative Skewed Distribution

***D* parameter:**

The RMSE values obtained for the *D* parameter are displayed presented in Table 5. As for the errors pertaining to the difficulty parameter obtained when the two-dimensional tests were estimated as unidimensional, it was observed that the errors increased as the correlation among the dimensions increased. In the case of standard normal distributions, the lowest error occurred when the correlation among the dimensions was 0.15 while the highest error occurred when the correlation was 0.75. However, no regular pattern was found regarding the errors under the condition with skewed distributions. Accordingly, in the case that distributions were skewed, and the angle was 15º and 75º, the errors decreased as the correlation increased. When the item's angle with the *x* axis was 30º, 45º and 60º, and the distribution was positively-skewed, RMSE values again produced a hyperbolic curve. Accordingly, errors decreased until the correlation of 0.45 and they increased again after the correlation of 0.45. The pattern that was obtained in the positively-skewed distribution was generally observed in the negatively-skewed distribution. When the correlations and distributions were fixed, and the angles increased, the errors did not exhibit a regular pattern. Under the condition with correlation of 0.15 between the dimensions and when the distribution was standard normal, considering the errors pertaining to the *b* parameter showed that the criterion values were closest to each other. Under this condition, almost all of the errors that were obtained by estimating the two-dimensional structures as unidimensional were lower than the criterion value.

θ_1 parameter:

The RMSE values obtained for the θ_1 parameter are presented in Table 6. Errors pertaining to the θ_1 parameter were affected by both correlation between ability parameters and angle of items. In this respect, the errors decreased as the correlation between the dimensions increased. In the case that distributions and correlations were held constant, the errors increased only when the angles increased. Specifically, the increase of the angle under the conditions of low correlation resulted in a significant increase in the errors; the increase of the angle under the conditions of high correlation had relatively lower effect on the errors. The highest errors were obtained when the correlation was 0.15 and the angle was 75º. Varying the distribution did not have a significant effect on the errors. Under all conditions, the errors obtained in standard normal distribution had lower values than in the positively and negatively skewed distributions. The errors acquired from the skewed distributions under the same conditions had similar values. The errors obtained for the θ_1 parameter were quite higher than the criterion values under all conditions. When the correlation was 0.75, the criterion RMSE and the obtained RMSE values were closest to each other, but the difference increased as the angle increased.

Table 5. RMSE Values for *D* Parameter

		Correlation of Abilities														
		ρ_1=0.15			ρ_1=0.30			ρ_1=0.45			ρ_1=0.60			ρ_1=0.75		
Angles	**Results of Unidimensional data**	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****
15^0	0.053	0.095	0.180	0.182	0.100	0.171	0.187	0.120	0.164	0.193	0.139	0.158	0.200	0.160	0.157	0.214
30^0	0.081	0.078	0.179	0.213	0.108	0.169	0.191	0.151	0.164	0.181	0.191	0.171	0.176	0.230	0.182	0.179
45^0	0.123	0.078	0.208	0.209	0.124	0.190	0.196	0.175	0.189	0.193	0.222	0.193	0.196	0.261	0.204	0.200
60^0	0.090	0.076	0.200	0.197	0.109	0.187	0.178	0.151	0.178	0.164	0.189	0.179	0.165	0.225	0.187	0.170
75^0	0.095	0.057	0.222	0.247	0.068	0.201	0.229	0.087	0.193	0.217	0.108	0.180	0.199	0.131	0.170	0.183

*SND: Standard Normal Distribution, **PSD: Positive Skewed Distribution, ***NSD: Negative Skewed Distribution

Table 6. RMSE Values for θ_1 Parameter

		Correlation of Abilities														
		ρ_1=0.15			ρ_1=0.30			ρ_1=0.45			ρ_1=0.60			ρ_1=0.75		
Angles	**Results of Unidimensional data**	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****
15^0	0.053	0.447	0.493	0.494	0.439	0.489	0.489	0.431	0.482	0.484	0.421	0.472	0.474	0.410	0.467	0.464
30^0	0.081	0.597	0.641	0.633	0.560	0.612	0.607	0.519	0.579	0.576	0.477	0.541	0.542	0.432	0.497	0.500
45^0	0.123	0.748	0.776	0.777	0.685	0.731	0.733	0.618	0.679	0.682	0.548	0.618	0.619	0.472	0.549	0.548
60^0	0.090	0.930	0.945	0.951	0.842	0.881	0.888	0.753	0.753	0.813	0.656	0.725	0.731	0.551	0.626	0.627
75^0	0.095	1.108	1.122	1.121	1.006	1.044	1.045	0.895	0.954	0.958	0.776	0.845	0.850	0.638	0.717	0.719

