



7th World Conference on Educational Sciences, (WCES-2015), 05-07 February 2015, Novotel Athens Convention Center, Athens, Greece

Invariability Of Item Parameters In The Sample With Restricted Real Data Results*

Nuri Dogan^{a*}, Meltem Yurtcu^a

^aHacettepe University, Beytepe Campus, Ankara, 06800, TURKEY

Abstract

There are two basic theories used in testing and item analysis in education and psychology, and they are called as the Classical Test Theory (CTT) and the Item Response Theory (IRT). Both theories have advantages and disadvantages changing according to the situation in which they are used. The most important advantage of the IRT, and its superiority to the CTT is that in the CTT the item parameters are based on sample and the ability parameters are based on the test. There have been many studies conducted to determine whether the IRT provides this advantage or not. However, in these studies, many different results have been reported. The item parameter invariance being multidirectional shows that continuing these types of studies is necessary. This study is based on this fact, and has been designed as a simulation study in which the distribution and parameter characteristics of the real data are used as restrictions. The number of the items have been kept stable, the sample size, the restrictions of distribution and item parameters, and the parameter prediction models have been taken as the simulation models. According to the results obtained, the restriction of the parameters and increasing the sample size have been determined as important factors in ensuring the parameter invariability.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of Academic World Education and Research Center.

Keywords: parameters invariance, irt, simulation study

1. Introduction

Tests are frequently made use of in the educational and psychological fields. Tests are used especially in determining the success levels, in selection and placement procedures of the students, guidance, defining and

* Nuri Dogan. Tel.: +90 312 297 62 00.
E-mail address: nuridogan2004@gmail.com

following the student characteristics, following them in their careers, granting certificates or diplomas, determining the proficiency levels etc. However, for whatever purpose tests are used, it is crucial that they are proven to be proper for the specific purpose and ensure a sensitive and sufficient measurement. In other words, it must be proven that tests are valid and reliable. The soundness of a test depends on its psychometric properties. Determining the psychometric properties of a test means analyzing the test in terms of the contents and items.

Various theories have been developed so far throughout the history to determine whether tests are valid, reliable and ensure the demanded properties. These theories have also been used in producing such tests. There are two frequently-used theories in producing the tests in education and psychology. These are the Classical Test Theory (CTT) and the Item Response Theory (IRT). Both theories have advantages and disadvantages. The most important advantage of the IRT and its superiority to the CTT is that in the CTT, the item parameters are based on sample, and the skill parameters are based on the test (its strength).

The IRT was started to be developed in mid-1900s after the criticism on the insufficient sides of the CTT, and only became widespread with the improvements in the computer technology and software development at the end of 1900s (Hambleton & Swaminathan, 1985).

The IRT establishes a relation based on the non-linear (logistic) function between the skills of an individual and the possibility of his/her replying an item accurately. The IRT has strong hypotheses (single-dimension, local dependency, etc.) that may be tested according to the CTT. The mathematical model (logistic function) used by the IRT supports the invariability of the parameter. In the CTT, the parameters like the item difficulty (p), distinctiveness (r), and possibility of answering the item by chance (c) which are calculated for the items vary according to the group to which the test is applied and to the skill levels of the individuals who will receive the test. However, according to the IRT, in case an agreement has been ensured between the factors such as being single-dimensional, local independency, the form of the item characteristic function and the model data is ensured (Hambleton, Swaminathan & Rogers, 1991), the item parameters such as the difficulty of the item (b) discrimination of the item (a) and the guessing parameter (c) will not change. In such a situation, the validity and reliability of the tests will be obtained in a high level and more accurate decisions will be made about the individuals.

