# A Comparison of Frequentist and Bayesian Approaches: The Power to Detect Model Misspecifications in Confirmatory Factor Analytic Models

## Emine Önen

Division of Educational Measurement and Evaluation, Faculty of Education, Gazi University, Turkey

**Abstract**   This simulation study was conducted to compare the performances of Frequentist and Bayesian approaches in the context of power to detect model misspecification in terms of omitted cross-loading in CFA models with respect to the several variables (number of omitted cross-loading, magnitude of main loading, number of factors, number of indicators per factor and sample size) and (to) investigate the efficiency of BSEM approach to detect cross-loadings. BSEM approach allows including and estimating certain number of cross-loadings by specifying informative piror with small-variance for cross-loadings in the model. By this way, BSEM approach enables researchers to come up with models that better represent the substantive theory. At this simulation study, model misspecification was considered as major misspecification (cl=0.3) and minor misspecification (cl=0.1) according to the amount of omitted cross-loading. Results of this study revealed that Frequentist approach was so sensitive to minor model misspecification whereas Bayesian approach with non-informative prior was so sensitive to the major model misspecification. Finally, it was concluded that the power of BSEM approach to detect cross-loading varied according to the both amount and number of cross-loadings and for large amount of cross-loading the performance of this approach was so well.

**Keywords**   Model Misspecification, Confirmatory Factor Analysis, Frequentist Approach, Bayesian Approach, BSEM Approach

## 1. Introduction

It is assumed that educational and psychological measures reflect underlying and non-observable latent construct(s). The information about these latent constructs can be obtained through their effects on observed variables. Observed variables here are the measures or indicators of the relevant constructs [1,2,3]. Examining the factor structure of measures in terms of exploring and describing the connections between the educational and psychological measures and latent variables underlying these measures is quite important in making accurate and appropriate decisions related to the measured construct. The oldest and most common models known in specifying the relationships between observed variables and underlying latent constructs are factor analytic models [1]. As in cited Bollen [4], factor analytic models are built upon the model which was developed on the basis of Spearman's (1904) studies and then named "common factor model". There are two factor analysis techniques based on the common factor model: Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) [5]. EFA and CFA adopt different approaches and assumptions. Whereas EFA is a data-driven approach, CFA is a theory-grounded approach. The main aim in EFA is to specify the number of the underlying factors and nature of the variables in an observed variable set according to the relationships between these variables [1,6]. In EFA, it is allowed that all observed variables load into all the factors, and all factor loadings are freely estimated [1,4]. In CFA, unlike EFA, a priori factor structure is specified for the relationships between the latent factors and observed variables/measures, and the level of fit between this factor structure and sample data is examined. The basic equation of a CFA model is as follows:

$$y_i = \tau + \Lambda\eta_i + \varepsilon_i \qquad (1)$$

In this equation, $y_i$ (i denotes person) indicates the vector of observed indicators with $p \times 1$ (p: number of measured variables) dimension, and $\eta_i$ indicates the random vector of interrelated factor scores with q dimension. $\Lambda$ represents the factor loading matrix with $p \times q$ dimension, $\varepsilon_i$ represents a vector of residual scores, which are peculiar to each

observed variable with p dimension, and $\tau$ represents a $p \times 1$ vector including the constants related to observed variables. In this study, it was assumed that $\eta_i$ and $\varepsilon_i$ were independent of each other, and these parameters were normally distributed ($\eta_i \sim N_q(0, \Phi)$; $\varepsilon_i \sim N_p(0, \Psi)$) [7,8].

In order to specify a CFA model, it is required to make $q^2$ restrictions on $\Lambda$ and $\Phi$ (factor variances-covariance matrix) in the model. On the basis of theory and empirical findings, a specific number of components in $\Lambda$ matrix in equation (1) are fixed to 0. This is made to reflect a hypothesis that only specific factors affect the specific indicators [9,10,11]. Traditionally in CFA, a simple factor structure, in which each indicator is affected by only one factor, and an indicator's loading in another factor is restricted to 0 (zero), is specified. This represents a model which does not include any cross-loadings; however, in practice, psychometric indicators are rarely the perfect indicators of the pure structure. Indicators are mostly in a complex factor structure [11,12]. Cross-loading models the significant connections between indicators and non-target factors. Thus, allowing the estimation of cross-loadings in a CFA model will make explicit these connections. Relying on such a type of CFA model will be a more realistic approach than relying on CFA models in which these cross-loadings are not taken into account [13]. Reference [11] state that generally in CFA practices, more cross-loadings than what is required are fixed to 0 in order to specify the model. Besides that, the researchers emphasize that specifying cross-loadings incorrectly as 0 in CFA models causes overestimation of factor correlations, obtaining distorted factors, and then distorted structural relations.

Reference [14] expressed that a model is under-parametrized misspecified model if one or more parameters, the population value of which is non-zero, are specified as 0; and a model is over-parametrized misspecified model if one or more parameters, the population value of which is "0" (zero), are freely estimated. Accordingly, that a cross-loading, the real value of which is non-zero, is fixed to "0" indicates a model misspecification [14]. Especially in CFA models which contains weak indicators, even little misspecification of model has relatively a considerable impact on the other parts of the model. Model misspecification may lead to the convergence of the parameter estimates into incorrect values, and this case can make a serious threat to the validity of the measures [15,16,17]. The misspecification of CFA models increases the possibility that the researchers make Type I and Type II errors in the process of testing structural models, and these errors constrains the ability of theory development/improvement [18,19]. CFA is commonly used in order to provide evidence on the reliability and validity of educational and psychological measures. In this respect, the specification of a model, which is appropriate to the factor structure of educational and psychological measures, has importance in terms of the

validity and practicality of the relevant measures [20]. These measures provide a basis for making inferences and important decisions about individuals in the educational setting. However, there is always a risk of making inaccurate inferences about the reality when the models are misspecified [21]. Accordingly, the researchers carry out a model modification sequence as an option by using Modification Indices (MI) in order to detect misspecification in CFA models. And by this way they try to find a model which shows good fit to the data [22]. However, sometimes, theoretical, reasonable and sufficient basis for modifications carried out in this approach may not be found, and this case may cause the researchers to get away from the theory and lead them to an incorrect model [12,13,23].