*SND: Standard Normal Distribution, **PSD: Positive Skewed Distribution, ***NSD: Negative Skewed Distribution

***θ_2* parameter:**

The RMSE values obtained for the θ_2 parameter are presented in Table 7. As seen in Table 7, errors pertaining to the θ_2 parameter significantly decreased as the correlation between dimensions increased. It can be suggested that the varying the distribution did not affect the errors significantly. When the distributions are compared to each other with other conditions being fixed, the lowest error values were obtained under the condition of standard normal distribution. Errors obtained in positively and negatively-skewed distributions under the same conditions were close to each other in general. As the angles increased, the errors obtained for θ_2 decreased. When all the results are considered together, it was observed that the lowest error occurred when the correlation was 0.75 and the angle was 75^o, and the highest error occurred when the correlation was 0.15 and the angle was 15^o. The difference between the criterion values and the estimated values for the θ_2 parameter increased as the angles and correlations increased; under all conditions, the criterion RMSE values were lower than the RMSE values obtained for the multidimensional data.

When the two-dimensional structures are estimated as unidimensional, the errors pertaining to the θ_2 parameter had similarities to the error values obtained for θ_1 under the same conditions. According to this, the errors were affected by the increase of the correlation and by the distributions in the same way. However, contrary to the situation observed in the θ_1 parameter, the errors of θ_2 decreased as the angle increased. The error patterns obtained for θ_1 and the error patterns obtained for θ_2 were opposite. In this respect, it can be suggested that the errors obtained for θ_1 and θ_2 when the total of the angles were 90^o were very close to each other. The error of θ_1 under the condition of 15^o angle was very close to the error of θ_2 under the condition of 75^o. Similarly, the error of θ_1 under the condition of 30^o angle was very close to the error of θ_2 under the condition of 60^o angle. Therefore, the errors obtained for both θ_1 and θ_2 under similar conditions and under the condition of 45^o angle were close to each other.

***θ_{avg}* parameter:**

The RMSE values obtained for the θ_{avg} parameter are presented in Table 8. Table 8 demonstrates the errors pertaining to the θ_{avg} parameter, which is the average of the θ_1 and θ_2 parameters. According to the table, the variations in angles and correlations affected the errors pertaining to the θ_{avg} parameter. However, this effect was not as high as in θ_1 and θ_2; yet, it was lower. Similarly, the errors decreased as the correlation increased. The increase of the angles had a varying effect on the errors. Accordingly, under all conditions, the errors initially decreased and then increased as the angles increased. The lowest errors were obtained under the conditions of 45^o angles. Variation in distributions did not significantly affect the error of θ_{avg}. Errors obtained in standard normal distribution had the lowest values while similar errors were obtained in positively and negatively-skewed distributions. This finding is similar to the one found for θ_1 and θ_2. The criterion RMSE values were found to be lower than the RMSE values obtained for multidimensional tests under all conditions. The condition in which the criterion values and the errors pertaining to the multidimensional data was closest to each other when the angles were 45^o.

ANOVA results about the comparison of results

According to ANOVA results, the average errors of discrimination parameter varied in accordance with distributions (for a_1 [$F_{2,7497}$=16.700, p<.05]; for a_2 [$F_{2,7497}$=150.015, p<.05]; for a_{avg} [$F_{2,7497}$=2960.506, p<.05]; for MDISC [$F_{2,7497}$=1679.966, p<.05]). Based on the results of post hoc comparisons, there was not any significant difference between errors obtained under positively and negatively skewed distribution conditions for a_1 and a_2, and the errors obtained under normal conditions were smaller. For MDISC and a_{avg}, errors obtained for all distribution conditions were different from each other; the lowest error values were obtained under standard normal distribution and the highest error values were obtained under negatively skewed distribution.

According to ANOVA results, the average errors of discrimination parameter varied by interdimensional correlation (for a_1 [$F_{4,7495}$=3.754, p<.05]; for a_2 [$F_{4,7495}$=3.279, p>.05]; for a_{avg} [$F_{4,7495}$=149.596, p<.05]; for *MDISC* [$F_{4,7495}$=224.635, p<.05]). Based on the conducted post hoc comparisons, for a_1, there was a significant difference only between errors obtained in correlation of 0.15 and 0.75. According to this, error values obtained under 0.15 correlation condition were lower. For a_2, it was observed that the errors obtained under the condition where correlation was 0.30 were higher than the errors obtained under the conditions where correlations were 0.15 and 0.75. No significant difference was obtained among the errors apart from other conditions. For a_{avg} and *MDISC*, errors obtained under all correlation conditions were not different from each other. According to this, the highest errors were obtained in 0.15 correlation value, and the lowest errors were obtained in 0.75 correlation value.