1.1. The question of the research

There have been many studies conducted to determine whether the IRT provides the advantage on the invariability of the item and ability parameters. In many of these studies, simulation has been used in order to examine the invariability of the item parameters in the CTT and IRT (Galdin & Laurencelle, 2010; Adedoyin, Nenty & Chilissa, 2008; Dogan & Tezbasaran, 2003; Nartgun 2002; Fan 1998; Somer, 1998; Stage 1998). According to the reports of these studies, in one study (Adedoyin, Nenty & Chillisa, 2008) it was found that the IRT provides the invariability property more than the CTT in terms of item difficulty; and in other studies it was found that the CTT provides invariability as much as the IRT in terms of item difficulty, and it was also found that the IRT cannot provide the invariability especially in terms of a and c parameters.

Based on various findings in the literature, it has become clear that the claim of the IRT providing the invariability of the parameters still needs further studies. The conditions influencing the item parameter invariability and the question of what precautions need to be taken in order to provide the item parameter invariability are among the most important issues on which the psychometric specialists who make use of IRT emphasize and examine continuously. So far no studies have been able to produce satisfying and consistent answers to the abovementioned problems. On the other hand, many variables, conditions and combinations of these that may affect the invariability of item parameters have not been used in the studies conducted so far. Depending on the abovementioned problems, in this study, an answer have been sought for the question "How does the use of real data result as the restrictions in producing simulation data affect the invariability of the parameters in the models with 1, 2 and 3 parameters

depending on the 100, 250 and 500 sample size?" In order to answer this critical question, the 3 sub-problems have been formed:

1. How are the bias error values calculated for the model b parameter with 1 parameter in 100, 250 and 500 exemplification size when the point distribution and item parameters are limited and unlimited?
2. How are the bias and error values calculated for the model a and b parameters with 2 parameters in 100, 250 and 500 exemplification size when the point distribution and item parameters are limited and unlimited?
3. How are the bias and error values calculated for the model a, b and c parameters with 3 parameters in 100, 250 and 500 exemplification size when the point distribution and item parameters are limited and unlimited?

2. Method

2.1. Type of research;

The study is a simulation study. In a simulation study, there is a target of limiting the distribution and parameters during the data production. The real data results have been used in the study with the thought that instead of determining the restrictions arbitrarily, using the values obtained from real data would be beneficial in terms of providing more information to test developers who are making field practice. For this purpose, the results of the analyses of the answers given to the Mathematics Sub-test consisting of 25 questions in the Booklet A in the Student Selection and Placement Exam for Secondary Schools (i.e. the OKS Exam in Turkey) that is applied to the 8th Graders by the Republic of Turkey, Ministry of National Education were used.

Firstly, the descriptive statistics on the data of the Booklet A of the Student Selection Exam, the study for determining whether IRT assumptions have been covered or not, and the IRT analyses were performed to be used in the conditions in the simulation studies. The descriptive statistics of the data are examined, it becomes obvious that the distribution is positively skewed and sharp, the group is accumulated in small points, the range is in maximum width, internal consistency reliability of the test is high at an acceptable level, and the average article discrimination strength is extremely high.

In order to determine whether the data cover the unidimensionality condition and local independency condition, which are among the MTK Conditions, the exploratory factor analysis has been made by using the Principal Axes Method. According to the factor analysis results, one factor was found whose Eigen Value is over 1, and there is a significant difference between the first eigenvalue (5.843) and the second eigenvalue (0.853). On the other hand, the first Eigen Value explains the 23,38% of the variance in the points. According to Hambleton, Swaminathan and Rogers (1991), in order to call a test as a unidimensional test, it is sufficient that the difference between the first Eigen Value and the second Eigen Value is 4-fold, and the first Eigen Value explains the 20% of the variability. According to the results of the analysis, it was determined that the loads of all the articles in the first factor are higher than 0,30. Again, according to the same authors, the proof of unidimensionality is also a proof for local independency. Based on these findings, it has been concluded that the data on the Student Selection and Placement Exam for Secondary Schools (i.e. OKS in Turkey) 2008 Booklet A are proper for the IRT analyses.