These problems and strict assumptions in CFA conducted in Frequentist approach restrict the researchers' ability to examine cross-loadings and correlated residuals parameters [13]. Bayesian SEM (BSEM) approach suggested by Muthen and Asparouhov [11] on the basis of Bayesian approach solves these problems. The basic difference between Bayesian approach and Frequentist approach is related to the nature of the unknown population parameters in the model. In Frequentist approach it is assumed that there is only one constant and unknown parameter in the population. But in Bayesian approach, model parameters are regarded as random variables. Instead of trusting null hypothesis in Frequentist approach, the posterior distributions are evaluated in Bayesian approach. In Bayesian approach, posterior distributions are obtained for model parameters through updating the researcher's prior knowledge with the observed data. Besides that, it provides an opportunity to test the models, which will be under-identified in Frequentist approach, through the estimation of cross-loadings and residual covariances [24,25,26]. Also, ML-CFA estimation in Frequentist approach is based on the large sample normal theory; on the other hand, Bayesian approach is not based on this theory; that's why it is efficiently used in small samples [11,25]. Finally, while model modifications are made by using MI in Frequentist approach. But there is no need to make sequential modifications in Bayesian approach since model parameters can be freely and concurrently estimated in this approach [11,13].

## 1.1. Bayesian Approach

The aim of Bayesian approach is to combine data likelihood and prior knowledge, which is obtained from theory, previous studies and experiences, for getting posterior distributions for the unknown model parameters. In this approach, it is considered that model parameters include an uncertainty, and each parameter has a distribution to capture the uncertainty related to this parameter value. Bayes Theorem is explained with the following equation by means of the probability

distributions terms:

$$f(\theta|data) = \frac{f(data|\theta).\ f(\theta)}{f(data)} \qquad (2)$$

In this equation, $f(\theta|data)$ indicates posterior distribution for population parameter ($\theta$), $f(data|\theta)$ represents sampling density, and $f(data)$ denotes marginal distribution [24,27]. Prior distribution ($f(\theta)$), is a key element in Bayesian analysis, and it is based upon prior beliefs about the possible values of a parameter. Priors may be informative or non-informative. Non-informative prior expresses a prior distribution with a high variance, which includes a large amount of ambiguity about the population parameter. In another words, non-informative prior contains a little of information (or no information). On the other hand, informative prior is a prior distribution with low variance which includes a large amount of information and low level of ambiguity about the population parameter [11,24]. In Bayesian approach, Markov Chain Monte Carlo (MCMC) algorithms are generally used to obtain posterior distributions based on priors and data likelihood. Posterior estimations for the parameters are provided as the mean, mode and median values of these distributions [28,29]. In MCMC sampling, convergence should be achieved to end iteration and get posterior distributions [11]. In order to assess whether the convergence is achieved or not, Potential scale reduction (PSR) is calculated by using Gelman-Rubin convergence diagnostic. PSR values lower than 1.1 show that convergence is achieved [11, 29]. Assessing model fit in Bayesian analysis is carried out with posterior predictive checking. The main logic here is that the difference between the data generated by the model and the real data should be very small. Any possible deviation between two data sets will point out the model misspecification. Accordingly, posterior predictive p-values (ppp-values), which show the degree of model fit, are calculated. Although there are no theoretical cutoffs, it is seen that different cut of values as 0.01, 0.05 and 0.10 are suggested in the literature [28,29]. In this study, 0.01 cut of value was considered. An approach proposed within the context of Bayesian approach and SEM framework [11] provides the ways of examining the possible sources of the differences between the specified model and data. But in Frequentist approach instead of this examination, it is given suggestions on model modifications to increase the model fit by carrying out only model test. Since this study is conducted on the basis of BSEM, BSEM approach is explained in the next section.

## 1.2. Bayesian Structural Equation Modelling (BSEM)

The existence of the significant cross-loadings and correlated residuals, which are included in population model in fact but may not be captured with a CFA conducted within Frequentist framework. However, it can be explored by means of Bayesian Structural Equation Modeling (BSEM). By this way, more appropriate and accurate inferences about the relevant structures could be

made by specifying and testing the models, which reflect the theory better. BSEM is quite useful in terms of providing opportunities to test a model, which will be under-identified when a parameter is added to a CFA model. In this approach, the models, which include more parameters, could be estimated effectively in even small samples. BSEM approach in this context is more flexible; for that reason, it has been more commonly used in Social and Behavioral Sciences. In BSEM, models are not tested by using the parameter specifications with "exact zero" or "exact equation" as it is done in Frequentist approach. Instead of this, in BSEM approach estimations are obtained by using "approximate 0" and "approximate equations" in accordance with the informative priors with low variance. As in EFA, this approach is highly important in terms of both allowing the estimation of cross-loadings and providing model specification which reflects the theory better since it is still based on theory [11,13]. BSEM approach can be used for both structural models and measurement models since it can be applied to any constrained parameter in SEM. This study focused on the estimation of CFA model within the frame of BSEM because CFA models were studied in this research. When $\theta$ is considered as a collection of the parameters ($\tau$, $\Lambda$, $\Phi$ and $\Psi$), which will be estimated, for a CFA model within the frame of BSEM, the posterior distribution for $\theta$ is obtained through the following equation:

$$p(\theta|Y) = p(Y|\theta)p(\theta)/p(Y) \propto p(Y|\theta)p(\theta) \qquad (3)$$

In equation (3), the term of $p(\theta|Y)$ shows the posterior distribution for $\theta$, and the term of $p(Y|\theta)$ indicates data likelihood and $p(\theta)$ prior distribution. Various prior distributions can be specified here for the parameters included in $\theta$. Generally, Inverse-Gamma (in case of univariate distribution) and Inverse-Wishart (in case of multivariate distribution) distributions are specified for variance and covariance parameters. In addition, normal distribution is specified for mean, intercepts and slopes parameters as conjugate priors:

$$\Phi \sim \text{Inv-Wishart}\ (\Phi_0, d),$$
$$\psi_{jj} \sim \text{Inv-Gamma}\ (\upsilon_\psi/2,\ \upsilon_\psi\psi_0/2)\ , \qquad (4)$$
$$\tau_j \sim N\ (\mu_\tau, \sigma_\tau^2)$$

Here $\Phi_0$ and $d$ (hyperparameters) are the parameters of Inverse-Wishart distribution, and $\Phi_0$ reflects the researcher's prior beliefs about the elements of $\Phi$. On the other hand, $d$ ($d \geq p+4$, p: number of observed) is manipulated to set the informativeness of the prior [11, 20, 30]. In this study, as in Lu, Chow and Loken's [20] study, it was assumed that factor covariance matrix is an ($\Phi$) positive definite matrix, and residual covariance matrix ($\Psi$) is a diagonal matrix (off-diagonal elements are fixed to 0). Accordingly, Inverse-Gamma was specified as prior distribution for only residual variances ($\psi_{jj}$). Within the frame of BSEM, the elements ($\lambda_{jk}$) in $\Lambda$ can be considered as two group in estimating m-factor CFA models. Main

loadings, which are specified according to substantive theory and freely estimated in Frequentist approach, are included in the first group. A prior distribution with $\sigma_\lambda^2$ variance is specified for those in the way allowing main loading values to differ considerably from 0.