It was determined that the average errors of discrimination parameter varied by angles (for a_1 [$F_{4,7495}$=9211.581, p<.05]; for a_2 [$F_{4,7495}$=7896.183, p<.05]; for a_{avg} [$F_{4,7495}$=736.080, p<.05]; for *MDISC* [$F_{4,7495}$=1372.812, p<.05]). Based on the results of post hoc test, errors obtained from all angles were different from each other. When means were examined, for a_1 and a_2, errors got lower up to 45^o, had the lowest value at 45^o, and got higher after 45^o. For MDISC, as angles increased errors also increased; and for a_{avg}, a systematic pattern couldn't be obtained.

According to the results of ANOVA carried out for *D* parameter, the average errors of this parameter varied by distributions [$F_{2,7497}$=917.760, p<.05]. Based on the results of post hoc test, errors obtained from all correlations were different from each other. When means were examined, it was observed that errors obtained under negatively skewed distribution conditions were the highest, and errors obtained under standard normal distribution conditions were the lowest.

According to the results of ANOVA conducted for *D* parameter, the average errors of this parameter varied by interdimensional correlation [$F_{4,7497}$=81.988, p<.05]. Base on the results of post hoc comparisons, errors obtained from all correlation values were different from each other. When means were examined, in general, as interdimensional correlation increased, errors also increased.

Finally, it was determined that the average errors of *D* parameter varied by angles [$F_{4,7495}$=69.682, p<.05]. Based on the results of post hoc test, only the errors under conditions in which the angles were 30^o and 60^o were not different from each other. Errors obtained under all other conditions were different from each other.

According to the results of ANOVA, it was determined that errors of ability parameter varied by distributions (for θ_1 [$F_{2,7497}$=67.582, p<.05]; for θ_2 [$F_{2,7497}$=61.608, p<.05]; for θ_{avg} [$F_{2,7497}$=344.435, p<.05]). Based on the results of post hoc comparisons, for ability parameter, there was not any difference in positively and negatively skewed distributions; errors obtained under standard normal distribution conditions were lower.

According to the results of ANOVA, the errors of ability parameter varied by correlations (for θ_1 [$F_{4,7495}$=448.577, p<.05]; for θ_2 [$F_{4,7495}$=349.489, p<.05]; for θ_{avg} [$F_{4,7495}$=310.452, p<.05]). Based on the results of post hoc comparisons, errors obtained from all correlation values were different from each other. When means were analyzed, as interdimensional correlation for all ability parameters under all conditions increased, errors decreased.

Finally, according to the results of ANOVA, the average errors of ability parameter varied by angles (for θ_1 [$F_{4,7495}$=4737.972, p<.05]; for θ_2 ([$F_{4,7495}$=6193.641, p<.05]; for θ_{avg} [$F_{4,7495}$=4705.022, p<.05]). Based on the results of post hoc comparisons, errors obtained from all correlation values were different from each other. When means were analyzed, it was observed that for θ_1, as angles increased, errors also increased; for θ_2 and θ_{avg}, as angles increased, errors decreased.

Table 7. RMSE Values for θ_2 Parameter

		Correlation of Abilities														
Angles	**Results of Unidimensional data**	**ρ_1=0.15**			**ρ_1=0.30**			**ρ_1=0.45**			**ρ_1=0.60**			**ρ_1=0.75**		
		SND*	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****
15^0	0.053	1.157	1.173	1.171	1.055	1.099	1.088	0.940	1.007	0.998	0.820	0.902	0.891	0.688	0.770	0.759
30^0	0.081	0.927	0.935	0.945	0.842	0.876	0.884	0.753	0.804	0.818	0.660	0.721	0.734	0.557	0.625	0.636
45^0	0.123	0.751	0.779	0.779	0.687	0.732	0.734	0.621	0.677	0.680	0.551	0.620	0.622	0.474	0.550	0.551
60^0	0.090	0.574	0.621	0.616	0.537	0.595	0.591	0.498	0.498	0.560	0.456	0.525	0.525	0.410	0.483	0.481
75^0	0.095	0.414	0.475	0.477	0.402	0.467	0.470	0.387	0.454	0.458	0.372	0.441	0.445	0.355	0.428	0.431

*SND: Standard Normal Distribution, **PSD: Positive Skewed Distribution, ***NSD: Negative Skewed Distribution