The analysis results of the mathematics sub-test consisting of 25 questions in the Student Selection and Placement Exam for Secondary Schools (the OKS Exam) 2008, Booklet A, according to the IRT, 1, 2 and 3 parameter models are given in Table 1.

Table 1. Summary of the results of the OKS 2008 Mathematics Exam Sub-test Item Analysis Theory

Model	Parameter	Mean	S.Dev.	Min.	Max.	Test info
1PLM*	b	0.932	0.59	-0.719	2.09	8.29
2PLM**	a	0.772	0.297	0.284	1.311	10.88
	b	1.078	0.905	-0.647	3.456	
3PLM***	a	2.174	0.738	1.059	3.842	41.32
	b	1.141	0.529	-0.102	2.123	
	c	0.139	0.055	0.045	0.271	

*1PLM: Only item difficulty parameter "b" value is calculated, a stable, c is accepted as zero.

**2PLM: Item difficulty parameter "b" and item differentiability parameter "a" is calculated, c is accepted as zero.

***3PLM: Item difficulty parameter “b”, item differentiability parameter “a” and chance parameter “c” is calculated.

2.2. Study Design;

While the study design was being formed, firstly the simulation conditions were decided. For this reason, the stable and changing conditions of the study were determined. Since the OKS Exam 2008 Mathematics sub-test, which will be used in restricting the simulation conditions, consists of 25 questions, the number of the questions was taken as the fixed variate. The positively skewed distribution data was produced as being normal in accordance with the expected value of the ability parameter when the point distribution was unrestricted, and when it was restricted, this data was produced with 7,38 average value and 5,53 standard deviation. The sample size, as it would be observed in usual school practices, was designed so as it would not be very big: 100, 250 and 500. When the restrictions was applied, the Item Parameter values were produced as being in accordance with the normal distribution with real data average and standard deviation for the b parameter; and in accordance with the uniform distribution between the real data maximum and minimum values for the a and c parameters; and when the restriction was not applied, these values were produced as a normal distribution with 0 average and 1 standard deviation for the b parameter; and in accordance with the uniform distribution for the a parameter between 0-2; and in accordance with the uniform distribution for the c parameter between 0-0.25. As the last step, the produced data were analyzed separately for the models with 1, 2 and 3 parameters which are among the Item Response Models which are considered as being proper for the dichotomous data. On the other hand, the number of sample replication for each condition was determined as 40. When the conditions were considered together, the study design is formed is shown in Table 2.

Table 2. The Study Design

Distribution	Parameter values	Model	Sample size	Replication	Numbers of items	Number of file	Total
Limited	Limited	1PLM	100				2*2*3*3*40=1440
Normal	Expected	2PLM	250	40	25		
		3PLM	500				

N.B.: The bold characters in the table are used in forming the names of the files.

2.3. Process;

The processes in the data production and analysis have been performed with certain steps. These steps may be summarized as follows;

Firstly, the wingen-3 (Han & Hambleton, 2007) software was used and 40 pieces of 1-0 answer files for each condition were formed. In the second stage, the answer files produced were analyzed with the Bilog-MG software, and the b parameter for the models with one parameter was calculated; the a and b parameters for the models with two parameters were calculated; and the a, b and c parameters for the models with three parameters were calculated. As the last step, the parameter estimations of the 40 samples (b for one parameter; a and b for two parameters; a, b and c for three parameters) that were calculated the bias values, the Root Mean Square Deviation(RMSD) and correlations.

As a consequence, the mean bias, RMSD and correlations obtained for the 40 repetitions were used in comparing and evaluating the results that were obtained as based on the conditions.