$$\lambda_{jm} \sim N\,(\mu_\lambda,\, \sigma_\lambda^2) \qquad (5)$$

The second-group elements in loading matrix ($\Lambda$) are cross-loadings which specify the relationships between an indicator and other non-target factors. Cross-loadings are fixed to 0 and not estimated in the context of model specification in testing CFA models within Frequentist framework whereas BSEM approach allows their estimation [11,30]. Reference [11] specified three main fields of study for CFA models within the context of BSEM approach, and one of these fields is cross-loadings in CFA. Other two fields are (a) residual correlations in CFA and (b) examination of measurement invariance with MIMIC modelling. As this study focused on cross-loadings in CFA models, only this field of study was taken into account in this research.

### 1.3. Cross-loadings in BSEM

Unlike the structure of main loadings based on the requirement of "0" cross-loadings in Frequentist approach, flexible cross-loading structure is allowed in the estimation of CFA models in BSEM approach as in EFA models. In this context, not only does BSEM provide the estimation of non-zero cross-loadings in reality, but also allows other cross-loadings to approximate "0". In BSEM, a specific number of cross-loadings with moderate magnitude can be added to the model. The key point here is to choose the variance of the prior distribution to be specified for cross-loadings. The specification of a prior with very low variance to cross-loadings may cause cross-loadings to not adequately differ from "0" which is the prior mean. This case may result in getting worse of ppp-values in terms of the model fit. On the other hand, high value of a prior variance may cause cross-loadings to have considerably high values. This also may cause a model to approximate a non-identified model. In this case, MCMC algorithm may not be convergence for the relevant parameter estimations. With a prior with "0 mean" and "low variance" to be specified for cross-loadings, a researcher reflects a prior belief that the probability of cross-loadings around "0" is higher [8,11,13]. Reference [11] elucidated the values of variance which can be chosen for "a normal prior with 0 mean". Accordingly, 0.01 prior variance will result in cross-loadings located between -0.2 and +0.2 with 95%. This is a highly informative prior and points out that cross-loadings are close to zero but not exactly zero. In BSEM approach, for a CFA model in which all parameters are freely estimated, it could be gathered information about model modification by examining 95% Bayesian Credibility Interval (95% BCI) for cross-loadings. When 95% BCI related to a $\lambda$ parameter which is cross-loadings

does not include 0, it is considered that this cross-loading points out a significant relationship between the relevant indicator and factor. Therefore this cross-loading should be specified and estimated in the model. That 95% BCI includes 0 is regarded as MI, and this case shows that the relevant cross-loadings is non-significant, and it may not be taken into consideration. In this regard, it was stated that BSEM performs better than ML-CFA in detecting model misspecification in terms of cross-loadings [8,11,13].

It is seen that there are studies carried out through Monte Carlo simulation and studies based on the real data in relation to cross-loadings in factor analysis within the frame of BSEM approach [8, 11, 20, 22, 25, 26, 31]. In these studies, generally, the methods based on BSEM approach and methods based on Frequentist approach were compared. The comparisions were made from the perspectives of exploring the existence of cross-loadings and correlated residuals and evaluating model modifications in this context [11,25,26], variable selection methods [8] and various model selection criteria [20], the impact of various prior distributions for cross-loadings (BSEM with Spake and Slob Prior: BSEM-SSP, BSEM with Ridge Proir: BSEM-RP) on factor solutions [8,11]. In these studies, the number and magnitude of cross-loadings [8,11,20], sample size [8, 11, 22], violations of assumptions on the distributions [20], magnitudes of factor loadings, factor correlations and residual correlations (Pan, Ip, & Dube, 2017) were manipulated. The performance of two approaches was examined under those conditions. The research results [8, 11, 20, 22, 25, 26] generally point out that when informative priors for factor analytic models are used, BSEM approach-based methods outperform Frequentist approach-based methods in terms of model fit, detection of cross-loadings, rejection of incorrect models. Simulation studies reveal that the power to detect cross-loadings in BSEM changes according to the magnitude of cross-loadings and sample size when informative priors are used [8,11,20]. In addition it has been found that the effects of different prior distributions (BSEM-SSP and BSEM-RP) according to the magnitude of cross-loadings and sample size differ [8] in terms of RMSEA and cross-loadings within the context of true-positive rates and false positive rates. Accordingly, the current study comparatively examines the performance of BSEM approach and performance of Frequentist approach in detecting model misspecification in view of cross-loadings in CFA. For that reason, next section is about the detection of model misspecification.

### 1.4. Detection of Model Misspecification

This study mainly examines the performance of Bayesian approach and performance of Frequentist approach comparatively in detection of "model misspecification" in CFA models in terms of cross-loadings. In the literature [14, 15, 16, 17, 32] it is

seen that "model misspecification" in CFA within the Frequentist approach is studied within the contexts of the association of an indicator with an incorrect factor, number and magnitude of cross-loadings, misspecification of factor structure. In a study conducted by French and Finch [15], it was observed that Type I error rate, fit indexes (CFI, RMSEA) and power to detect misspecification in terms of the association of the indicators with incorrect factors change according to the some variables: Sample size, number of factors, and number of indicators; and generally for more complex models, this error is lower and, the power is higher. Simulation studies on the detection of model misspecification in terms of cross-loadings point out that the power to detect misspecification differs in accordance with the number of cross-loadings and sample size [32]. Also the decisions on accepting or rejecting misspecified models may change as a function of sample size [14]. In a study [16] it was found that RMSEA is mostly insensitive to multiple omitted cross-loadings. However, RMSEA sensitivity to misspecification increases when factor loadings increase; on the other hand, this sensitivity decreases when model size increases [16]. It is seen that in the literature [11,22,33] related to BSEM approach, Frequentist approach and BSEM approach were compared for CFA and EFA models under the various conditions. These conditions were changed according to the number and magnitude of the cross loadings, number of the factors, sample size, violence of the assumptions on the distribution, magnitude of main-loadings factor correlations and residuals by examining "model misspecification" with regard to "cross-loadings" and "correlated residuals" [11,22,33]. The simulation study carried out by Muthen and Asparauhov [11] showed that ppp-values within the condition of low level misspecification in large samples were more robustness by comparison with p values in Frequentist CFA. The researchers stated that 95% coverage changes accordance with the magnitude of the cross-loadings, and the impact of cross-loading magnitude on ppp-values differs according to the sample size. In parallel with this finding, it was found in the study carried out by Pan, Ip and Dube [22] that the power to detect model misspecification and Type I error rates change in accordance with the magnitude of factor loadings and factor correlations. The results of a study [20] in which the methods of selecting variables were compared, point out that Bayesian model selection criteria (Bayesian Information Criteria: BIC, Bayes Factor: BF) equilibrates well, by comparison with EFA, between TP-Rates and FP-Rates. Also it was observed that the differences between Bayesian MCC and LRT estimations depend on the sample size and magnitude of cross-loadings. In another study [33] it was found that the rejection rates for the models with minor misspecification increased in accordance with the levels of misspecification and sample size in both frequentist CFA and Bayesian CFA.