Table 8. RMSE Values for θ_{avg} Parameter

		Correlation of Abilities														
Angles	**Results of Unidimensional data**	**ρ_1=0.15**			**ρ_1=0.30**			**ρ_1=0.45**			**ρ_1=0.60**			**ρ_1=0.75**		
		SND*	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****	**SND***	**PSD****	**NSD*****
15^0	0.053	0.586	0.604	0.601	0.550	0.581	0.573	0.511	0.555	0.546	0.475	0.527	0.519	0.442	0.499	0.489
30^0	0.081	0.426	0.447	0.453	0.401	0.431	0.439	0.379	0.419	0.428	0.364	0.409	0.419	0.351	0.404	0.414
45^0	0.123	0.367	0.400	0.399	0.347	0.387	0.390	0.330	0.383	0.383	0.320	0.377	0.381	0.314	0.380	0.381
60^0	0.090	0.414	0.439	0.442	0.386	0.424	0.427	0.363	0.363	0.414	0.345	0.402	0.404	0.333	0.396	0.396
75^0	0.095	0.527	0.545	0.546	0.486	0.519	0.521	0.448	0.494	0.496	0.411	0.465	0.469	0.376	0.438	0.442

*SND: Standard Normal Distribution, **PSD: Positive Skewed Distribution, ***NSD: Negative Skewed Distribution

DISCUSSION and CONCLUSION

The studies in the literature have suggested that errors pertaining to discrimination parameter increase as the correlation between the dimensions increases (Ansley & Forsyth, 1985; Ackerman, 1989; Zhang, 2008). In this study, *MDISC*, one of the discrimination parameters, displayed such a pattern. In addition to *MDISC*, the errors pertaining to the a_1 parameter under the conditions that items' angles were smaller than 45^o were in line with these studies in the literature, and an opposite pattern was observed on the error values for the conditions with angles, higher than 45^o. Since the a_2 parameter had an opposite pattern with a_1, the a_2 parameter under the conditions of angles larger than 45^o is in line with these studies in the literature, and the errors decreased as the correlation among the dimensions increased under these conditions. Thus, it can be suggested that in this study, the most noticeable value especially for the a_1 and a_2 parameters was the 45 point (45^o angle and 0.45 correlation). The RMSE values calculated for a_{avg}, which is the average of the a_1 and a_2 parameters, showed a different pattern than the existing studies' values in the literature. Accordingly, the lowest errors for a_1, a_2 and a_{avg} were generally obtained under the conditions in which the angle was 45^o, and the errors pertaining to the a_1 and a_2 parameters produced a hyperbolic curve when the correlations were kept constant. Gocer Sahin, Walker, and Gelbal (2015) and Gocer Sahin (2016) reported that the average angles of the items they used for their studies were around 45^o. The errors pertaining to the discrimination parameter produced a hyperbolic curve in these authors' studies, too. In this respect, the findings obtained in this study are in line with the studies of Gocer Sahin, Walker, and Gelbal (2015) and Gocer Sahin (2016). If the angles were bigger than 45^o, then the errors increased as the correlation increased. And, this finding was consistent with the findings of Kahraman (2013). All the discussions above are valid for the conditions in which distributions are standard normal; while the pattern obtained in skewed distributions is similar to the one in the standard normal distribution, the conditions in which the lowest RMSE values were obtained in skewed distributions are different.

Although the pattern of the a_1 and a_2 parameters were found to be contrary to previous studies in the literature, the *MDISC* parameter had a pattern that is similar to the ones reported in the studies of Ansley and Forsyth (1985), Ackerman (1989), Zhang (2008), Gocer Sahin, Walker, and Gelbal (2015), Gocer Sahin (2016). According to findings, the errors decreased as the correlation among the dimensions increased. Besides, as the angles of the items increased, (i.e. as the complexity of the items increased), the RMSE values increased. This is an expected result since *MDISC* corresponds to the discrimination of the multidimensional IRT model when it is considered a unidimensional IRT model.

With the 45^o angle being the breakpoint, when the angles for a_1 were higher than 45^o (when the angles are 60^o and 75^o), the RMSE values obtained under the conditions of standard normal distribution were found to be higher than the errors obtained under the conditions of skewed distribution. When the angle for a_1 was 45^o or smaller, the error values obtained in conditions with the skewed ability distributions were higher. The pattern for the a_2 parameter was exactly the opposite of this pattern. It can be suggested that the a_1 and a_2 parameters were not generally affected by the skewed distribution. Although skewed distributions did not affect a_1 and a_2 parameters, the a_{avg} and *MDISC* parameters were affected by skewed distributions. The RMSE values obtained for the a_{avg} and *MDISC* parameters under all conditions of standard normal distribution were lower than the RMSE values obtained under the conditions of skewed distribution, but this difference was not very large. It was also mentioned in the study of Kirisci, Hsu, and Yu (2001) that especially the *MDISC* parameter was not affected by skewed distributions. In the studies of Gocer Sahin, Walker, and Gelbal (2015) and Gocer Sahin (2016), in which the distributions were manipulated as standard normal or only normal, it was reported that the mentioned distributions did not affect the discrimination parameter.

