International Online Journal of Educational Sciences, 2019, 11 (4), 131-140

iojes
www.iojes.net

International Online Journal of Educational Sciences

ISSN: 1309-2707

Comparison of Different Estimation Methods Used in Confirmatory Factor Analyses in Non-Normal Data: A Monte Carlo Study

Research Article

Funda NALBANTOĞLU YILMAZ[1]

[1] *Nevşehir Hacı Bektaş Veli University, Faculty of Education, Educational Science Department, Nevşehir, Turkey*, ORCID: 0000-0002-3228-4605

To cite this article: Nalbantoğlu-Yılmaz, F. (2019). Comparision of Different Estimation Methods Used in Confirmatory Factor Analyses in Non-Normal Data: A Monte Carlo Study. *International Online Journal of Educational Sciences*, 11 (4), 131-140.

ARTICLE INFO

Article History:

Received: 27.03.2019

Available online: 06.08.2019

ABSTRACT

In social sciences, education and psychology, it is difficult to obtain distributional assumption. For these reasons, in situations where normality assumptions are not met, it is important to compare the effects of estimation methods used in CFA. So, in this study, it was aimed to compare CFA results obtained from ML, MLR and WLS estimation methods in data sets with different sample sizes where normality assumption is not met. The study was conducted with data produced after Monte Carlo simulations. In this direction, non-normal data sets were produced 300, 500, 1000, 1500 and 2000 sample sizes including 14 items in total; 7 items in the first factor and including 7 items in the second factor. In the study, fit indices were examined in order to find the efficiency of the same model which was tested with different estimation methods and changing sample sizes. Additionally, as a result of CFA conducted using different estimation methods, the average relative bias related to factor loads and the average relative bias related to standard error or factor loads were calculated. When analyses conducted with three estimation methods with the sample sizes are compared, it was found that the best fit is obtained with the MLR method in small sample size. In the comparison of WLS and MLR prediction methods in big sample sizes where distribution is not normal, WLS estimations were found to give less biased and truer factor estimations compared to MLR, especially ML.

Keywords:
Confirmatory factor analysis, estimation methods, non-normal data

Introduction

Confirmative Factor Analysis (CFA) is a statistical method which helps to investigate structure-relationship among latent-observable variables and to determine whether the structure is confirmed or not as a model (Jöreskog & Sörbom, 1993). In education and psychology, it is used widely in developing scales and investigating the structure among variables.

In confirmative factor analysis, it is aimed to test structures based on theoretical approaches (Şimşek, 2007). In examining the model, which was established based on theoretical approaches, with CFA, it is the first priority to state fit of model and parameter estimations (Lei, 2009). There are several estimation methods used in determining model fit and parameter estimates. Some of these necessitate normality assumptions and some others can be used in situations where normality assumptions are not met. For example; estimation methods

[1] Corresponding author's address: Nevşehir Hacı Bektaş Veli University, Educational Science Department, Nevşehir, Turkey.
Telephone: 03842281000-21034
e-mail: fundan@nevsehir.edu.tr
DOI: https://doi.org/10.15345/iojes.2019.04.010

like Maximum Likelihood (ML) and Generalized Least Square (GLS) necessitate normality assumptions (Schumacker & Lomax, 2016). According to some references (Baghdarnia, Soreh & Gorji, 2014; Çelik & Yılmaz, 2016; Schumacker & Lomax, 2016), Unweighted Least Squares (ULS) estimation method does not require distributional assumption, but in some other references (Kline, 2005, 2016) it requires a multivariate normal distribution.

In conditions where the normality assumption is met, the most frequently used method in parameter estimation is ML (Li, 2016). One reason for this is that ML estimation method provides unbiased and consistent parameter estimates (Bollen, 1989 cited in Curran, West & Finch, 1996). Another reason is that LISREL program, by which analyses are conducted, makes ML estimation when a method choice is not made. When the normality assumption is not met, different estimation methods can be chosen and used. Finney, Distefano and Kopp (2016) suggest using Weighted Least Squares (WLS), Robust Diagonal Weighted Least Squares (Robust DWLS), Robust Maximum Likelihood (MLR) and Satorra-Bentler (S-B) χ^2 when normality assumptions are not met and/or in categorical data.

