LOOK ACADEMIC PUBLISHERS
OPEN ACCESS

INTERNATIONAL JOURNAL OF ENVIRONMENTAL & SCIENCE EDUCATION
2016, VOL. 11, NO.18 13093-13110,

Restricted With Parameters Which Are Estimated From Measurement Test Theory

[a] Ayfer SAYIN [b] Selahattin GELBAL

[a]Gazi University, TURKEY
[b]Hacettepe University, TURKEY

ABSTRACT

Even though the structural equation modeling has been used often by the researchers, it is observed that the assumptions are not examined before the analysis, the appropriate parameter estimation method was not performed. In cases when the estimations are not met in structural equation modeling, estimates below or above the parameter values are conducted. In this study, it is aimed at investigating the fit indexes of the model, estimated with a different parameter estimation method and sample size, based on item subtraction and restricted parameters (1, CTT and IRT values) in the confirmatory factor analysis model with multicollinearity problem. As a result of the estimation, the model was detected to make biased estimations when multicollinearity and sample size assumptions are not met. In order to prevent item loss, item parameters are suggested to be restricted with the values estimated from the classical test theory and item response theory. Based on the results of this research, it is asserted that all assumptions, as well as sample size and multicollinearity problem, are required to be examined before the confirmatory factor analysis is estimated. Otherwise there may be biased predictions. One of the multicollinearity problem-causing items may be subtracted or the items may be integrated. In order for the item not to be lost, the item parameters may be restricted. As information is obtained before the items, the values predicted with the classical test theory may be used in parameter restriction instead of 0 or 1.

KEYWORDS
multicollinearity,
confirmatory factor analysis,
item response theory
restricted parameter
classical test theory

ARTICLE HISTORY
Received 20 September 2016
Revised 13 November 2016
Accepted 29 November 2016

Introduction

Serving various purposes such as observing human behavior, clarifying the underlying causes of behaviors, and to reveal the causative relations, the concept of measurement was defined in many ways. Stevens (1946)

CORRESPONDENCE Ayfer Sayın ayfersayin@yahoo.com

defines measurement as the assignment of numerals to objects or events according to rules (cited in Crocker and Algina, 1986). Extending the definition of measurement, Stevens (1961) defined measurement as the observation of a quality, and its expression of the observation results with a number or symbol (cited in Baykul, 2010). Objects, behaviors or incidents may not be directly observed in the studies conducted in social sciences. Structural equation modeling is used in order to measure features that are not directly observed, namely the latent structures, test simultaneously the multiple regression equations, examine the causal processes and conduct the confirmatory factor analysis and path analysis processes.

Structural equation modeling is a multivariate statistic, and it bears some basic assumptions such as normality, extreme value, multicollinearity, and sample size (Byrne, 2010). If the variables observed in the structural equation modeling were not measured with correct measuring devices, or if the data set does not meet the assumptions of the model, all results will be biased estimate when the observed and latent variables in the model are to be explained (Quesnel, Scherling and Wallis, 2007). In this study, it is aimed at investigating how the multicollinearity and sample size assumptions reflect on the results.

Multicollinearity occurs in cases when there are variables which measure the same thing, but look like different variables. For instance, one of the variables with the correlation coefficient 0.95 can be analyzed, but two of them cannot be analyzed at the same time. Multicollinearity among the variables is demonstrated if the square of the multiple correlation coefficient calculated among all the variables is greater than 0.90, if the tolerance coefficient is smaller than 0.10, or if the variance inflation factor (VIF) is greater than 10. When there is multicollinearity among the variables, either one of the variables must be eliminated, or the two variables must be integrated (Kline, 2011). Tabachnick and Fidell (2007) indicate that a correlation coefficient 0.90 or more among the variables and the variables show the same tendency; moreover, the variables among which there is a correlation of 0.70 or more may cause multicollinearity problem. In case of multicollinearity, especially if the just identified or over-identified model turns into unidentified model, another method is to make restricted in the model parameters (Brown, 2006).

High correlations among the variables indicate the existence of a multicollinearity problem. Removal of one of the variables or integration of the variables in case of multicollinearity is based on the fact that highly correlated variables bear the same features. On the other hand, there may be variables in the model identification which do not have the exact same features even though they are highly correlated, and which the researcher consider necessary in the model, and it may be found appropriate for the variables to be in the theoretical structure of the model. However, it is observed in this case that the studies intended for the variables to stay in the model are restricted. Within the scope of this research, a model model specification which has multicollinearity items, alternatives were formed in order for the variables to stay in the model, and investigations were conducted (Raykoy and Marcoulides, 2000).