45^o angle and 0.45 correlations can be suggested to be the critical values for the discrimination parameters of the tests with a semi-mixed structure, especially for the a_1, a_2 and a_{avg} parameters. If a test parameter with few errors is desired in the estimation of a multidimensional test with a semi-mixed structure as unidimensional, then it can be recommended to use a test in which the items' angles are 45^o. If the correlation is 0.45 in such a test, then it is possible to obtain minimum errors.

As for the errors pertaining to the difficulty parameter obtained when the two-dimensional tests were estimated as unidimensional, it is observed that the errors increased as the correlation among the dimensions increased. In the case of standard normal distributions, the lowest error occurred when the correlation among the dimensions was 0.15 while the highest error occurred when the correlation was 0.75. However, no regular pattern was found regarding the errors under the condition of skewed distributions. Almost all of the errors that were obtained by estimating the two-dimensional structures as unidimensional were lower than the criterion value.

The errors obtained for difficulty parameter were generally lower than errors of other parameters. According to that result it can be concluded that difficulty parameter is the robust parameter. This result is similar to the literature. It did not matter whether the distribution was positively or negatively skewed for the difficulty parameter; instead, the main concern was whether the distribution was standard normal or not.

The errors for ability parameters increased as correlation between dimensions increased. This result is similar to the literature (Ackerman, 1989; Ansley & Forsyth, 1985; Doody, 1985; Drasgow & Parsons, 1983; Gocer Sahin, 2016; Zhang, 2008). Interestingly, although items' angles increased the RMSE decreased for θ_1, and although items' angles decreased the RMSE increased for θ_2. It did not matter whether the distribution was positively or negatively skewed for the ability parameters; instead, the main consideration was whether the distribution was standard normal or not. Because when the distributions were skewed, higher errors were obtained than standard normal distributions. This result is similar to the literature. For example, in Gocer Sahin's (2016) study, errors for θ_{avg} were between the errors for θ_1 and θ_2.

Limitations and Suggestions

This study is limited by its research design that has two dimensional data, and two-parameter logistic and compensatory model. The generalizability of the results is limited to the studied conditions; which were a test with 25 items, a sample size with 2,000 examinees, correlations between dimensions with 0.15, 0.30, 0.45, 0.60 and 0.75; angles that items have with *x* axis are 15^o, 30^o, 45^o, 60^o and 75^o; and lastly, distributions which were standard normal, positively skewed and negatively skewed. Another limitation of this study is that the results are based on only one software. Multiple software programs may result in differences in parameter estimates. In this study only RMSE statistics was used to evaluate the results. Bias or other statistics could also be calculated for this purpose.

Based on the conditions of this study, a multidimensional test which has a high correlation between dimensions is suggested for the researchers who aim to scale the abilities of individuals to a one-level scale. However, if the aim is to develop a qualified test, for a two-dimensional test, items that have 0.45 interdimensional correlation and have 45^o angles with *x* axis should be used. If the estimation is carried out through BILOG program, ability distribution should be standard normal or normal.

REFERENCES

Abdel-fattah, A. A. (1994, April). *Comparing BILOG and LOGIST estimates for normal, truncated normal, and beta ability distributions.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Ansley, T. N., & Forsyth, R. A. (1985). An examination of the characteristics of unidimensional IRT parameter estimates derived from two-dimensional data. *Applied Psychological Measurement*, *9*(1), 37-48. https://doi.org/10.1177/014662168500900104

Ackerman, T. A. (1989). Unidimensional IRT calibration of compensatory and noncompensatory multidimensional items. *Applied Psychological Measurement, 13*(2), 113-27. https://doi.org/10.1177/014662168901300201

Ackerman, T. A., Gierl, M. J., & Walker, C. M. (2003). Using multidimensional item response theory to evaluate educational and psychological tests. *Educational Measurement: Issues and Practice*, *22*(3), 37-51. https://doi.org/10.1111/j.1745-3992.2003.tb00136.x

Bulut, O. (2013). *Between-person and within-person subscore reliability: Comparison of unidimensional and multidimensional IRT models* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database (UMI No. 3589000).

Davey, T., Nering M. L. & Thompson, T. (1997). *Realistic simulation of item response data* (ACT Research Report 97-4). Iowa City, IA: ACT, Inc.