When the results in all these studies [11, 20, 22, 33] are holistically evaluated, it can be understood that the power

to detect model misspecification differs according to the number and magnitude of the cross-loadings, number of factors and number of indicators for per factor, magnitude of the factor loadings, level of misspecification and sample size. However, in the result of reviewing the related literature, when model misspecification is taken into account in terms of the cross-loadings, any studies that compare CFA conducted in the Frequentist framework and CFA conducted in the Bayesian framework under the conditions of sample size, number of factors, number of indicators per factor and magnitude of factor loadings could not be reached. On the other hand, simulation studies [14,16,32], point out that misspecified models can produce model fit indexes in a way that they will indicate the models which show good fit. In addition, it is stated that model size is an effective factor in detecting model misspecification. These cases may lead to the rejection of a model which includes misspecification at a negligible level or acceptance of a model, which includes important misspecification level, as a valid model. However, finding a correct model related to the factor structure of psychological and educational measures has a considerable importance in making appropriate decisions especially in educational settings and getting accurate results in the research studies. For these reasons, it is considered that it is important to compare the performance of these approaches in detecting "model misspecification" and specify a more accurate way in view of what approach is adopted, and under what condition this approach is adopted. Accordingly, a simulation study, in which misspecification was regarded as the magnitude of "omitted cross-loadings", and the level of misspecification was considered as "minor misspecification" and "major misspecification", [111,32,33] was carried out. In this simulation study, the power to detect model misspecification in CFA models of Frequentist and Bayesian approaches and BSEM (in terms of rejection of incorrect models) was compared under various conditions specified for the number of factors, number of indicators for each factor, sample size. In addition, the power of BSEM approach to detect cross-loadings was investigated.

## 2. Methods

### 2.1. Simulation Design and Data Generation

In this study, various simulation settings were designed in order to compare the classic ML CFA and CFA within Bayesian framework in terms of the power to detect model misspecification. Accordingly, power of these two approaches to detect model misspecification was examined under the conditions of the number of different factors, number of different indicators for per factor, different factors' loadings, magnitude of different cross-loadings (indicates misspecification level) and number of different

cross-loadings, and size of different samples. In this study, on the basis of the literature review [11,33], for the level of misspecification, the conditions, in which cross-loadings are specified as 0.1, are considered as "minor misspecification", and the conditions, in which cross-loadings are specified as 0.3, were considered as "major misspecification". For each level of model misspecification, in terms of cross-loadings two conditions (single cross-loading: one omitted cross-loading and multiple cross-loadings: 3 omitted cross-loadings), two factor numbers (3 and 4), two conditions in terms of the number of indicators for per factor (3 and 4 indicators), two sample size (200 and 500), and magnitude of two factor loadings ($\lambda=0.3$ and $\lambda=0.8$) were considered, which results in 2 (cross-loading number) $\times$ 2 (sample size) $\times$ 2 (factor number) $\times$ 2 (number of indicator for per factor) $\times$ 2 (factor loading magnitude)=32 conditions (in total 64 conditions). Accordingly, under the conditions of model misspecification, different factor structures were specified in accordance with the factor number, number of indicator for per factor, and number of cross-loadings.

In Figure 1, those denoted by "X" indicate main loadings, and those denoted by "x" indicate cross loadings. For these factor structures, data were generated for each condition (with continuous indicators) by manipulating the magnitudes of main loadings and cross-loadings. In this simulation study, both for data generation and model testing processes, factor variances were fixed at 1.00 for identification purposes, and for simplicity, the indicators' intercepts were fixed at 0. All factor correlations and all residual variances for indicators were 0.3. First of all in this process, data sets were generated on the basis of the factor structures in which cross-loadings were not found for the relevant sample size under the conditions (3 factors-9 indicators, 3 factor 12 indicators, 4 factors with 12 indicators) specified on the basis of the number of factor and number of indicator. Afterwards, 500 iterations were used for each condition in accordance with the factor structures specified in the study, level of misspecification, and sample size. Conditions, which were studied separately both for minor and major misspecification levels in this research, are presented in Table 1 below.
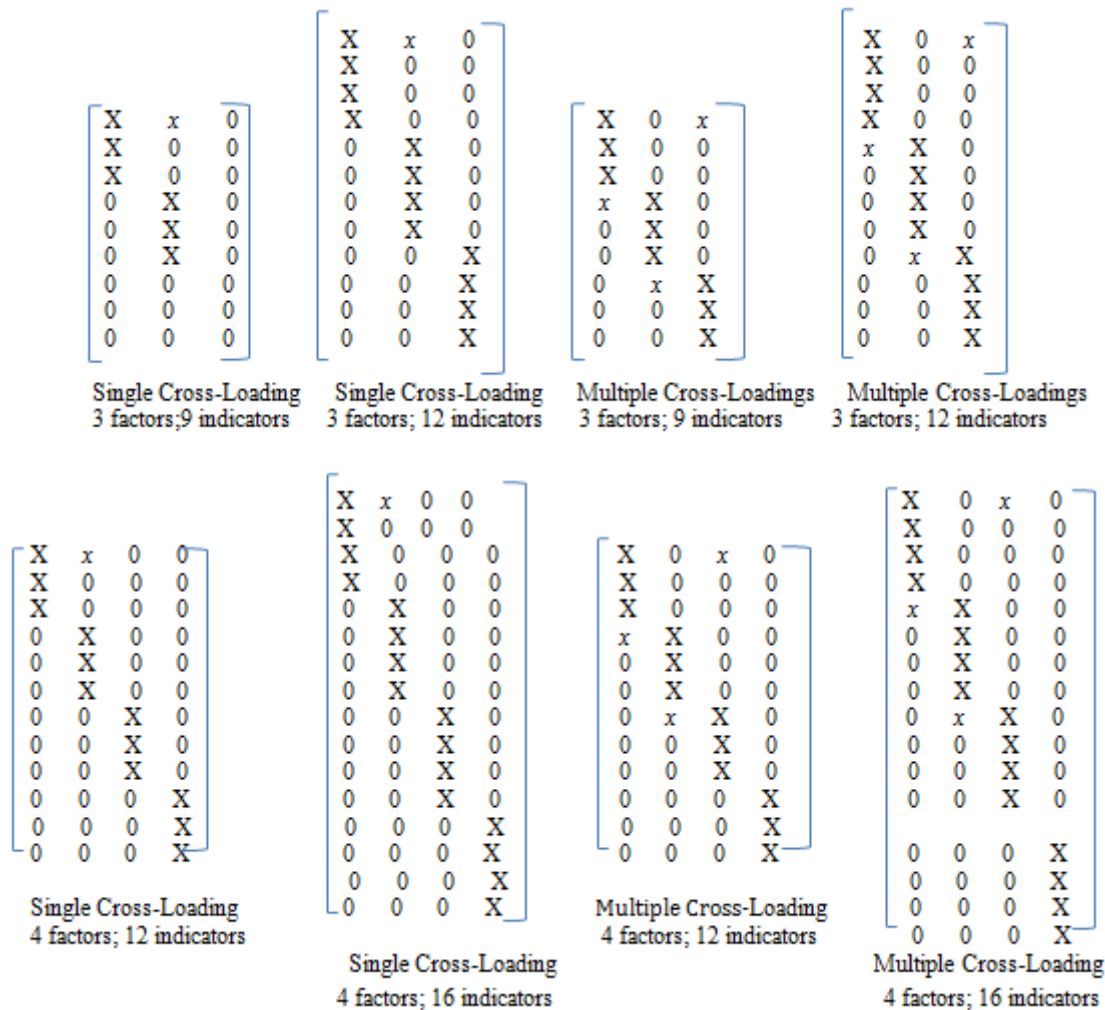