Widely used by the researchers nowadays, Classical Test Theory (CTT) is established on a linear-by-linear association between the observed score and true score (Crocker and Algina, 1986; Hambleton and Swaminathan, 1985; McDonald, 1999; Kline, 2005). Called the true score model and based on X=T+E basic equation, CTT has various assumptions. Some of these assumptions include the normal distribution of the random errors estimated with regards to the true score, the zero correlation between the random errors and the observed score or the true score, the equality of the standard deviation of the random error distribution and the standard error of the measurement, the equality of the observed score variance and the total of the true score variance and the error score variance, and the equality of the true score variance-observed score variance to reliability (Kline, 2005). In order for the analyses to be conducted in the CTT, the participants are required to respond to the items. Accordingly, the statistics of the test items (such as item difficulty) depend on the sample, and the calculated statistics are interpreted according to the group (Embretson and Reise, 2000). Item Response Theory (IRT) is a mathematical model defining the correlation between the possibility of an item to be correctly responded, and the ability level of the individual which the item aims at measuring. The established mathematical model provides information with regards to the possibility of responders at various ability levels to give response to an item (Crocker and Algina, 1986; Embretson and Reise, 2000; Baker, 2011). Hambleton and Swaminathan (1985) indicate that there are two main features on which the IRT is based. The test performance of a person can be estimated by way of a factor that can be expressed as trait or ability, and the correlation between the performance of a person about one item and the (performance-determining) trait enabling him/her to give response to that item can be explained with the item characteristic curve. Item characteristic curve is the graph of the possibility of the item's being correctly responded to, as the factor determining the test performance of an item or the function of the implicit feature. It provides information about a person's possibility, who is at a specific ability level, to correctly respond to an item. Theoretically, the ability of the individual is between the negative infinite and positive infinite. In terms of ease of application, the item characteristic curve is placed between -3 and +3 (Crocker and Algina, 1986; Baker, 2011). There are assumptions on which the IRT is based; such as unidimensionality, local independence, and the nature of the item characteristic curve (Lord and Novick, 1968; Hambleton and Swaminathan, 1985; Embretson and Reise, 2000; Baker, 2011).

In this study, it is aimed at investigating the fit indexes of the model, estimated with a different parameter estimation method and sample size, based on material subtraction and restricted parameters (1, CTT and IRT values) in the confirmatory factor analysis model with multicollinearity problem. Within the scope of the research, a total of 60 cases were examined, comprised of sample size (4) x parameter estimation method (3) x restricted parameters or integration/subtraction (5).

Method

This part is comprised of the data generation and data analysis.

Data generation

Model

Within the scope of the research, the unidimensional model comprised of six items, which are in the scale of mathematics learning-related opinions in the TIMSS 2011 study, was specificationed. The model is shown in the Figure 1.

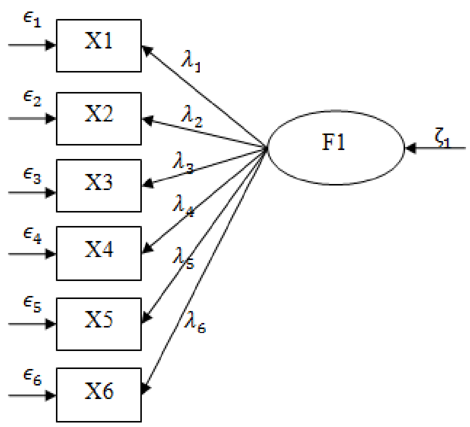

Figure 1. The model with multicollinearity problem that was examined within the scope of the research

In accordance with the responses of 6148 8th grade students who participated in the TIMSS 2011 study and answered all items on the scale, exploratory factor analysis was primarily calculated. As a result of the analysis, it was determined that the scale items, as shown in the Figure 1, became unidimensional, and they clarified 52.734% of the total variance. It was detected that the factor loadings of the items ranged between 0.843 and 0.465. The Cronbach's alpha reliability coefficient of the responses given to the scale items was estimated as 0.813. The items in the scale are shown in the Table 1.

Table 1. Scale items

Math	X1. Enjoy learning mathematics.
	X2. Wish have not to study mathematics.***
	X3. Math is boring.
	X4. Learn interesting things.
	X5. Like mathematics.
	X6. Important to do well in math.

***reverse items.

It was detected that there was a positive and high correlation among the responses given to the items X1 and X5 in the scale, and that this correlation caused a multicollinearity problem.

The degree of freedom was calculated for the single-factor model established with the abovementioned items. The model has five observed variables, and there are 6*(6+1) / 2 = 21 variables in the variance-covariance matrix of the model. As seen in the Figure 1, a total of 12 parameters will be estimated in Model 1 - six factor loadings and six error variances. As the degree of freedom of Model 1 is 21-12=9, the model is seen to be an "over-identified model".

Model Parameter

Data generation model parameters were determined in accordance with the responses given by 6148 8th grade students who participated in the TIMSS 2011 study in Turkey and who answered all the scale items. The correlation matrix and descriptive statistics referenced in data generation are shown in Table 2 and Table 3.