Doody, E. N. (1985, April). *Examining the effects of multidimensional data on ability and item parameter estimation using the three-parameter logistic model*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Drasgow, F., & Parsons, C. K. (1983). Application of unidimensional item response theory models to multidimensional data. *Applied Psychological Measurement*, *7*, 189-199. https://doi.org/10.1177/014662168300700207

Fleishman, A. I. (1978). A method for simulating non-normal distributions. *Psychometrika, 43,* 521-532. https://doi.org/10.1007/BF02293811

Gocer Sahin, S. (2015). *Yarı karışık yapılı çok boyutlu yapıların tek boyutlu olarak ele alınması durumunda kestirilen parametrelerin incelenmesi* (Doctoral Dissertation). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/giris.jsp

Gocer Sahin S., Walker, C. M., & Gelbal, S. (2015). The Impact of model misspecification with multidimensional test data. In L. A. van der Ark, D. M. Bolt, S. M. Chow, J. A. Douglas & W. C. Wang (Eds.), *Quantitative Psychology Research: The 79th Annual Meeting of the Psychometric Society* (pp.133-44). New York, NY: Springer.

Hambleton, R. K. (1989). Principles and selected applications of item response theory. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 147–200). New York, NY: Macmillan.

Harrison, D. A. (1986). Robustness of IRT parameter estimation to violations of the unidimensionality assumption. *Journal of Educational Statistics, 11*(2), 91-115. doi: 10.2307/1164972

Kahraman, N. (2013) Unidimensional interpretations for multidimensional test items. *Journal of Educational Measurement, 50*(2), 227-246. https://doi.org/10.1111/jedm.12012

Kim, K. Y., & Lee, W. (2014, July). *Recovery of item parameters under various IRT item calibration methods*. Paper presented at the International Meeting of the Psychometric Society, Madison, WI.

Kirisci, L., Hsu, T., & Yu, L. (2001). Robustness of item parameter estimation programs to assumptions of unidimensionality and normality. *Applied Psychological Measurement, 25*(2), 146-162. https://doi.org/10.1177/01466210122031975

Leucht, R. M & Miller, T. R. (1992). Unidimensional calibrations and interpretations of composite traits for multidimensional tests. *Applied Psychological Measurement, 16*, 279-293. https://doi.org/10.1177/014662169201600308

Reckase, M. D. (2009). *Multidimensional item response theory (statistics for social and behavioral sciences)*. New York: Springer.

Reckase, M. D., Ackerman, T. A., & Carlson, J. E. (1988). Building a unidimensional test using multidimensional items. *Journal of Educational Measurement, 25*, 193-203. https://doi.org/10.1111/j.1745-3984.1988.tb00302.x

Reckase, M. D., & McKinley, R. L. (1991). The discriminating power of items that measure more than one dimension. *Applied Psychological Measurement, 15*(4), 361-373. https://doi.org/10.1177/014662169101500407

Seong, T-J. (1990). Sensitivity of marginal maximum likelihood estimation of item and ability parameters to the characteristics of the prior ability distributions. *Applied Psychological Measurement, 14*(3), 299-311. https://doi.org/10.1177/014662169001400307

Sheng Y. & Wikle C. K. (2007). Comparing multidimensional and unidimensional item response theory models. *Educational and Psychological Measurement, 68*(3), 413-430. https://doi.org/10.1177/0013164406296977

Toland, M. (2008). *Determining the accuracy of item parameter standard error of estimates in BILOG-MG 3.* (Doctoral Dissertation). Avaible from ProQuest Dissertations and Theses database (UMI No. 3317288).

Yen, M. Y. (1987). A comparison of the efficiency and accuracy of BILOG and LOGIST. *Psychometrika*, *52*, 275-291.

Zhang, B. (2008). Application of unidimensional item response models to tests with items sensitive to secondary dimension. *The Journal of Experimental Education*, *77*(2), 147-166.

Zhang, J. (2005). *Estimating multidimensional item response models with mixed structure* (ETS Research Report 05-04). Princeton, NJ: Educational Testing Service.

Zhang, J. (2012). Calibration of response data using MIRT models with simple and mixed structures. *Applied Psychological Measurement, 36*(5), 375-398. https://doi.org/10.1177/0146621612445904

Boyutlar Arası Korelasyon ile Madde Ayırt Ediciliği Arasındaki Etkileşimin Parametre Kestirimi Üzerine Etkisi

Giriş

Testlerin uygulanması, verilerin analizi ve yorumlanmasından önce test boyutluluğunun titizlikle incelenmesi gerekir. Tek boyutluluk sayıltısının MTK için bu denli önemli olması ve tek boyutluluğa dayanan modellerin uygulanması ve yorumlanmasının daha kolay olması, araştırmacıları çok boyutlu modellerin tek boyutlu olarak ele alındığı çalışmalara yöneltmektedir. Çok boyutlu testlerin tek boyutlu olarak kestirilmesi ile ilgili çalışmaların 1980'li yıllardan itibaren yapıldığı görülmektedir. Bu tür çalışmalar genel olarak modelin yanlış tanımlanması (model misspecification) olarak adlandırılmaktadır.