**Figure 1.**  Factor structures tested in this study

**Table 1.** Conditions studied both for minor misspecificaton and major misspecification levels at this research

| Conditions | Number of cross-loadings | Number of factors | Number of indicators per factor | Magnitude of main loadings | N |
|---|---|---|---|---|---|
| 1 | Single | 3 | 3 | 0.4 | 200 |
| 2 | Multiple | 3 | 3 | 0.4 | 200 |
| 3 | Single | 3 | 3 | 0.7 | 200 |
| 4 | Multiple | 3 | 3 | 0.7 | 200 |
| 5 | Single | 3 | 3 | 0.4 | 500 |
| 6 | Multiple | 3 | 3 | 0.4 | 500 |
| 7 | Single | 3 | 3 | 0.7 | 500 |
| 8 | Multiple | 3 | 3 | 0.7 | 500 |
| 9 | Single | 3 | 4 | 0.4 | 200 |
| 10 | Multiple | 3 | 4 | 0.4 | 200 |
| 11 | Single | 3 | 4 | 0.7 | 200 |
| 12 | Multiple | 3 | 4 | 0.7 | 200 |
| 13 | Single | 3 | 4 | 0.4 | 500 |
| 14 | Multiple | 3 | 4 | 0.4 | 500 |
| 15 | Single | 3 | 4 | 0.7 | 500 |
| 16 | Multiple | 3 | 4 | 0.7 | 500 |
| 17 | Single | 4 | 3 | 0.4 | 200 |
| 18 | Multiple | 4 | 3 | 0.4 | 200 |
| 19 | Single | 4 | 3 | 0.7 | 200 |
| 20 | Multiple | 4 | 3 | 0.7 | 200 |
| 21 | Single | 4 | 3 | 0.4 | 500 |
| 22 | Multiple | 4 | 3 | 0.4 | 500 |
| 23 | Single | 4 | 3 | 0.7 | 500 |
| 24 | Multiple | 4 | 3 | 0.7 | 500 |
| 25 | Single | 4 | 4 | 0.4 | 200 |
| 26 | Multiple | 4 | 4 | 0.4 | 200 |
| 27 | Single | 4 | 4 | 0.7 | 200 |
| 28 | Multiple | 4 | 4 | 0.7 | 200 |
| 29 | Single | 4 | 4 | 0.4 | 500 |
| 30 | Multiple | 4 | 4 | 0.4 | 500 |
| 31 | Single | 4 | 4 | 0.7 | 500 |
| 32 | Multiple | 4 | 4 | 0.7 | 500 |

## 2.2. Estimation and Prior Specification

In this study, ML estimation method for Frequentist CFA and Bayesian estimation method for BSEM approach were used. In this study, as in a study [20], it was assumed that factor covariance matrix ($\Phi$) positive definite matrix and residual covariance matrix ($\Psi$) are a diagonal matrix (off-diagonal elements are zero). For identification purpose, the factor variances of the tested models were specified at 1, and latent means were specified at 0. Accordingly, within the Bayesian framework, conjugate priors were specified for main loadings, cross-loadings, residual variances and factor covariances in CFA. Default priors in *Mplus 7* were taken into consideration as priors for the parameters except for cross-loadings. Accordingly, N $(0, \infty)$ for main loadings, IG $(-1,0)$ for residual variances ($\psi_{jj}$), and IW $(0.000, -4)$ for factor variances and covariances were specified as prior distributions. For prior specifications for cross-loadings parameters, Muthén and Asparauhov's [11] study was taken into account. Researchers states that a normal prior with 0 mean and 0.01 variance will enable cross-loadings to approximate 0; however, this prior is evaluated as a strong informative prior since it is not exactly 0. Accordingly, in this simulation study, an informative prior as $\lambda \sim N$ $(0, 0.01)$ for cross-loadings parameters was specified. Default settings in *Mplus* were used in the estimation of all models examined in the study [11, 34].