Table 2. Correlation Matrix Referenced for Data Generation

	X1	X2	X3	X4	X5	X6
X1	1					
X2	,432	1				
X3	,543	,580	1			
X4	,470	,274	,326	1		
X5	,867	,489	,623	,499	1	
X6	,333	,139	,188	,289	,334	1

Table 3. Average and Standard Deviation Value Referenced for Data Generation

Items	$\overline{X}$	SS
X1	1,77	0,946
X2	2,37	1,187
X3	2,25	1,123
X4	1,80	0,950
X5	1,93	1,043
X6	1,22	0,590

As a standard, the generated data was turned into raw scores, and they were rounded up the closest whole number possible.

Sample Size

In this research, investigations were conducted on the sample sizes of 100, 250, 500 and 1000.

Iteration Number

In the research, each datum was generated with 20 iterations. For 20-iteration purposes, 20*12=240 data were generated in the analyses.

Monte Carlo Study

As Monte Carlo simulation studies can be used for the purposes of creating independent data sets under relevant conditions, making test statistics or calculations for all data sets created, determining the true sample size, and summarizing the statistics estimated in the sample sizes (Fan et al, 2012; Davidian, 2005), data were created in this study by way of Monte Carlo simulation method. So it is the same equations of the correlation, co-variance, and the whole mathematical model all the data which include different sample size.

Data Analysis

Assumptions examination

Assumptions of four different data sets created for each model were examined. The fundamental assumptions of the structural equation modeling – missing data, outliers, non-normality of data distributions, linearity, homoscedasticity – were examined. As sample size is the variable of the research, sample size assumptions was not emphasized. Moreover, the multicollinearity problem of the model was taken into consideration.

Analyses

Following the data creation, the estimations were made by way of maximum likelihood (ML), unweighted least squares (ULS), generalized least squares (GLS) parameter estimation methods. The analyses are shown in Table 4.

Table 4. Depiction of Analysis Stages

Analysis Stages	Sample size			
	100	250	500	1000
RP*	X	X	X	X
IS**	X	X	X	X
II***	X	X	X	X
PEM****	X	X	X	X
FI******	X	X	X	X

* RP: Restricted Parameters (1, CTT- IRT)
** IS: Item subtraction
*** II: Item integration
**** PEM: Parameter Estimation Methods (ML- ULS-GLS)
*****FI: Fit Indexes (X^2/sd, RMSEA, GFI, CFI, SRMR, NFI)

In this study, the item parameters causing multicollinearity problem in the model were narrowed down. Based on the CTT measurement model, item-total correlation coefficients were used in the restrictions. aij parameter, defining item discrimination, was used in the IRT measurement model. aij, which is equal to the item characteristics curve slope, provides information about the quality of the item. The increase in the estimated aij parameter shows the increase in item discrimination (Embretson and Reise, 2000).

SAS 9.1.3 package program was used in data creation and verification of the creation. SAS 9.1.3, LISREL 8.7 program was used in the parameter estimation and estimations of the models, and the obtained results were compared. SPSS 21.0 and MULTILOG programs were used in the CTT-based parameter estimations and IRT-based parameter estimations, respectively, and the results were tabulated and reported.

Findings

1. What are the fit indexes of the model with a multicollinearity problem, estimated with different parameter estimation methods?

Within the scope of the research, the fit indexes of the model with a multicollinearity problem determined as a result of the estimations made by using different parameter estimation methods are shown in Table 5.

Table 5. Fit Indexes of the Model with a Multicollinearity Problem, Estimated with Different Parameter Estimation Methods

Sample Size	Parameter Estimation Methods	Fit Indexes						
		X^2 (sd)	X^2/sd	RMSEA	SRMR	GFI	CFI	NFI
100	ML	34,15 (9)	3,79	0,17	0,11	0,90	0,89	0,87
	ULS	16,6 (9)	1,84	0,13	0,08	0,98	0,97	0,94
	GLS	16,05 (9)	1,78	0,12	0,18	0,82	0,77	0,76
250	ML	50,94 (9)	5,66	0,16	0,092	0,92	0,93	0,92
	ULS	39,93 (9)	4,44	0,12	0,072	0,98	0,96	0,95
	GLS	34,50 (9)	3,83	0,11	0,11	0,87	0,84	0,81
500	ML	97,56 (9)	10,84	0,11	0,063	0,94	0,95	0,94
	ULS	49,33 (9)	5,48	0,09	0,07	0,99	0,96	0,95
	GLS	52,05 (9)	6,01	0,09	0,09	0,89	0,89	0,84
1000	ML	246,92 (9)	27,44	0,11	0,061	0,94	0,95	0,95
	ULS	154,83 (9)	17,2	0,098	0,058	0,99	0,98	0,97
	GLS	147,48 (9)	16,39	0,098	0,087	0,90	0,90	0,89

* Fit Indexes: X^2 (sd): Ki-kare (serbestlik derecesi), RMSEA: Root Mean Square Error of Approximation, SRMR: Standardized Root Mean Square Residual, GFI: Goodness of Fit Index, CFI: Comparative Fit Index, NFI: Normed Fit Index, NNFI: Non-Normed Fit Index

X^2/sd fit indexes of the model in which the multicollinearity hypothesis is not met, as seen in Table 5, increased in all parameter estimation methods (ML, ULS, GLS), depending on sample size. RMSEA and SRMR fit indexes were detected to decrease depending on the sample size. A general increase was detected in the GFI, CFI and NFI indexes as well, parallel to the increase in the sample size. Even though there is a multicollinearity problem in the model, the model-data concordance is obtained in 500 and 1000 samples, depending on the sample size.