3. Findings

The findings are given in order in this section depending on the sub-problems

3.1. The mean bias and error values calculated for the model with one parameter (parameter b)

The bias calculated for the b parameter in the model with one parameter, Root Mean Square Deviation-RMSD,

and the correlation values are given in Table 3. According to the table, when the bias values calculated for the b values are examined it is observed that the bias values increase in each condition where the distribution is restricted. It is observed that the highest bias value is obtained in the condition where the distribution is restricted and the parameters are as expected, and the smallest bias value is obtained in the condition where the distribution is normal and the parameters are restricted. It was observed that the bias values were extremely low in conditions where the distribution and the parameters were not unrestricted. The size of the sample is extremely influential on the bias values in conditions where the distribution is restricted, while in conditions where the distribution is unrestricted, the sample size does not have an important effect on the bias values.

Table 3. The bias, Root Mean Square Deviation-RMSD, and the correlation values calculated for the b parameter in the model with one parameter

SAMPLES	BIAS	RMSD	R
DRPR100*	0.363	1.865	3.675
DRPR250	0.673	1.24	1.442
DRPR500	0.833	1.033	1.073
DRPE100	0.888	1.895	2.548
DRPE250	0.955	1.514	1.568
DRPE500	0.967	1.248	1.312
DNPE100	0.843	0.007	0.398
DNPE250	0.899	0.047	0.318
DNPE500	0.938	0.054	0.270
DNPE100	0.961	0.142	0.528
DNPE250	0.978	0.167	0.320
DNPE500	0.989	0.161	0.239

*N.B: While the names of the files are being read in the tables, the first and second character show the situation of the distribution; the third and fourth character show the situation of the parameter; and the latest three characters show the situation of the sample size. In this context, the first sample in Table 3 should be read like this: The mean bias value at the condition where the distribution is restricted, the parameter is restricted, the sample size is 100 for the model b parameter with one parameter. Again, the last sample in Table 3 should be read like this: The mean bias value at the condition where the distribution is normal, the parameter is expected, the sample size of the exemplification is 500 for the model b parameter with one parameter.

When the RMSD for the model with one parameter given in Table 3 is examined, it is observed that there is a similarity with the results on the bias values. The RMSD values increase in conditions where the distribution is restricted, while they decrease in conditions where the distribution is unrestricted. The biggest RMSD value is obtained in the condition where the distribution and parameters are restricted, while the smallest RMSD values are obtained in the condition where the distribution and parameter are unrestricted. It is observed that the sample size has a significant effect on the RMSD values in conditions where the distribution is restricted, on the other hand, it is also observed that it does not have a significant effect in conditions where the distribution is unrestricted. It is clear that the results of the “condition where the distribution is normal and the parameter is restricted” and the results of the “condition where the distribution is normal and the parameter is expected” are highly similar to each other.

When the correlation values given in Table 3 are examined, it is observed that the smallest correlation average calculated for the b value in model with one parameter is obtained in the condition where “the distribution and parameters are restricted and the sample size of the exemplification is 100”. The biggest correlation average value is obtained for the condition where “the distribution and parameters are not restricted and the sample size is 500”. In a condition where the distribution and parameter are restricted, the sample is expanded in order to expand the correlations in a sufficient level. The correlations were found to be high for the three exemplification sizes in the condition where the distribution was normal and the parameter was as expected. In conditions where the distribution is restricted and the parameter is unrestricted; and in the conditions where the distribution is normal and the parameter is restricted, the correlations were found to be high at a sufficient level. When the four condition on the restricted distribution and the parameters are evaluated together, it is possible to suggest that as the sample size increases, so will the correlations in a relative manner.

3.2. The bias, error and correlation values calculated for the model with two parameters (a and b parameters)

The bias, the Root Mean Square Deviation-RMSD and the averages of the correlation values calculated for a

and b parameters in the model with two parameters are given in Table 4.