## 2.3. Analytic Strategy

In this study, the performance of CFA based on Bayesian approach and the performance of CFA based on Frequentist approach were compared in terms of the model fit, power to reject the incorrect model (indicates power to detect model misspecification), parameter coverage and power to detect cross-loadings. Accordingly, in the first stage, ML-CFA and Bayes CFA were conducted on the basis of the data sets which are generated according to the model which does not include cross-loadings for each condition specified in the simulation design. Afterwards on the basis of the data sets generated under the conditions of cross-loading=0.1 and cross-loading=0.3, ML-CFA and Bayes CFA were carried out by ignoring cross-loadings (not estimated) and without specifying informative prior. In the second stage, CFA models were tested within the frame of BSEM by specifiying informative small-variance priors for cross-loadings specified under each condition [11, 34]. In Bayesian approach, MCMC sampling was utilized to sample posterior distributions related to parameters. During MCMC sampling, it was assessed whether convergence was achieved or not on the basis of PSR value. PSR values smaller than 1.1 indicate that convergence is achieved [28,29]. Fit of the models, which were tested on the basis of both two approaches within the context of the simulation study, were compared on the basis of RMSEA and SRMR values for Frequentist CFA and ppp-values for Bayes CFA. The power to detect model misspecification of Bayesian approach with non-informative prior and Frequentist approach was examined based on ML LRT rejection rates and ppp-rejection rates, respectively. In order to make comparison related to parameter estimations, the 95% coverage values were taken into consideration. Within Frequentist approach, these values are considered as 95% coverage of the related parameters. However, these values indicate at what percentage (%) of the simulation iterations in Bayesian approach BCI covers the population parameters utilized in data generation. Besides that, average estimates, standard deviations, standard error and

mean-square-error values related to parameter estimations based on both two approaches were examined. % Significant Coefficient (%Sig Coeff) was utilized in evaluating BSEM approach's power to detect cross-loadings [11, 34].

# 3. Results

In this section, firstly convergence results in all analyses carried out in the simulation studies are presented and interpreted. Following that, Frequentist and Bayesian approaches are compared in terms of the findings related to model fit and parameter coverage for all models tested respectively in the conditions with no-cross-loadings, in which cross-loading is not estimated (minor and major misspecification), and BSEM with formative priors for cross-loadings. After that, the findings related to power of both approaches to detect model misspecification are presented and comparatively interpreted for both two approaches. Finally, the findings related to BSEM approach's power to detect cross-loadings are presented and interpreted.

## 3.1. Convergence Assessment

PSR values for convergence were assessed for each analysis conducted on the basis of Bayesian approach. It has been seen that convergence was achieved for all models tested on the basis of Bayesian approach. PSR values for the models with no cross-loadings varied between 1.041 and 1.09, for the models with minor misspecification varied between 1.036 and 1.093 and for the models with major misspecification varied between 1.035-1.093. It has been observed that convergence was quickly achieved for the models with no-cross loading whereas the convergence was more slowly achieved for the

models with major misspecification. PSR values in the analyses based on BSEM approach changed between 1.041 and 1.098.

## 3.2. Model fit and Parameter Coverage

Initially, the analyses were carried out for the data sets generated on the basis of the models with no cross-loadings in order to control the performance of the both two approaches for the models with no-cross loadings. Model fit indexes (RMSEA=0.008- 0.044, SRMR=0.020- 0.048) calculated at the end of testing these models with ML-CFA indicate that the relevant models showed a good fit to the data. When model fit was assessed on the basis of posterior predictive p-values (in Bayes-CFA with non-informative priors), it was seen that ppp-values were higher than 0.1 cutoffs (ppp-values varied between 0.010 and 0.026) for all conditions in a way that would indicate acceptable model fit. It has been observed that there was not a considerable change in the model fit according to the number of factors and indicators, magnitude of factor loading, and sample size when the relevant models were tested on the basis of Frequentist approach. However, generally higher ppp-values were obtained for more complex models when Bayes CFA was conducted (with non-informative priors). In view of model fit assessment, it has been understood that ML-CFA performed better in comparison with Bayes CFA in testing models with no cross-loadings. ML-CFA based on Frequentist approach and Bayes CFA based on Bayesian approach were compared in terms of parameter estimations and parameter coverage in addition to model fit. Because of the space limitation, only factor loading value related to y3 indicator for main loadings parameters, and average estimations and 95% coverage values related to the correlation calculated for the relationship between factor1-factor3 for factor correlations were calculated and presented here in Table 2.

**Table 2.** ML-CFA and Bayes CFA Results: No Cross-Loading

| Situation | Parameters | ML-CFA | | | BAYES CFA | | |
|---|---|---|---|---|---|---|---|
| | | Estimates | 95% Coverage | %Sig. Coeff | Estimates | 95% Coverage | %Sig. Coeff |
| 3 factors, 9 indicators, λ=0.4, N=200 | Main Loading | 0.406 | 0.942 | 1.000 | 0.401 | 0.918 | 1.000 |
| | Factor correlation | 0.310 | 0.928 | 0.692 | 0.300 | 0.944 | 0.726 |
| 3 factors, 9 indicators, λ=0.7, N=200 | Main Loading | 0.701 | 0.956 | 1.000 | 0.701 | 0.960 | 1.000 |
| | Factor correlation | 0.305 | 0.936 | 0.946 | 0.303 | 0.956 | 0.940 |
| 3 factors, 9 indicators, λ=0.4, N=500 | Main Loading | 0.402 | 0.966 | 1.000 | 0.402 | 0.942 | 1.000 |
| | Factor correlation | 0.302 | 0.938 | 0.970 | 0.298 | 0.940 | 0.990 |
| 3 factors, 9 indicators, λ=0.7, N=500 | Main Loading | 0.701 | 0.964 | 1.000 | 0.703 | 0.940 | 1.000 |
| | Factor correlation | 0.302 | 0.936 | 1.000 | 0.301 | 0.962 | 1.000 |
| 3 factors, 12 indicators, λ=0.4, N=200 | Main Loading | 0.405 | 0.936 | 1.000 | 0.409 | 0.936 | 1.000 |
| | Factor correlation | 0.299 | 0.946 | 0.778 | 0.294 | 0.942 | 0.820 |
| 3 factors, 12 indicators, λ=0.7, N=200 | Main Loading | 0.701 | 0.948 | 1.000 | 0.707 | 0.958 | 1.000 |
| | Factor correlation | 0.300 | 0.964 | 1.000 | 0.305 | 0.946 | 0.954 |
| 3 factors, 12 indicators, λ=0.4, N=500 | Main Loading | 0.403 | 0.954 | 1.000 | 0.399 | 0.926 | 1.000 |
| | Factor correlation | 0.302 | 0.950 | 0.994 | 0.299 | 0.944 | 0.996 |
| 3 factors, 12 indicators, λ=0.7, N=500 | Main Loading | 0.701 | 0.958 | 1.000 | 0.699 | 0.948 | 1.000 |
| | Factor correlation | 0.301 | 0.938 | 1.000 | 0.303 | 0.954 | 1.000 |
| 4 factors, 12 indicators, λ=0.4, N=200 | Main Loading | 0.410 | 0.952 | 0.998 | 0.390 | 0.928 | 1.000 |
| | Factor correlation | 0.294 | 0.944 | 0.658 | 0.305 | 0.950 | 0.726 |
| 4 factors, 12 indicators, λ=0.7, N=200 | Main Loading | 0.702 | 0.952 | 1.000 | 0.695 | 0.936 | 1.000 |
| | Factor correlation | 0.206 | 0.936 | 0.936 | 0.305 | 0.958 | 0.944 |
| 4 factors, 12 indicators, λ=0.4, N=500 | Main Loading | 0.403 | 0.942 | 1.000 | 0.398 | 0.926 | 1.000 |
| | Factor correlation | 0.301 | 0.950 | 0.986 | 0.305 | 0.954 | 0.990 |
| 4 factors, 12 indicators, λ=0.7, N=500 | Main Loading | 0.699 | 0.946 | 1.000 | 0.701 | 0.954 | 1.000 |
| | Factor correlation | 0.299 | 0.948 | 1.000 | 0.306 | 0.962 | 1.000 |
| 4 factors, 16 indicators, λ=0.4, N=200 | Main Loading | 0.406 | 0.946 | 1.000 | 0.401 | 0.944 | 1.000 |
| | Factor correlation | 0.300 | 0.926 | 0.766 | 0.306 | 0.944 | 0.806 |
| 4 factors, 16 indicators, λ=0.7, N=200 | Main Loading | 0.701 | 0.956 | 1.000 | 0.699 | 0.950 | 1.000 |
| | Factor correlation | 0.297 | 0.930 | 0.950 | 0.310 | 0.962 | 0.948 |
| 4 factors, 16 indicators, λ=0.4, N=500 | Main Loading | 0.403 | 0.954 | 1.000 | 0.399 | 0.952 | 1.000 |
| | Factor correlation | 0.299 | 0.952 | 1.000 | 0.301 | 0.960 | 1.000 |
| 4 factors, 16 indicators, λ=0.7, N=500 | Main Loading | 0.701 | 0.962 | 1.000 | 0.700 | 0.964 | 1.000 |
| | Factor correlation | 0.298 | 0.958 | 1.000 | 0.302 | 0.960 | 1.000 |