2. What are the fit indexes estimated with different parameter estimation methods as a result of the integration/subtraction of multicollinearity problem-causing items?

Multicollinearity problem occurs in cases when the correlation among variables is high. This shows that the variables actually bear the same structures, and that they measure the structure to be measured at the same level and under the same condition (Brown, 2006). In this case, subtraction of one of the multicollinearity problem-causing items, or the elimination of items is suggested (Kline, 2005). Accordingly, the first and fifth items which cause the multicollinearity problem were respectively subtracted, the items were eliminated, their average was calculated, as well as the fit indexes. Estimation results are shown in Table 6.

Table 6. Fit Indexes Estimated with Different Parameter Estimation Methods as a Result of Item Integration/Subtraction

SS	PEM	Analysis Stages	Fit Indexes X^2 (sd)	X^2/sd	RMSEA	SRMR	GFI	CFI	NFI
100	ML	X1-X5 integration	21,37 (5)	4,27	0,22	0,085	0,92	0,88	0,85
		X1 subtraction	22,64 (5)	4,53	0,22	0,087	0,91	0,87	0,85
		X5 subtraction	21,41 (5)	4,28	0,18	0,085	0,92	0,86	0,83
	ULS	X1-X5 integration	11,36 (5)	2,27	0,15	0,085	0,99	0,95	0,92
		X1 subtraction	11,86 (5)	2,37	0,15	0,091	0,97	0,93	0,92
		X5 subtraction	12,00 (5)	2,40	0,12	0,091	0,97	0,94	0,90
	GLS	X1-X5 integration	12,21 (5)	2,44	0,16	0,13	0,87	0,74	0,74
		X1 subtraction	12,80 (5)	2,56	0,15	0,13	0,86	0,76	0,74

		X5 subtraction	12,04 (5)	2,41	0,120	0,12	0,87	0,80	0,74
250	ML	X1-X5 integration	49,87 (5)	9,97	0,18	0,078	0,91	0,88	0,87
		X1 subtraction	51,69 (5)	10,34	0,19	0,077	0,92	0,88	0,88
		X5 subtraction	47,08 (5)	9,42	0,18	0,080	0,92	0,87	0,86
	ULS	X1-X5 integration	29,90 (5)	5,98	0,14	0,080	0,99	0,93	0,93
		X1 subtraction	34,97 (5)	7,99	0,16	0,080	0,99	0,93	0,92
		X5 subtraction	29,36 (5)	5,87	0,14	0,085	0,98	0,94	0,92
	GLS	X1-X5 integration	32,07 (5)	6,41	0,15	0,10	0,88	0,76	0,73
		X1 subtraction	32,97 (5)	6,59	0,15	0,10	0,88	0,76	0,75
		X5 subtraction	27,66 (5)	5,53	0,13	0,10	0,89	0,78	0,74
500	ML	X1-X5 integration	83,72 (5)	16,74	0,19	0,071	0,93	0,91	0,90
		X1 subtraction	82,80 (5)	16,56	0,19	0,073	0,92	0,90	0,89
		X5 subtraction	83,05 (5)	16,61	0,18	0,067	0,93	0,91	0,90
	ULS	X1-X5 integration	39,39 (5)	7,88	0,12	0,072	1,00	0,94	0,94
		X1 subtraction	44,85 (5)	8,97	0,13	0,080	0,99	0,95	0,93
		X5 subtraction	45,72 (5)	9,14	0,13	0,080	0,99	0,94	0,93
	GLS	X1-X5 integration	46,83 (5)	9,37	0,13	0,090	0,89	0,81	0,74
		X1 subtraction	44,44 (5)	8,89	0,13	0,095	0,88	0,81	0,76
		X5 subtraction	44,97 (5)	8,99	0,13	0,084	0,91	0,80	0,76
1000	ML	X1-X5 integration	245,88 (5)	49,18	0,18	0,063	0,94	0,92	0,91
		X1 subtraction	237,98 (5)	47,60	0,18	0,062	0,94	0,92	0,92
		X5 subtraction	246,95 (5)	49,39	0,18	0,064	0,94	0,91	0,90
	ULS	X1-X5 integration	122,84 (5)	24,57	0,11	0,065	1,00	0,96	0,95
		X1 subtraction	121,36 (5)	24,27	0,12	0,065	0,99	0,96	0,95
		X5 subtraction	123,27 (5)	24,65	0,12	0,068	0,99	0,95	0,94
	GLS	X1-X5 integration	129,93 (5)	25,99	0,12	0,081	0,91	0,82	0,81
		X1 subtraction	120,23 (5)	24,05	0,13	0,081	0,91	0,84	0,83