Modeli yanlış tanımlama çalışmalarında incelenen koşullardan biri testin yapısı olup (yaklaşık basit veya karmaşık) bunun dışında en çok ele alınan koşullar, boyutlar arası korelasyon ve dağılımların çarpıklığıdır (Ackerman, 1989; Ansley ve Forsyth; 1985; Drasgow ve Parsons; 1983; Harrison, 1986; Kirisci, Hsu ve Yu, 2001, Leucht ve Miller, 1992; Reckase, Ackerman ve Carlson, 1988; Zhang, 2008; Zhang, 2012). Kahraman (2013) tarafından yapılan bir çalışmada, çok boyutlu bir testin ikinci boyutunun ihmal edilerek tek boyutlu kestiriminde, korelasyon arttıkça ayırt ediciliğe ait hatanın arttığı belirtilmiştir.

Son yıllarda yapılan çalışmalarda yarı karışık (semi-mixed) yapılı çok boyutlu testlerin tek boyutlu olarak kestirilmesinde, boyutlar arası korelasyon arttıkça madde parametrelerine ait hataların da artmasının, maddelerin analitik düzlemdeki açıları ile boyutlar arasındaki korelasyonun etkileşiminin bir sonucu olduğu düşünülmektedir. Bu çalışma, bu hipotezin doğru olup olmadığını test etmek üzere yapılmıştır. Dolayısıyla bu çalışmanın amacı, iki boyutlu testlerin tek boyutlu olarak ele alınması durumunda kestirilen parametrelerin, farklı yetenek dağılımları, boyutlar arası korelasyon ve maddenin x ekseni ile yaptığı açı değişkenlerinin kombinasyonlarından nasıl etkilendiğini belirlemektir.

Yöntem

Bu çalışmada, bir testte yer alan maddelerin x ekseniyle yaptığı açılar ile boyutlar arası korelasyonlar manipüle edilerek, boyutlar arası korelasyon ile maddelere ait açıların etkileşiminin parametre kestirimi üzerine etkisi incelenmiştir. Çalışmada simülasyon yoluyla yarı karışık yapılı, 25 maddeden oluşan iki boyutlu testler üretilmiştir. Tüm desende örneklem büyüklüğü 2000 olacak şekilde sabitlenmiştir. Ele alınan iki boyutlu testlerde yetenek parametreleri arasındaki korelasyon düşük ilişkiden yüksek ilişkiye doğru sıralanacak biçimde (ρ=0,15; ρ=0,30; ρ=0,45; ρ=0,60; ρ=0,75) değişimlenmiştir.

Bu çalışmada madde vektörlerinin x ekseniyle yaptığı açı, korelasyonlar ile aynı sayısal değerlerde olmak üzere 15^o, 30^o, 45^o, 60^o ve 75^o şeklinde manipüle edilmiştir. Bu şekilde oluşturulan desende açıları (15^o ve 30^o) olan maddeler öncelikli olarak θ_1 yeteneğini, açıları 45^o olan maddeler hem θ_1 hem θ_2 yeteneğini ve açıları 60^o ve 75^o olan maddeler ise öncelikli olarak θ_2 yeteneğini ölçmektedir. Yetenek parametreleri ise standart normal, sağa çarpık ve sola çarpık dağılım olmak üzere üç farklı dağılımdan elde edilmiştir.

Bu şekilde düzenlenen çalışmada yetenek dağılımları 3; maddelerin x ekseniyle yaptığı açılar 5 ve boyutlar arası korelasyon 5 koşul olmak üzere toplam (3 x 5 x 5) 75 hücreli bir desen oluşturulmuştur. Veriler, SAS programı aracılığıyla telafisel, 2 parametreli lojistik modele dayanarak üretilmiştir. Veri üretiminde 100 replikasyon yapılmıştır.

Çok boyutlu yapıların tek boyutlu olarak ele alınması durumunda kestirilen parametrelerin içerdiği hataların değerlendirilmesinde RMSE istatistiğinden faydalanılmıştır. RMSE değerleri, tüm parametreler için ayrı ayrı hesaplanmıştır.

Çalışmada çok boyutlu testler dışında gerçekte tek boyutlu olan 25 maddeli ve 2000 kişilik bir test tek boyutlu olarak kestirilmiştir. Tek boyutlu test oluştururken, çok boyutlu testlere ait parametrelerden yararlanılmıştır. Buna göre çok boyutlu testlere ait *MDISC* ve *D* parametresi, tek boyutlu teste ait gerçek *a* ve *b* parametrelerini oluşturmuştur.