In situations where the distribution is not normal, according to suggestions of using different estimation methods, comparing these estimation methods is one of the most important issues which should be investigated. Because in social sciences, education and psychology, it is difficult to obtain distributional assumption. Estimation methods influence fit indices and model parameters obtained from the model. For these reasons, in situations where normality assumptions are not met, it is important to compare the effects of estimation methods used in CFA.

In literature, there are studies comparing different estimation methods. Bandalos (2014) compared confirmative factor analysis results made by ML, MLR, and categorical DWLS estimation methods. Beauducel and Herzberg (2006) compared results obtained in confirmative factor analysis ML and mean and variance adjusted WLS methods. Curran, West and Finch (1996) examined sample size, multivariate normality effects by using ML, distribution-free method (ADF) and S-B χ^2. Çelik, Saraçlı and Yılmaz (2011) compared ML, ULS and GLS estimation methods under normality assumption. Fitriana (2011), in the CFA study made for Bandung Family Relations Test, compared ML, DWLS and WLS estimation methods in data where normality assumption is not met. Flora and Curran (2004), in CFA, made with data in which normality assumption is met and not met, compared WLS and DWLS (robust WLS) estimations. Kılıç, Uysal and Atar (2017) compared estimation methods which can be used based on Pearson and tetrachoric correlation matrix in confirmative factor analysis. Koğar and Yılmaz-Koğar (2015), in CFA study conducted with data at the level of classification and ranking, compared ML, ULS and DWLS methods. In some studies, in data where normality assumption is not met, results obtained from ML and DWLS estimation methods were compared (Baghdarnia et al., 2014; Lei, 2009; Mindrila, 2010; Wang & Cunningham, 2005). In some other studies, model fit indices obtained from ML, ULS, GLS estimation methods were compared (Sayın, 2016; Schumacker & Beyerlein, 2000).

In this study, differently from literature, it was aimed to compare CFA results obtained from ML, MLR and WLS estimation methods in data sets with different sample sizes where normality assumption is not met. In this direction, results of the study would help us predict about using which estimation method would give a better factor load in conditions where normality assumption is not met and which estimation method would be more appropriate. The study also extends applications to understand the differences and similarities between estimation methods. In accordance with this purpose, CFA was made with ML, MLR and WLS estimation methods on data sets with different sample sizes. Obtained results were discussed within the frame of research questions below:

1. How fit indices obtained from ML, MLR and WLS estimation methods change?
2. How average relative bias obtained from ML, MLR and WLS estimation methods for factor loadings change?
3. How average relative bias obtained from ML, MLR and WLS estimation methods for standard errors of factor loadings change?

Method

Type of the Study

This study aim to comparison estimates methods which used in CFA under non-normal data. In this respect, the study is a fundamental (basic) study.

Data Generation

The study was conducted with data produced after Monte Carlo simulations. The Monte Carlo studies can be used for methodological investigations of the performance of statistical estimators under various conditions (Muthén & Muthén, 2002). In line with the objective of study, Mplus 7.0 program was used and data sets were produced in different sample sizes which did not meet normal distribution assumption.

Data sets produced as a result of Monte Carlo study were created by utilizing the study of Muthén and Muthén (2002) and with reference to information about scales in literature and by choosing a replication number of 100. In this direction, non-normal data sets were produced 300, 500, 1000, 1500 and 2000 sample sizes including 14 items in total; 7 items in the first factor and including 7 items in the second factor. In each data set, factor loads were identified as 0.80, the residual variances of the factor indicators were identified as 0.55 -0.57 and factor correlation was identified as 0.40.

In the study, in five different sample sizes in which multivariate normality assumptions are not met, the effect of 3 different estimation methods was investigated. In this direction, results obtained from estimation methods in 15 (5*3) different conditions were compared.

In literature, there are different suggestions about sample size necessary in the structural equation model. Sample size can differentiate depending on the variable number in the model, the complexity of the model and estimation model used. In ML estimations which will be conducted with multivariate normal distribution data, it was stated that a sample size 10 times more than free parameter number is sufficient (Bentler & Chou, 1987). In conditions where there is no normal distribution, a larger sample size is necessary (Çelik & Yılmaz, 2016). Boomsma and Hoogland (2001), Schermelleh-Engel, Moosbrugger and Müller (2003) underlined that in the MLR method, the sample size must be at least 400. In asymptotically distribution free estimation methods (ADF), if the sample size is less than 500, standard error could be negative or/and biased (Finch, West & MacKinnon, 1997). In addition, it was suggested that in the WLS method, the sample size must be at least 1000 (Muthen, 1993 cited in Wang & Cunningham, 2005). Also, in literature, it is proposed that the ideal ratio of sample size and item number could be 20 (Jackson, 2003 cited in Kline, 2015). In this direction, for the model consisting of two-factor and 14 observed indicators, 5 different sample sizes (300, 500, 1000, 1500, and 2000) were determined.