	X5 subtraction	124,20 (5)	24,84	0,12	0,082	0,91	0,81	0,80

When the table is examined, an increase in the X^2/sd fit index depending on the sample size was observed. The increase observed before the estimation (in case when the multicollinearity assumption is violated) was also caused by the integration/subtraction of the multicollinearity problem-causing items. RMSEA and SRMR value was detected to decrease in line with the sample size, and the fit index was detected to be fixed after a sample size of 500. GFI, CFI and NFI fit indexes were detected to increase in line with the sample size, and to be generally fixed after a sample size of 500. In this study, GFI goodness of fit value was detected to have been predicted higher than the increasing CFI and NFI goodness of fit values.

The model-data fit estimated after processes such as parameter prediction methods and item integration/subtraction based on the sample size was detected to be lower than the model-data fit estimated in case of the violation of multicollinearity assumption (before the process). Especially in case the assumption is violated as the sample size decreases, it was detected that the model-data fit was inclined to be estimated higher, and created biased results. In other words, the multicollinearity assumption, especially in groups with small sample size, may result in model-data concordance. Within the scope of the research, it was determined that in case of multicollinearity, the items causing multicollinearity can be integrated or subtracted.

An addition during the process, it have been studied how the change of other variable parameters and it was determined to be minor differences. For example lambda coeffient of X^2 of the model in which the multicollinearity hypothesis is not met in 100 sample size, is estimated 0,57. It was estimated 0,63; 0,62 and 0,61 respectively when X1-X5 integrationed, X1 subtractioned and X5 subtractioned.

3. What are the fit indixes estimated with different parameter prediction methods as a result of the restriction of multicollinearity problem-causing items?

There is item loss in case of item integration/subtraction, and, especially in the adaptation studies, subtraction of items from the scale or the integration of item scores may cause problems. In order for the estimations to be made without item loss, the parameters of the first and fifth items were fixed to 1, CTT values and IRT values, respectively. Estimations were made with ML, ULS and GLS parameter prediction method at 100, 250, 500 and 1000 sample sizes. The estimation results are shown in Table 7.

Table 7. Fit Indexes Estimated with Different Parameter Prediction Methods as a Result of Item Restriction

SS	PEM	Analysis Stages (restriction)	Fit Indexes						
			X^2 (sd)	X^2/sd	RMSEA	SRMR	GFI	CFI	NFI
100	ML	λ_1=1,00; λ_5=1,00	34,38 (10)	3,44	0,17	0,100	0,90	0,89	0,87
		λ_1=0,795; λ_5=0,852	34,52 (10)	3,45	0,16	0,093	0,90	0,89	0,86
		λ_1=4,370; λ_5=7,155	69,98 (10)	7,00	0,30	0,330	0,81	0,77	0,75
	ULS	λ_1=1,00; λ_5=1,00	17,35 (10)	1,73	0,14	0,085	0,97	0,97	0,94
		λ_1=0,795; λ_5=0,852	17,51 (10)	1,75	0,13	0,081	0,98	0,95	0,94
		λ_1=4,370; λ_5=7,155	32,18 (10)	3,22	0,24	0,140	0,84	0,84	0,84
	GLS	λ_1=1,00; λ_5=1,00	15,62 (10)	1,56	0,14	0,25	0,82	0,65	0,62
		λ_1=0,795; λ_5=0,852	16,30 (10)	1,63	0,13	0,20	0,82	0,70	0,66
		λ_1=4,370; λ_5=7,155	37,00 (10)	3,70	0,32	0,84	0,56	0,47	0,45
250	ML	λ_1=1,00; λ_5=1,00	82,84 (10)	8,28	0,17	0,087	0,91	0,91	0,91
		λ_1=0,806; λ_5=0,840	75,89 (10)	7,59	0,16	0,09	0,91	0,92	0,91
		λ_1=3,550; λ_5=4,430	53,79 (10)	5,38	0,25	0,26	0,83	0,79	0,76
	ULS	λ_1=1,00; λ_5=1,00	53,20 (10)	5,32	0,13	0,008	0,98	0,94	0,94
		λ_1=0,806; λ_5=0,840	50,55 (10)	5,06	0,13	0,081	0,98	0,95	0,94
		λ_1=3,550; λ_5=4,430	42,88 (10)	4,29	0,17	0,12	0,96	0,92	0,90
	GLS	λ_1=1,00; λ_5=1,00	55,46 (10)	5,55	0,13	0,22	0,84	0,71	0,70
		λ_1=0,806; λ_5=0,840	49,26 (10)	4,93	0,12	0,19	0,86	0,75	0,74
		λ_1=3,550; λ_5=4,430	48,38 (10)	4,84	0,26	0,67	0,61	0,65	0,61
	ML	λ_1=1,00; λ_5=1,00	141,26 (10)	14,13	0,16	0,083	0,90	0,91	0,90
		λ_1=0,790; λ_5=0,851	113,84 (10)	11,38	0,16	0,076	0,92	0,92	0,92