Sonuç ve Tartışma

Literatürde yapılan çalışmalarda boyutlar arası korelasyon arttıkça ayırt ediciliğe ait hataların azaldığı belirtilmiştir (Ackerman,1989; Ansley ve Forsyth, 1985; Zhang, 2008). Bu çalışmada ise ayırt edicilik parametrelerinden *MDISC*'in bu örüntüye sahip olduğu görülmüştür. *MDISC*'in yanı sıra maddelerin açılarının 45^o'den küçük olduğu koşullarda a_1 parametresine ait hatalar alan yazındaki bu çalışmalar ile paralellik göstermekte, boyutlar arası korelasyon arttıkça hatalar azalmaktadır. a_2 parametresi a_1 ile ters bir örüntü göstermiştir. Bu çalışmada özellikle a_1 ve a_2 parametreleri için en önemli değerin 45 noktası (45^o'lik açı ve 0,45 korelasyon) olduğu söylenebilir. a_1 ve a_2 bu iki parametrenin ortalaması olan a_{ort} için hesaplanan RMSE değerleri alan yazından farklı bir örüntü göstermiştir. Çarpık dağılımlarda elde edilen örüntü standart normal dağılım ile benzer olmakla birlikte çarpık dağılımlarda en düşük RMSE değerlerinin elde edildiği koşullar farklılık göstermektedir.

Bu çalışmada a_1 için açıların 45^o'den (açılar, 60^o ve 75^o) yüksek ve dağılımın standart normal olduğu koşullarda elde edilen RMSE değerleri, çarpık dağılım koşullarında elde edilen hatalardan daha yüksek olmakla beraber bu fark çok fazla değildir. a_1 için açı 45^o ve 45^o'den küçükken çarpık dağılımlarda elde edilen hata değerleri daha yüksektir. Bu durum a_2 parametresi için tam tersidir. Ancak yine de genel olarak çarpık dağılımın a_1 ve a_2 parametresini etkilemediği söylenebilir. Her ne kadar çarpık dağılımlar a_1 ve a_2 parametrelerini etkilemese de a_{ort} ve *MDISC* parametreleri çarpık dağılımlardan etkilenmektedir. Dağılımın standart normal olduğu bütün koşullarda a_{ort} ve *MDISC* parametreleri için elde edilen RMSE değerleri çarpık dağılım koşullarındaki RMSE değerlerinden düşüktür.

Yarı karışık yapılı testler için özellikle a_1, a_2 ve a_{ort} parametrelerine ilişkin açının 45^o ve boyutlar arası korelasyonun 0,45 olduğu koşulların kritik RMSE değerine sahip olduğu söylenebilir. Buna göre çok boyutlu yarı karışık yapılı bir test tek boyutlu olarak kestirildiğinde, madde açılarının 45^o olduğu testlerde test parametresinin düşük miktarda hata içerdiği görülmüştür. Bu test ile beraber boyutlar arası korelasyon 0,45 olduğunda ise hatalar en düşük değerlerini almıştır.

Güçlük parametresi için elde edilen hata değerleri, diğer parametrelerinkinden genel olarak daha azdır. Buna göre bu çalışmada da alan yazına benzer olarak güçlük parametresinin daha dayanıklı olduğu söylenebilir. Güçlük parametresi için de dağılımın sağa veya sola çarpık olması önemli olmayıp; dağılımın standart normal olması veya olmaması önemlidir.

Yetenek parametrelerine ait hatalar, boyutlar arası korelasyon arttıkça azalmıştır. Bu bulgu alan yazındaki benzer çalışmalar ile paraleldir (Ackerman, 1989; Ansley & Forsyth, 1985; Doody, 1985; Drasgow & Parsons, 1983; Gocer Sahin, 2016; Zhang, 2008). θ_1 için maddelerin açıları arttıkça hatalar artmasına rağmen, θ_2 için açı arttıkça hatanın azalması ilginç bir sonuçtur. Yetenek parametreleri için dağılımın sağa veya sola çarpık olması önemli olmamakla birlikte dağılımın standart normal olması önemli bir koşuldur. Çünkü dağılım çarpıklaştığında yetenek parametrelerine ait hatalar artmaktadır. Bu durum alan yazın ile benzerlik göstermektedir. Gocer Sahin (2016)'nın çalışmasına benzer olarak θ_{ort} için elde edilen hata değerleri θ_1 ve θ_2 için elde edilen hataların arasında değer almıştır.