Data Analysis

In the study, when normality assumption is not met, it was aimed to compare different estimation methods used in confirmative factor analysis. Accordingly, for data sets with different sample sizes used in the research study, univariate and multivariate normality assumptions were investigated.

For univariate normality assumption, firstly skewness and kurtosis coefficients and histograms were examined. And also z-statistics obtained by dividing the coefficient of skewness to standard error was examined. A great majority of z-statistics obtained for items in all data sets were found greater than +2.58 at 0.01 probability level. Accordingly, in data sets including sample sizes of 300, 500, 1000, 1500 and 2000, it can be said the data slightly non-normal. In determining whether data sets created to meet or do not meet multivariate normality assumption, Mardia's Test of Multivariate Normality obtained from LISREL program was used. According to multivariate normality test results, it was found that data sets with different sample sizes do not meet multivariate normality assumption (Mardia Kurtosis and Skewness Test $p<0.01$).

In literature, when normality assumption is not met, with reference to estimation method suggestions which could be used in CFA, estimations obtained from MLR and WLS methods were compared in the study (Baghdarnia et al., 2014; Browne, 1984; Finney & Distefano, 2013; Finney, Distefano & Kopp, 2016; Kline, 2011; Mindrila, 2010; Schermelleh-Engel et al., 2003). In addition, ML estimation was also used in the study. It is known that ML estimation is used in continuous data in which multivariate normality assumption is met (Boomsma & Hoogland, 2001; Kline, 2016). However, in some behavioral studies, it is not logical to expect data to show a normal distribution (Curran et al., 1996). On the other hand, in literature, it is stated that ML predictions are relatively resistant against violation of normality assumption (Çelik & Yılmaz, 2016; West, Finch & Curran, 1995 cited in Schermelleh-Engel et al., 2003). Furthermore, because the ML estimation method is frequently used in CFA studies in literature, ML estimation is also included in the study. In CFA that is conducted to compare estimation methods, the model given Figure 1 was used.

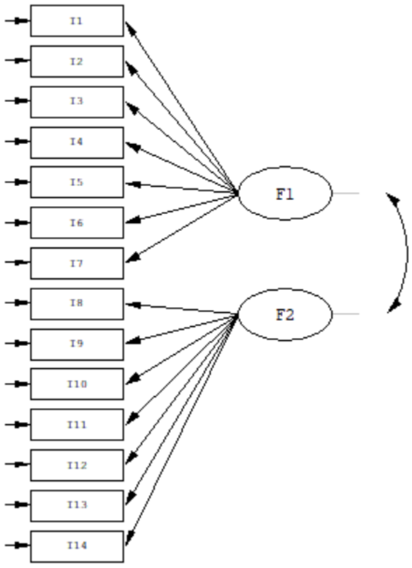

Figure 1. Estimation Model

In the study, fit indices were examined in order to find the efficiency of the same model which was tested with different estimation methods and changing sample sizes. In the study, indices obtained from different estimation methods like $\chi2/df$, RMSEA (Root Average Square Error of Approximation), SRMR (Standardized Root-Mean-Square-Error), AIC (Akaike Information Criterion) were examined. For these fit indices, the criteria below were used.

If $\chi2/df$ is lower from 3 this means excellent fit, if it is lower than 5 it means an acceptable fit (Sümer, 2000). RMSEA value below 0.08 means an acceptable fit and if it is below 0.05 it means excellent good fit (Browne & Cudeck, 1993 cited in Schermelleh-Engel et al., 2003; Schumacker & Lomax, 2016). SRMR value below 0.05 means a good fit and if it is below 0.10 it means an acceptable fit (Kline, 2011). AIC (Akaike Information Criteria) is used to compare models obtained from the same sample. As the AIC value decreases, the fit of the model increases (Kline, 2011).