500		λ_1=3,230; λ_5=7,180	451,94 (10)	45,19	0,24	0,22	0,93	0,80	0,80
	ULS	λ_1=1,00; λ_5=1,00	74,38 (10)	7,44	0,11	0,078	0,98	0,95	0,94
		λ_1=0,790; λ_5=0,851	60,52 (10)	6,05	0,10	0,074	0,98	0,97	0,95
		λ_1=3,230; λ_5=7,180	291,21 (10)	29,12	0,15	0,010	0,96	0,91	0,91
	GLS	λ_1=1,00; λ_5=1,00	81,84 (10)	8,18	0,12	0,17	0,84	0,76	0,74
		λ_1=0,790; λ_5=0,851	62,67 (10)	6,27	0,10	0,12	0,86	0,83	0,80
		λ_1=3,230; λ_5=7,180	508,60 (10)	50,86	0,20	0,55	0,66	0,79	0,74
1000	ML	λ_1=1,00; λ_5=1,00	312,06 (10)	31,21	0,16	0,076	0,91	0,92	0,92
		λ_1=0,804; λ_5=0,869	271,47 (10)	27,15	0,14	0,067	0,93	0,94	0,94
		λ_1=3,510; λ_5=10,300	593,46 (10)	59,35	0,15	0,065	0,94	0,95	0,94
	ULS	λ_1=1,00; λ_5=1,00	193,94 (10)	19,39	0,086	0,066	0,99	0,96	0,96
		λ_1=0,804; λ_5=0,869	176,17 (10)	17,62	0,087	0,061	0,99	0,97	0,97
		λ_1=3,510; λ_5=10,300	300,80 (10)	30,08	0,11	0,071	0,99	0,96	0,95
	GLS	λ_1=1,00; λ_5=1,00	189,03 (10)	18,90	0,075	0,17	0,85	0,91	0,82
		λ_1=0,804; λ_5=0,869	165,75 (10)	16,58	0,088	0,11	0,89	0,91	0,81
		λ_1=3,510; λ_5=10,300	673,39 (10)	67,34	0,12	0,093	0,89	0,84	0,82

When the information in the Table 7 is examined, the X^2/sd value was detected to increase, depending on the sample size, even after the restriction of item parameters. RMSEA and SRMR decreasing fit indexes, as expected, were detected to be predicted lower as the sample size decreased. GFI, CFI and NFI increasing fit indexes were detected to have been predicted higher with ULS parameter prediction method, when compared to ML and GLS. In addition, GFI fit index is observed to have higher values when compared to CFI and NFI fit indexes.

As a result of the fixation of the item parameters to 1 and the CTT values, higher goodness of fit values was predicted when compared to the fixation to IRT values. In the model formed by the fixation of model parameter values to 1, it was detected that the ML and item-data concordance was obtained in 250 samples.

If item loss is not desired, and if it is considered necessary for the item not to be subtracted in accordance with an expert opinion, it is observed that the parameters can be predicted with CTT and be restricted to these values. It was determined that the restrictions can be made with the values obtained from IRT in great sample sizes.

An addition during the process, it have been studied how the change of other variable parameters and it was determined to be major differences. For example lambda coeffient of X^2 of the model in which the multicollinearity hypothesis is not met in 100 sample size, is estimated 0,57. It was estimated 0,63; 0,62 and 3,11 respectively when restricted 1, CTT, IRT parameters.