Table 4. The bias error and correlation values calculated for the a and b parameters in the model with two parameters

PARAMETER A				PARAMETER B			
SAMPLES	BIAS	RMSD	R	SAMPLES	BIAS	RMSD	R
DRPR100	0.818	1.584	0.464	DSPS100	3.595	6.649	0.245
DRPR250	0.757	1.154	0.569	DSPS250	2.053	3.344	0.560
DRPR500	0.693	0.805	0.735	DSPS500	1.210	1.451	0.802
DRPE100	0.588	0.890	0.750	DSPB100	1.955	2.441	0.785
DRPE250	0.635	0.788	0.871	DSPB250	1.752	1.974	0.830
DRPE500	0.619	0.688	0.927	DSPB500	1.648	1.764	0.824
DNPR100	0.427	0.516	0.557	DNPS100	0.500	0.856	0.692
DNPR250	0.428	0.494	0.655	DNPS250	0.424	0.621	0.833
DNPR500	0.389	0.439	0.749	DNPS500	0.375	0.561	0.860
DNPE100	0.326	0.618	0.788	DNPB100	0.680	1.414	0.850
DNPE250	0.298	0.440	0.857	DNPB250	0.504	0.847	0.863
DNPE500	0.291	0.405	0.887	DNPB500	0.531	0.940	0.833

When the bias and RMSD values calculated for the a parameter in the model with two parameters given in Table 4 are considered, it is observed that the highest value is obtained in the condition where the distribution and the parameters are restricted; and the smallest values are obtained in the condition where the distribution and the parameters are unrestricted. It is also observed that in general, the sample size has an effect on the RMSD values calculated for the a parameter in the model with two parameters and on the bias; and it increases in conditions where the distribution is restricted.

When the bias and RMSD values calculated for the b parameter in the model with two parameters are examined, it is observed that the highest values are obtained in the condition where the distribution and the parameters are restricted; and the smallest values are obtained in the conditions where the distribution was normal and the parameter was restricted. It may be suggested that the bias averages and the RMSD values calculated for the b parameter in the model with two parameters are affected negatively by the restriction of distribution; and positively by restriction of the parameter. The smallest bias and RMSD values were obtained in conditions where the distribution was normal and the parameter was restricted. It may be suggested that the sample size and restriction of the distribution have a significant effect on the bias and RMSD values calculated for the b parameter in the model with two parameters; and that the effect of the sample size in other conditions is not significant.

When the correlations for the a value in the model with two parameters given in Table 4 are examined, it is observed that the lowest correlation was obtained for the condition where the distribution and the parameter were restricted, and the highest correlation was obtained in the condition where the distribution was restricted and the parameter was as expected. The a value was found to be low in each condition where the parameter was restricted, and was found high in conditions where the distribution was normal. It can be suggested that the correlation averages obtained for the a parameter in this model are affected at a significant level for each condition by the sample size. The smallest value of the correlation averages calculated for the b parameter in the model with two parameters was observed in the condition where the distribution and the parameter were restricted; and the highest value was observed in the condition where the distribution was normal and the parameters were unrestricted. The sample size has a significant role on the correlations calculated for the b parameter in the model with two parameters in the condition where the distribution was restricted. The sample size does not have a significant effect in the other conditions.

3.3. The bias, error and correlation values calculated for the model with three parameters (a, b and c parameters)

The average values of the bias, Root Mean Square Deviation-RMSD and correlation calculated for the a, b and c parameters in the model with three parameters are given in Table 5.

Table 5. The bias, error and correlation values calculated for the a, b and c parameters in the model with three parameters.