Additionally, as a result of CFA conducted using different estimation methods, the average relative bias related to factor loadings and the average relative bias related to standard errors of the factor loadings were calculated. If the relative bias is below 5%, it is negligible, but if it is between 5-10% it has an average bias and if it is over 10%, it is interpreted as a high level of bias (Curran et al., 1996; Flora & Curran, 2004). In data analyses, LISREL 8.80 and Mplus 7.0 programs were used.

Findings

Comparison of Fit İndices

In Table 1, model fit indices obtained after CFA for the same model using ML, MLR and WLS estimation methods with different sample sizes are given.

Table 1. Model fit indices obtained from different estimation methods

Sample Size	Est. Methods	χ2/df	RMSEA	SRMR	AIC
300	ML	1.218	0.027	0.029	150.54
	MLR	1.216	0.027	0.029	150.43
	WLS	1.313	0.032	0.076	157.80
500	ML	1.269	0.023	0.022	154.44
	MLR	1.268	0.023	0.022	154.34
	WLS	1.364	0.027	0.050	161.68
1000	ML	1.455	0.021	0.021	168.55
	MLR	1.454	0.021	0.021	168.55
	WLS	1.362	0.019	0.031	161.49
1500	ML	1.203	0.012	0.015	149.44
	MLR	1.203	0.012	0.015	149.42
	WLS	1.171	0.011	0.019	146.96
2000	ML	1.158	0.009	0.013	146.00
	MLR	1.158	0.009	0.013	146.04
	WLS	1.141	0.0084	0.016	144.74

When χ2/df, which is one of the fit indices given in Table 1, is examined, all χ2/df values show a perfect fit. The smallest χ2/df value obtained in 300 and 500 sample sizes was through MLR, ML and WLS estimation methods respectively. The smallest χ2/df value in 1000, 1500 and 2000 sample sizes were obtained with WLS, MLR and ML estimation methods respectively.

RMSEA is important in terms of providing predictions independent from sample size (Şimşek, 2007). When RMSEA values obtained from estimation methods are examined, it is found that RMSEA values estimated in all methods show a perfect fit. In analyses conducted using ML and MLR methods in 300 and 500 sample sizes, better RMSEA values were obtained compared to WLS. On the other hand, with sample sizes of 1000, 1500 and 2000, as a result of CFA conducted using WLS prediction method, better RMSEA values were obtained.

SRMR values show a good fit in all estimation methods except the WLS method with a sample size of 300 (SRMR<0.05). AIC fit index is used for choosing the most appropriate model from a series of models for same data (Schermelleh-Engel et al., 2003), and choosing the closest model to the real with existing data (Cudeck & Browne, 1983 cited in Çelik & Yılmaz, 2016). AIC value is obtained when the best MLR estimation is used in 300 and 500 sample sizes. The same fit index with these sample sizes is obtained with ML and WLS respectively then. In 1000, 1500 and 2000 sample sizes, AIC value is obtained best with WLS prediction method. In this direction, it can be said that in big sample sizes (N=1000, 1500 and 2000), the closest estimations to reality is obtained in the model obtained with the WLS method.

Comparison of The Average Relative Bias According to Estimation Methods

Table 2 shows the average relative bias (ARB) for factor loadings by estimation method and sample size.

Table 2. ARB for factor loadings according to estimation methods

	Factor Loadings		
Sample Size	**ARB_{ML} (%)**	**ARB_{MLR} (%)**	**ARB_{WLS} (%)**
300	3.30	3.30	5.45
500	2.14	2.14	3.04
1000	1.43	1.43	1.16
1500	1.07	1.07	0.98
2000	0.80	0.80	0.63

As it is shown in Table 2, the average relative bias for factor loadings obtained from ML and MLR estimation methods change in between 0.80% and 3.30%. The average relative bias for factor loadings obtained from ML and MLR estimations are similar. The predicted bias values for these methods are equal because the factor loadings for ML and MLR were similar. In the literature, it is stated that similar factor load is obtained from these two estimation methods (Li, 2014).

The average relative bias for factor loadings obtained from MLR estimation range between 0.8% and 3.30%. Because relative bias obtained is less than 5%, the average relative bias for factor loadings obtained from MLR estimation is negligible. The average relative bias for factor loadings obtained from WLS estimation range in between 0.6% and 5.4%. The bias for factor loadings obtained with WLS estimation with all sample sizes except sample size of 300 is negligible. Factor loads obtained from WLS with a sample size of 300 shows moderately bias.