Results

In this study, the effect of predictions made by using different parameter prediction methods and at different sample sizes on the model-fit indexes was examined, in case when the multicollinearity assumption is violated in confirmatory factor analysis. As a result of the examination, the model-data fit values estimated by violating the multicollinearity assumption were detected to be higher and biased when compared to the concordance values estimated based on item integration/subtraction. In the models with multicollinearity problem, it was determined that similar model-data fit values were obtained as a result of the subtraction of one of the multicollinearity problem-causing items and the integration of items. As the sample size increases (e.g. 500 and more), it was detected that the model-data fit values estimated as a result of item parameter restriction became closer to the values estimated with the integration or subtraction of items. There is not a certain acceptance with regards to sample size in the structural equation modeling analyses, and it is known that sample size has an effect on the parameter prediction methods and fit indexes. Determining the minimum sample size is considered an important problem in structural equation modeling (Jackson, Voth and Frey, 2013). Many researchers (Bentler, 1990; Fan, Thompson and Wang, 1999; Kim, 2009; Iacobucci, 2009; Kline, 2011; Jackson et al, 2013) studied on sample size, but a certain suggestion could not be made with respect to sample size. Schermelleh-Engel, Moosbrugger and Müller (2003) indicate that 400 and more observations are needed for predictions to be made with maximum likelihood at any situation. Anderson and Gerbing (1984) express that three and more indicators, and 100 observations are to be found for each factor, and that a sample size of 150 is adequate for the analyses (cited in Iacobucci, 2009, p.92). According to another view, the minimum sample size needed for the structural equation modeling can be unconditional 200 persons or can be conditionally determined based on the features of the model. In the first studies conducted for the minimum sample size condition to be determined, sample size (n) was expressed as n/q, depending on the parameter number to be estimated (q) (Jackson et al, 2013). There are researches accepting this rule at various levels (Bollen, 1989; Herzog and Boomsma, 2009; Kim, 2009 and Bentler, 2006; Kline, 2011; Marsh et al, 1988; Mueller, 1996; Nevitt and Hancock, 2004; Ullman, 1996); and there are ones that still do not reach a consensus on the n/q rule (Jackson, 2003, 2007; Marsh, Hau, Balla, and Grayson, 1998). Conditioned studies in the determination of sample size were then tried depending on values such as factor number and factor load value (cited in Jackson et al., 2013, p.87). They show that a sample size 500 is adequate for the confirmatory factor analysis in this study.

X^2/sd value was detected to monotonously increase in all prediction methods, depending on the sample size. In other researches, the chi-square fit index, which was detected to generate insensitive results in small samples as well, was detected to become insensitive as the sample size increases. Iacobucci (2009) indicates that the X^2 fit index needs great sample sizes; and similarly, Kenny and McCoach (2003) express that the statistical strength of X^2 fit index is low in small samples, and that differentiations cannot be made between the good and bad models of the fit index. In this case, it is suggested to report the X^2/sd value in both small and big sample groups in evaluating the model-data concordance, and to take the result into consideration. It was detected to be more coherent when compared to RMSEA, a decreasing concordance index, and sample size, a SRMR value. Kline (2011) indicates that the X^2 fit index is dramatically affected by the sample width. Diamantopoulos and Siguaw (2000) explain that especially the RMSEA decreasing fit index is one of the most informative values, and Byrne (1998) states that he makes estimations based on the universe covariance of the fit index. The fact that the decreasing goodness of fit values –RMSEA, which generally does not take a value smaller than 0.08 in any sample size, and SRMR, which generally does not take a value smaller than 0.05– decrease depending on the sample size, contrary to X^2 goodness of fit, shows that the model co-variations formed at different sample sizes are also similar.

As a result of the subtraction and integration of the items, and the fixation of the parameters to 1, CTT values, and IRT values, it was determined that higher estimations were made when compared to ML and GLS, by using the ULS parameter prediction method based on the asymptotic co-variance matrix at all sample sizes. Even though Schumacker and Lomax (2004) indicate that ML and GLS parameter methods create similar results under the multivariate normalcy assumption, the estimations show that GLS parameter prediction method makes lower fit index estimations when compared to ML, especially in small sample groups.

The fit value parameters predicted as a result of the fixation of item parameters to 1 in small sample groups were detected to be higher than the values estimated as a result of fixing them to CTT and IRT values. As the sample size increases, it was detected that the models with parameter values fixed to 1 had similar fit values with parameters with values fixed to CTT values; and that in cases when the sample size reaches 1000, the models with parameters fixed to IRT values create similar results with regards to model-data fit with the modes formed with restrictions. Item parameters are suggested to be fixed to 1 or another value (Schumacker and Lomax, 2004). In the study he conducted, Sörbom (1975) examined the goodness of fit values estimated as a result of the parameters fixed to values other than zero. The difference between X^2 values in the models in which the item parameters are free and restricted was examined, and it was determined whether there is a meaningful difference between the fit indexes. It was expressed that as the absolute value to be fixed increases, the fit value of the model will increase. Schumacker and Lomax (2004) indicate that predictable values, if any, can be used in fixing the item parameters. Stoel, Garre, Dolan and Wittenboer (2006) indicate that the unequal restrictions must be used in the structural equation modeling, that the items can be equalized to factor load values or higher values, and that this will increase the statistical strength of the model. Stoel et al (2006) state that adequate information is required to be needed for parameter restriction. Andrews (1999) reached the conclusion that the variances of some variables are equalized to zero as a result of restrictions with random numbers. Rindskopf (1983) suggests that parameters must be restricted with unequal values instead of equal or fixed values. In case of a multicollinearity problem in the model examined under different conditions within the scope of this research, as researchers (Kline, 2011, Brown, 2006) indicate, the problem-causing items may be subtracted or integrated. If item loss is not desired or it is considered necessary for the item not to be subtracted, it is understood that the parameters can be predicted with CTT and be restricted to these values. It was determined that the restrictions can be made with the values obtained from IRT in great sample sizes.