When Table 2 is examined, it has been understood that coverage values, which were very close to 95% related to main loadings and factor correlation parameters, were obtained for both ML-CFA and Bayes CFA. Also there was not a considerable difference in parameter coverage values between conditions and between Frequentist and Bayesian Approaches.

Fit indexes (RMSEA=.012- .041 and SRMR =.024- .050) calculated based on Frequentist approach for models with minor misspecification indicate a good fit. However, it was observed that all models with major misspecification tested in single cross-loading conditions (number of omitted cross-loading=1) showed a good fit to the data (RMSEA= .030- .068 and SRMR= .034- .053). But many of the models with major misspecification (*except condition 26*) tested in multiple cross-loadings conditions (number of omitted cross-loadings=3) showed a poor fit to the data (RMSEA= *.055*- .121 and SRMR= *.047*- .111). In order to compare two approaches in terms of parameter estimates, initially average estimates and 95% coverage values obtained on the basis of ML-CFA are given in Table 3 below and interpreted.

**Table 3.** ML-CFA Results for Minor Misspecification and Major Misspecificaton Levels

| Condition | Misspec | Parameter | Estimate | 95% Coverage | %Sig.Coef. | Condition | Misspec. | Parameter | Estimate | 95% Coverage | %Sig.Coef. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Minor | Main Loding | 0.405 | 0.958 | 1.000 | 17 | Minor | Main Loading | 0.407 | 0.954 | 1.000 |
| | | Factor correlation | 0.316 | 0.924 | 0.742 | | | Factor correlation | 0.301 | 0.944 | 0.698 |
| | Major | Main Loading | 0.406 | 0.962 | 1.000 | | Major | Main Loading | 0.407 | 0.972 | 1.000 |
| | | Factor correlation | 0.280 | 0.870 | 0.764 | | | Factor correlation | 0.271 | 0.886 | 0.744 |
| 2 | Minor | Main Loading | 0.405 | 0.968 | 1.000 | 18 | Minor | Main Loading | 0.407 | 0.952 | 1.000 |
| | | Factor correlation | 0.376 | 0.938 | 0.914 | | | Factor correlation | 0.365 | 0.950 | 0.920 |
| | Major | Main Loading | 0.405 | 0.968 | 1.000 | | Major | Main Loading | 0.406 | 0.960 | 1.000 |
| | | Factor correlation | 0.444 | 0.838 | 0.980 | | | Factor correlation | 0.441 | 0.858 | 0.990 |
| 3 | Minor | Main Loading | 0.701 | 0.958 | 1.000 | 19 | Minor | Main Loading | 0.702 | 0.954 | 1.000 |
| | | Factor correlation | 0.312 | 0.942 | 0.960 | | | Factor correlation | 0.304 | 0.940 | 0.952 |
| | Major | Main Loading | 0.701 | 0.964 | 1.000 | | Major | Main Loading | 0.702 | 0.968 | 1.000 |
| | | Factor correlation | 0.321 | 0.944 | 0.974 | | | Factor correlation | 0.313 | 0.942 | 0.972 |
| 4 | Minor | Main Loading | 0.701 | 0.958 | 1.000 | 20 | Minor | Main Loading | 0.702 | 0.956 | 0.986 |
| | | Factor correlation | 0.353 | 0.928 | 0.984 | | | Factor correlation | 0.345 | 0.946 | 1.000 |
| | Major | Main Loading | 0.701 | 0.960 | 1.000 | | Major | Main Loading | 0.702 | 0.966 | 1.000 |
| | | Factor correlation | 0.438 | 0.716 | 1.000 | | | Factor correlation | 0.433 | 0.728 | 0.998 |
| 5 | Minor | Main Loading | 0.402 | 0.968 | 1.000 | 21 | Minor | Main Loading | 0.402 | 0.946 | 1.000 |
| | | Factor correlation | 0.309 | 0.936 | 0.982 | | | Factor correlation | 0.308 | 0.952 | 0.996 |
| | Major | Main Loading | 0.402 | 0.972 | 1.000 | | Major | Main Loading | 0.402 | 0.962 | 1.000 |
| | | Factor correlation | 0.269 | 0.848 | 1.000 | | | Factor correlation | 0.274 | 0.884 | 0.998 |
| 6 | Minor | Main Loading | 0.402 | 0.968 | 1.000 | 22 | Minor | Main Loading | 0.402 | 0.942 | 1.000 |
| | | Factor correlation | 0.375 | 0.866 | 1.000 | | | Factor correlation | 0.374 | 0.908 | 1.000 |
| | Major | Main Loading | 0.401 | 0.974 | 1.000 | | Major | Main Loading | 0.402 | 0.960 | 1.000 |
| | | Factor correlation | 0.452 | 0.542 | 1.000 | | | Factor correlation | 0.454 | 0.532 | 1.000 |
| 7 | Minor | Main Loading | 0.701 | 0.970 | 1.000 | 23 | Minor | Main Loading | 0.700 | 0.950 | 1.000 |
| | | Factor correlation | 0.310 | 0.940 | 1.000 | | | Factor correlation | 0.308 | 0.956 | 1.000 |