Table 3 shows the average relative bias (ARB) for standard errors of the factor loadings by estimation method and sample size.

Table 3. ARB for standard errors of the factor loadings according to estimation methods

	Standard Errors of Factor Loadings		
Sample Size	**ARB_{ML} (%)**	**ARB_{MLR} (%)**	**ARB_{WLS} (%)**
300	12.81	9.51	19.30
500	12.68	7.43	9.77
1000	13.36	6.77	7.25
1500	13.65	7.41	6.86
2000	15.31	4.82	4.72

Table 3 provides the average bias for the standard errors of the factor loadings. The average relative bias for standard errors of factor loadings obtained from ML change in between 12.81% and 15.31%. The ARB for standard errors obtained from MLR change between 4.82% and 9.51%. Although the ARB of factor loadings of two estimation methods are similar, ARB related to standard errors of factor loadings are different. ML estimation is show a high level of bias for the standard errors of the factor loadings under all sample size. But MLR estimation is show average bias under most sample size. And in 2000 sample size, ARB of standard error of factor loading is negligible for MLR estimator.

The ARB for standard errors obtained from WLS change between 4.72% and 19.30%. At a sample size of 300, WLS estimation is show a high level of ARB for the standard errors of factor loading. In 500, 1000 and 1500 sample sizes, WLS are showing an average bias. In 2000 sample size, the bias of WLS estimator is negligible.

ARB for standard errors of factor loadings in ML increased with increasing sample size. However, ARB for standard errors of factor loadings in MLR and WLS reduced with increasing sample size. When ARB for the standard errors of factor loadings of MLR and WLS estimators are compared, in small sample sizes, ARB in standard errors of factor loadings from MLR was smaller than WLS.

When three estimation methods are compared with the same sample size, in analyses with small sample sizes, the less bias is seen with MLR estimator. However, in big sample sizes, it can be said that WLS estimation contains less bias compared to MLR and ML. In small sample size, MLR estimation gives a truer factor load estimation compared to ML and WLS. In big sample size, WLS estimation gives truer factor load estimations compared to MLR and ML.

Discussion and Conclusion

In this study, it was aimed to compare CFA results obtained from ML, MLR and WLS estimation methods in data sets with different sample sizes where normality assumption is not met. In the direction of this purpose, first of all, fit indices obtained as a result of CFA found by using different estimation methods were compared. As a result of this comparison, it was found that fit indices, which were obtained after CFA found by using different estimation methods for the same model, show a good fit for each sample sizes. And in the study, obtaining small differences between fit indices according to the estimation methods may result from low/slightly non-normality.

When analyses conducted with three estimation methods with the sample sizes are compared, it was found that the best model-data fit is obtained with the MLR method in small sample size (N=300 and 500). In big sample sizes (N=1000, 1500 and 2000), it was found that the best model-data fit is obtained through the WLS estimation method. Nalbantoğlu-Yılmaz (2019), in her comparison study with real data, found that the model-data fit index for WLS estimations with a size of approximately 1000 samples was better.

In addition, in the study, the average relative bias for factor estimations obtained according to ML, MLR and WLS methods were compared. ML estimation gave biased standard errors of factor loadings estimations in all sample sizes in which normality assumptions are negligible. Studies in literature also support this condition (Bandalos, 2014; DiStefano, 2002; Mindrila, 2004; Curran et al., 1996; Finch, West & MacKinnon, 1997). It is known that deviation from normality can affect standard error, t coefficients and model data fit indices obtained through the ML method (Schermelleh-Engel et al., 2003). Results of this study supported these findings.

This study showed that out of three estimation methods, which are used when normality assumption is neglected, WLS and MLR methods give truer and more unbiased estimations compared to ML. Curran, West and Finch (1996) suggest using asymptotic distribution free (ADF) method or Satorra-Bentler (SB) chi-square in conditions when multivariate normal distribution cannot be met. In the study, WLS prediction is used out of ADF methods and within the scope of MLR, Satorra- Bentler (SB) chi-square was used. In this respect, these findings are consistent with the results of the study of Curran, West and Finch (1996).