Based on the results of this research, it is asserted that all assumptions, as well as sample size and multicollinearity problem, are required to be examined before the confirmatory factor analysis is estimated. Otherwise there may be biased predictions. One of the multicollinearity problem-causing items may be subtracted or the items may be integrated. In order for the item not to be lost, the item parameters may be restricted. As information is obtained before the items, the values predicted with the classical test theory may be used in parameter restriction instead of 0 or 1.

Disclosure statement

No potential conflict of interest was reported by the authors.

Acknowledgement

This study is summarized from a thesis named "The Comparison Of Goodness Of Fit Indices In Structural Equation Model Restricted Wıth Parameters Which Are Estimated From Classical Test Theory And Item Response Theory..

Notes on contributors

Ayfer SAYIN holds Phd.Gazi University, Gazi Educational Faculty, Measurement and Evaluation in Education, Ankara, TURKEY.

Selahattin GELBAL is a professor at Hacettepe University, Educational Faculty, Measurement and Evaluation in Education, Ankara, TURKEY.

References

Andrews, D. W. K. (1999). Estimation when a parameter is on a boundary. *Econometrica, 67*(6), *1341-1383.*

Baker, F. B. (2001). The basics of item response theory. Eric Clearinghouse on Assessment and Evaluation.

Baykul, Y. (2010). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. PegemA Akademi, Ankara.

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), *238-246.*

Brown, T. A. (2006). Confirmatory factor analysis for applied research. In: Kenny, A. D. (Eds.) *Methodology in the social sciences*. Guilford Press, London.

Byrne, B. M. (1998). *A primer of LISREL basic applications and programming for confirmatory factor analytic models.* Springer-Verlag New York.

Byrne, B. M. (2010). *Structural equation modeling with AMOS basic concepts, applications, and programming*. Taylor & Francis Group, New York.

Crocker, L. M. & Algina, L. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston.

Diamantopoulos, A. & Siguaw, J. A. (2000). *Introducing LISREL*, SAGE Publications. California.

Embretson, S. E. & Reise, S. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum Publishers.

Fan, X., Thompson B. & Wang, L. (1999). Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling, 6*(1), 56-83.

Hambleton, R. K. & Swaminathan, H. (1985). *Item response theory principles and applications*. Boston: Kluwer.

Herzog, W., & Boomsma, A. (2009). Small-sample robust estimators of noncentrality-based and incremental model fit. *Structural Equation Modeling*, *16*, 1–27.

Iacobucci, D. (2009). Structural equations modeling: Fit Indices, sample size, and advanced topics, *Journal of Consumer Psychology*, *20, 90-98.*

Jackson, D.L., Voth, J. & Frey, M.P. (2013). A note on sample size and solution propriety for confirmatory factor analytic models. *Structural Equation Modeling*, *20*, 86-97.

Kenny, D. A., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling, 10*, 333–351.

Kim, K.H. (2009). The relation among fit indexes, power, and sample size in structural equation modeling. *Structural Equation Modeling*, *12*(3), 368-390.

Kline, R. X. (2005). Classical test theory assumptions, equations, limitations, and item analyses Loken (Chp. 5). *In Psychological testing: A practical approach to design and evaluation*, SAGE Publications, California.

Kline, R. B. (2011). Principals and practice of structural equation modeling. New York. The Guilford Press.

Lord, F. M. & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading MA: Addison-Welsley Publishing Company.

Marsh, H. W., Hau, K. T., & Grayson, D. (2005). Goodness of fit in structural equation models. In A. Maydeu-Olivares & J. J. McArdle (Eds.). *Contemporary psychometrics: A festschrift for Roderick P. McDonald, 275–340.* Mahwah, NJ: Erlbaum.

McDonald, R.P. (1999). Test theory: a unified treatment. Mahwah, NJ: Lawrence Erlbaum.

Quesnel, C., Scherling, C. & Wallis, N. (2007). *Structural equation modeling: A simple-complex multivariate technique*. SEMWHORKSHOP Presentation.

Raykoy, T. & Marcoulides, G.A. (2000). *A first course in structural equation modeling. Lawrence Erlbaum Associates*, Inc., New Jersey.

Rindskopf, D. (1983). A general framework for using latent class analysis to test hierarchical and nonhierarchical learning models. *Psychometrika*, *48*, 85-97.

Schermelleh-Engel, K., Moosbrugger, H. & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*(2), 23-74.

Schumacker, R.E. & Lomax, R.G. (2004). A *beginner's guide to structural equation modeling*. Lawrence Erlbaum Associates, New Jersey.

Sörbom, D. (1975). Detection of correlated errors in longitudinal data. *British Journal of Mathematical and Statistical Psychology, 27*, 229-239.

Stoel, R. D., Garre, F. G., Dolan C. & Wittenboer G. (2006). On the likelihood ratio test in structural equation modeling when, parameters are subject to boundary constraints. *Psychological Methods, 11*(4), 439-455.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics*. Boston: Allyn and Bacon