PARAMETER A	PARAMETER B	PARAMETER C
-------------	-------------	-------------

SAMPLES	BIAS	RMSD	R	SAMPLES	BIAS	RMSD	R	SAMPLES	BIAS	RMSD	R
DRPR100	1.174	1.934	0.183	DSPS100	5.225	9.259	0.356	DSPS100	0.036	0.076	0.060
DRPR250	1.082	1.619	0.250	DSPS250	3.635	6.295	0.604	DSPS250	0.020	0.067	0.283
DRPR500	1.115	1.625	0.301	DSPS500	4.133	7.108	0.721	DSPS500	0.027	0.067	0.222
DRPE100	0.334	0.815	0.712	DSPB100	1.828	2.940	0.575	DSPB100	0.077	0.094	0.364
DRPE250	0.348	0.622	0.814	DSPB250	1.704	2.279	0.706	DSPB250	0.074	0.098	0.540
DRPE500	0.338	0.606	0.740	DSPB500	1.866	2.584	0.807	DSPB500	0.098	0.119	0.330
DNPR100	0.533	1.050	0.455	DNPS100	0.487	1.404	0.799	DNPS100	0.015	0.053	0.613
DNPR250	0.471	0.732	0.629	DNPS250	0.100	0.291	0.949	DNPS250	0.004	0.038	0.839
DNPR500	0.342	0.611	0.691	DNPS500	0.057	0.211	0.963	DNPS500	0.005	0.033	0.879
DNPE100	0.148	0.400	0.745	DNPB100	0.353	0.733	0.813	DNPB100	0.105	0.124	-0.007
DNPE250	0.157	0.344	0.839	DNPB250	0.479	0.865	0.787	DNPB250	0.110	0.137	-0.089
DNPE500	0.118	0.257	0.907	DNPB500	0.275	0.486	0.923	DNPB500	0.075	0.100	0.094

When the bias values given in Table 5 are examined it becomes obvious that the highest bias and RMSD values calculated for the a parameter in the model with three parameters were obtained in the condition where the distribution and the parameters are restricted, and the lowest bias and RMSD values were obtained in the condition where the distribution and the parameters are unrestricted. The bias and RMSD values of the a parameter in the model with three parameters increase in the condition where the parameters are restricted, and decrease in the other conditions. No systematic effect of the sample size was observed on the bias values of the a parameter in the model with three parameters.

It was observed that the highest bias and RMSD values for the b parameters of the model with three parameters occurred in the condition where the distribution and the parameters were restricted; and the lowest bias and RMSD values were obtained in the condition where the distribution was normal and the parameters were restricted. It is observed that the restriction of the distribution in the model with three parameters increased the bias and RMSD values of the b parameters; and decreased in other conditions. No systematic and important effect of the sample size was observed on the bias and RMSD values calculated for the b parameter in the model with three parameters.

We can suggest that the smallest bias and RMSD values calculated for the c parameter in the model with three parameters were observed in the condition where the distribution was normal and the parameters were restricted; and the highest bias and RMSD value was observed in the condition where the distribution and the parameters were expected. We can also suggest that in case the parameters are restricted in the model with three parameters the bias and RMSD values for the c parameter decrease, and increased the bias values in other conditions; and affect the RMSD value at a less level. No systematic and important effect of the sample size was observed on the bias and RMSD values calculated for the c parameter in the model with three parameters.

When the correlations given in Table 5 are examined it is observed that the lowest correlation for the a parameter in the model with three parameters was calculated in the condition where the distribution and the parameters are restricted, and the highest correlation is calculated in the condition where the distribution and the parameters are unrestricted. The correlations decreased in the conditions where the parameters were restricted for the a parameter in the model with three parameters; and the correlations increased in other conditions. It is possible to suggest that the sample size has a partial effect on the correlations calculated for the a parameter in the model with three parameters in all conditions. The highest correlation for the b parameter in the model with three parameters was observed in the condition where the distribution and parameters were restricted, and the highest correlation was observed in the condition where the distribution was normal and the parameters were restricted. It can be suggested that the correlation obtained for the b parameter in the model with three parameters decrease in conditions where the distribution is restricted, and increase in other conditions. It may be suggested that the sample size has a significant effect on the correlations obtained for the b parameter in the model with three parameters, and that this effect even increases in the conditions where the distribution is restricted. The smallest correlation calculated for the c parameter in the model with three parameters was obtained in the condition where the distribution and the parameters were unrestricted, and the highest correlation was obtained in the condition where the distribution was normal and the parameters were restricted. No systematic effect of the sample size was observed on the correlations calculated for the c parameter in the model with three parameters.