In comparison of WLS and MLR prediction methods in big sample sizes where distribution is not normal, WLS estimations were found to give less biased and truer factor estimations compared to MLR estimations. However, in small sample size where distribution is not normal, it was found that factor loads estimated with MLR give truer and unbiased factor load estimations compared to WLS method. In literature, it was found that in small sample size (200-500) MLR estimations are better compared to WLS (Curran, West, & Finch, 1996). Chou and Bentler (1995) stated that WLS method is inappropriate in conditions where model is complicated and data is limited (cited in Schermelleh-Engel et al., 2003). Flora and Curran (2004) found in their study that WLS give incorrect results in small sample size. In addition, in literature, it was underlined that in data which does not meet normality assumption, WLS estimation method does not reveal correct results in small sample size and it necessitates bigger sample sizes (Bentler & Dudgeon, 1996; Hoogland & Boomsma, 1998; Muthén & Kaplan, 1992). Results obtained from this study are consistent with literature.

With reference to the results obtained from this study, in sample sizes of 300, 500, 1000, 1500 and 2000 where normality assumptions are not met, ML biased standard error for factor load estimations are conducted. In Curran et. all (1996) studies, when the increasing non-normality, ML showed increasingly bias of $\chi2$. In Bandalos' (2014) study, standard error bias for ML estimator showed increasing with the level of asymmetry.

Based on the study findings and literature, it is recommended that ML estimation should not be preferred in non-normal data, especially in moderately/severely/extremely non-normal distribution.

MLR gave better results than WLS in small sample size where normality assumptions are not met. In this direction, in small sample size where normality assumptions are not met, MLR estimation could be used. It is suggested to use WLS estimation in big sample size (1000, especially 1500 and 2000) which does not show normal distribution. In this study, effect of three different estimation methods in five different sample sizes, where normality assumption is not met, was examined. In this study, for unbiased parameter estimations, minimum sample size determination, which is necessary according to estimation methods, was not conducted. In another study, minimum sample size determination could be made for correct parameter estimations for estimation methods. In addition, in following studies, effect of different estimation methods could be searched in different model and sample size. In complicated models, CFA results obtained from three estimation methods can be compared. In literature, as the non-normality level increases, it is seen that the differences between methods increases. So, re-comparison can be made to other data that is severely/extremely non-normal distribution.

REFERENCES

Baghdarnia, M., Soreh, R. F., & Gorji, R. (2014). The comparison of two methods of maximum likelihood (ML) and diagonally weighted least squares (DWLS) in testing construct validity of achievement goals. *Journal of Educational and Management Studies*, 4(1), 22-38.

Bandalos, D. L. (2014). Relative performance of categorical diagonally weighted least squares and robust maximum likelihood estimation. *Structural Equation Modeling: A Multidisciplinary Journal,* 21, 102 – 116.

Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus average and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling*, 13, 186–203.

Bentler, P. M., & Chou, C.-P. (1987). Practical issues in structural equation modeling. *Sociological Methods and Research*, 16, 78–117.

Bentler, P. M., & Dudgeon, P. (1996). Covariance structure analysis: statistical practice, theory and directions. *Annual Review of Psychology*, *47*(1), 563-592. doi:10.1146/annurev.psych.47.1.563

Boomsma, A., & Hoogland, J. J. (2001). The robustness of LISREL modeling revisited. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural equation models: Present and future. A Festschrift in honor of Karl Jöreskog* (pp. 139–168). Chicago: Scientific Software.

Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37, 62-83.

Curran, P. J., West, S. G., & Finch, G. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*, 16-29. doi:10.1037/1082-989X.1.1.16

Çelik, H. E., Saraçlı, S., & Yılmaz, V. (2011). An application of structural equaiton modeling under the assumption of multivariate normality. *e-Journal of New World Sciences Academy Physical Sciences,* 6(4), 112-123.

Çelik, H. E., & Yılmaz, V. (2016). *LISREL 9.1 ile Yapısal Eşitlik Modellemesi.* Ankara: Anı Yayıncılık.

DiStefano, C. (2002). The impact of categorization with confirmatory factor analysis. *Structural Equation Modeling*, 9, 327–346.

Finch, J. F., West, S. G., & MacKinnon D. P. (1997) Effects of sample size and nonnormality on the estimation of mediated effects in latent variable models. *Structural Equation Modeling: A Multidisciplinary Journal,* 4:2, 87-107, DOI:10.1080/10705519709540063

Finney, S. J., & DiStefano, C. (2013). Non-normal and categorical data in structural equation modeling. G. R. Hancock ve R. O. Mueller (Ed.), *Structural Equation Modeling: A Second Course* (pp. 439–492). Greenwich, CT: Information Age Publishing.