4. Results and discussion

Within the scope of the study and based on the findings on the sub-problems, and in terms of IRT, distribution and/or parameter restrictions are applied for the parameter invariance studies. In this process, making the decisions according to the model which suits the data best may be a proper decision. Because restricting the distribution or the parameter according to the findings may have different effects on the parameters for each model. For example, restricting the distribution decreases the correlations of the b parameter, but increases the correlations of the a parameter. The basic reason of this result may be the differences in the ranges of the a and b parameters. Future studies in which this situation is controlled may produce beneficial results. On the other hand, restricting the parameters increases the correlations of the a and c parameters, and decreases in correlations of the b parameter.

When the findings are examined according to the bias values, it was observed that restricting the distribution increased the bias values of the a and b parameters, and restricting the parameters decreases the bias values of the a, b and c parameters. In order to test the accuracy of this situation, using the parameter restrictions that will be obtained from other distribution forms will be beneficial. In this context, it is clear that new studies in which parameter restrictions obtained from different distributions are needed.

When the findings are examined according to the RMSD values, it is observed that limiting the distribution increases the RMSD values of the a and b parameters, and decreases the RMSD values of the a, b and c parameters. In situations where there are unrestrictions for the distribution, the bias and RMSD values of the c parameter increase.

As the latest item, it must be stated that unrestricting the distribution and parameters increases the correlations of especially the a parameter, but decreases the bias and RMSD values. Unrestricting the distribution and the parameters may prevent the excessive increase of the bias and RMSD values calculated for the b parameters.

In the literature, there are findings that report the fact that the parameter invariance is affected by the size of the exemplification and the distribution type (Dogan & Tezbasaran, 2003), and findings that report that the IRT is insufficient in covering the condition of the parameter invariance of the IRT (Fan, 1998; Stage, 1998).

References

- Adedoyin, O.O., Nenty, H.J. & Chillisa, B. (2008). Investigating the invariance of the item difficulty parameter estimates based on CTT and IRT. *Educational Research and Review*. 3(2), 83-93
- Dogan, N. & Tezbasaran A.(2003). *Klasik test kurami ve ortuk ozellikler kuraminin orneklemler baglaminda karsilastirilmesi*. Hacettepe Universitesi egitim fakultesi dergisi. 25, 58-67.
- Fan, X. (1998). Item response theory and classical test theory: An empirical comparison of their Item-person statistics. *Educational and psychological measurement*, 58, 3, 357-381
- Galdin, M. & Laurencelle, L. (2010). Assessing parameter invariance in item response theory's logistic two item parameter model: A monte carlo investigation. *Tutorials In quantitative methods for psychology*. 6(2), 39-51
- Hambleton, R. K. & Swaminathan, H. (1985). *Item response theory: principles and application*. Boston: kluwer-nijhoff publishing.
- Hambleton, R. K., Swaminathan, H. & Rogers, J. H. (1991). *Fundamentals of item response theory*. Boston: sage publications.
- Han, K. T., & Hambleton, R. K. (2007). User's Manual: WinGen (*Center for Educational Assessment Report No. 642*). Amherst, MA: University of Massachusetts, School of Education.
- Nartgun, Z. (2002). Ayni tutumu olmeye yönelik likert tipi olcek ile metrik olcegin madde ve olcek ozelliklerinin klasik test kurami ve ortuk ozellikler kuramina gore incelenmesi. *Yayimlanmamis Doktora Tezi, Hacettepe Universitesi, Ankara*.
- Somer, O. (1998). Kislilik testlerinde klasik ve modern test kuramlari ile madde analizi. *Turk Ppsikoloji Dergisi*. 13, 41, 1-15
- Stage, C. (1998). Item analysis based on item response theory and on classical test theory: a comparison. *The international SweSAT conference*. San Diego.