Finney, S. J., Distefano, C., & Kopp, J. P. (2016). Overview of estimation methods and preconditions for their application with structural equation modeling. In K. Schweizer & C. Distefano (Eds.), *Principles and methods of test construction: Standards and recent advances* (pp. 135-165). Boston: Hogrefe.

Fitriana, E. (2011). Confirmatory factor analysis of the Bandung family relation test: A simulation study comparing ML, DWLS, and WLS estimation. Ph.D. thesis, Radboud University Nijmegen, Nijmegen. Retrieved from https://repository.ubn.ru.nl/handle/2066/87163

Flora, D.B., & Curran, P.J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological methods,* 9(4), 466-491.

Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: An overview and a meta-analysis. *Sociological Methods & Research*, *26*, 329–367.

Jöreskog, K., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language.* Chicago, IL: Scientific Software International Inc.

Kılıç, A. F., Uysal, İ., & Atar, B. (2017). *Doğrulayıcı faktör analizinde kullanılan kestirim yöntemlerinin karşılaştırmalı olarak incelenmesi.* IVth International Eurasian Educational Research Congress Conference Proceedings, Ankara: Anı Yayıncılık.

Kline, B.R. (2005, 2011, 2016). *Principles and practice of structural modeling.* New York-London: The Guilford Press.

Koğar, H., & Yılmaz-Koğar, E. (2015). Comparison of different estimation methods for categorical and ordinal data in confirmatory factor analysis. *Journal of Measurement and Evaluation in Education and Psychology,* 6(2), 351-364.

Lei, P. (2009). Evaluating estimation methods for ordinal data in structural equation modeling. *Quality & Quantity, 43,* 495 – 507.

Li, C. H. (2014). *The performance of MLR, USLMV, and WLSMV estimation in structural regression models with ordinal variables (doctoral disertation).* Michigan State University, USA.

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behav Res.* 48: 936–949. DOI 10.3758/s13428-015-0619-7

Mindrila, D. (2004). Maximum Likelihood (ML) and Diagonally Weighted Least Squares (DWLS) Estimation Procedures: A Comparison of Estimation Bias with Ordinal and Multivariate Non-Normal Data. *International Journal of Digital Society* (IJDS), 1(1): 1-7.

Muthén, B. O., & Kaplan, D. (1992). A comparison of some methodologies for the factor-analysis of non-normal Likert variables: A note on the size of the model. *British Journal of Mathematical and Statistical Psychology,* 45, 19–30.

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling: A Multidisciplinary Journal, 9*, 599-620.

Nalbantoğlu-Yılmaz, F. (2019). Gerçek bir veride doğrulayıcı faktör analizinde kullanılan alternatif kestirim yöntemlerinin karşılaştırılması. Vith International Eurasian Educational Research Congress Abstract Book, Ankara: Anı Yayıncılık.

Sayın, A. (2016). The Examination of model fit indexes with different estimation methods under different sample sizes in confirmatory factor analysis. *Journal of Measurement and Evaluation in Education and Psychology*, 7(2), 432-442.

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

Schumacker, R. E., & Beyerlein, S. T. (2000). Confirmatory factor analysis with different correlation types and estimation methods. *Structural Equatıon Modeling*, 7(4), 629–636.

Schumacker, R. E., & Lomax, R. G. (2016). *A beginner's quide to structural equaiton modeling.* New York: Routledge.

Sümer, N. (2000). Yapısal eşitlik modelleri: Temel kavramlar ve örnek uygulamalar. *Türk Psikoloji Yazıları*, 3(6), 49-74.

Şimşek, Ö.F. (2007). *Yapısal eşitlik modellemesine giriş: Temel ilkeler ve LISREL uygulamaları.* Ankara: Ekinoks Yayınları.

Mindrila, D. (2010). Maximum likelihood (ML) and diagonally weighted least squares (DWLS) estimation procedures: A comparison of estimation bias with ordinal and multivariate non-normal data. *International Journal of Digital Society (IJDS),* 1(1): 60-66.

Wang, W. C., & Cunningham, E. G. (2005). Comparison of alternative estimation methods in confirmatory factor analyses of the general health questionnaire. *Psychological Reports*, 97(1), 3-10.