| | Major | Main Loading | 0.701 | 0.974 | 1.000 | | Major | Main Loading | 0.700 | 0.960 | 1.000 |
| | | Factor correlation | 0.319 | 0.942 | 1.000 | | | Factor correlation | 0.317 | 0.950 | 1.000 |
| 8 | Minor | Main Loading | 0.701 | 0.970 | 1.000 | 24 | Minor | Main Loading | 0.700 | 0.952 | 1.000 |
| | | Factor correlation | 0.352 | 0.872 | 1.000 | | | Factor correlation | 0.345 | 0.900 | 1.000 |
| | Major | Main Loading | 0.701 | 0.976 | 1.000 | | | Main Loading | 0.700 | 0.960 | 1.000 |
| | | Factor correlation | 0.443 | 0.322 | 1.000 | | Major | Factor correlation | 0.440 | 0.330 | 1.000 |
| 9 | Minor | Main Loading | 0.405 | 0.940 | 1.000 | 25 | Minor | Main Loading | 0.406 | 0.944 | 1.000 |
| | | Factor correlation | 0.309 | 0.958 | 0.814 | | | Factor correlation | 0.306 | 0.934 | 1.000 |
| | Major | Main Loading | 0.405 | 0.954 | 1.000 | | Major | Main Loading | 0.406 | 0.954 | 1.000 |
| | | Factor correlation | 0.310 | 0.962 | 0.892 | | | Factor correlation | 0.308 | 0.916 | 0.874 |
| 10 | Minor | Main Loading | 0.405 | 0.940 | 1.000 | 26 | Minor | Main Loading | 0.405 | 0.944 | 1.000 |
| | | Factor correlation | 0.363 | 0.940 | 1.000 | | | Factor correlation | 0.360 | 0.938 | 0.930 |
| | Major | Main Loading | 0.404 | 0.952 | 0.956 | | Major | Main Loading | 0.406 | 0.958 | 1.000 |
| | | Factor correlation | 0.459 | 0.954 | 1.000 | | | Factor correlation | 0.456 | 0.822 | 0.996 |
| 11 | Minor | Main Loading | 0.701 | 0.826 | 0.992 | 27 | Minor | Main Loading | 0.701 | 0.954 | 1.000 |
| | | Factor correlation | 0.307 | 0.946 | 1.000 | | | Factor correlation | 0.304 | 0.936 | 0.952 |
| | Major | Main Loading | 0.701 | 0.970 | 0.976 | | Major | Main Loading | 0.701 | 0.956 | 1.000 |
| | | Factor correlation | 0.318 | 0.952 | 1.000 | | | Factor correlation | 0.316 | 0.938 | 0.970 |
| 12 | Minor | Main Loading | 0.701 | 0.976 | 0.986 | 28 | Minor | Main Loading | 0.701 | 0.952 | 1.000 |
| | | Factor correlation | 0.339 | 0.962 | 1.000 | | | Factor correlation | 0.336 | 0.934 | 0.986 |
| | Major | Main Loading | 0.701 | 0.948 | 0.990 | | Major | Main Loading | 0.701 | 0.952 | 1.000 |
| | | Factor correlation | 0.410 | 0.828 | 1.000 | | | Factor correlation | 0.117 | 0.091 | 0.867 |
| 13 | Minor | Main Loading | 0.402 | 0.954 | 1.000 | 29 | Minor | Main Loading | 0.403 | 0.956 | 1.000 |
| | | Factor correlation | 0.311 | 0.960 | 1.000 | | | Factor correlation | 0.309 | 0.964 | 1.000 |
| | Major | Main Loading | 0.402 | 0.954 | 0.998 | | Major | Main Loading | 0.402 | 0.972 | 1.000 |
| | | Factor correlation | 0.315 | 0.944 | 1.000 | | | Factor correlation | 0.313 | 0.954 | 1.000 |
| 14 | Minor | Main Loading | 0.402 | 0.954 | 1.000 | 30 | Minor | Main Loading | 0.403 | 0.954 | 1.000 |
| | | Factor correlation | 0.366 | 0.912 | 1.000 | | | Factor correlation | 0.363 | 0.916 | 1.000 |
| | Major | Main Loading | 0.402 | 0.956 | 1.000 | | Major | Main Loading | 0.402 | 0.970 | 1.000 |
| | | Factor correlation | 0.469 | 0.388 | 1.000 | | | Factor correlation | 0.476 | 0.418 | 1.000 |
| 15 | Minor | Main Loading | 0.700 | 0.958 | 1.000 | 31 | Minor | Main Loading | 0.701 | 0.960 | 1.000 |
| | | Factor correlation | 0.308 | 0.948 | 1.000 | | | Factor correlation | 0.305 | 0.958 | 1.000 |
| | Major | Main Loading | 0.700 | 0.962 | 1.000 | | Major | Main Loading | 0.701 | 0.962 | 1.000 |
| | | Factor correlation | 0.320 | 0.948 | 1.000 | | | Factor correlation | 0.317 | 0.950 | 1.000 |
| 16 | Minor | Main Loading | 0.700 | 0.958 | 1.000 | 32 | Minor | Main Loading | 0.701 | 0.962 | 1.000 |
| | | Factor correlation | 0.340 | 0.918 | 1.000 | | | Factor correlation | 0.338 | 0.920 | 1.000 |
| | Major | Main Loading | 0.700 | 0.960 | 1.000 | | Major | Main Loading | 0.701 | 0.966 | 1.000 |
| | | Factor correlation | 0.411 | 0.518 | 1.000 | | | Factor correlation | 0.542 | 0.542 | 1.000 |

When ML-CFA was conducted, coverage values very close to 95% were obtained for the estimations related to main loadings parameters at both minor and major model misspecification levels. Also a considerable difference has not been found in 95% coverage values in accordance with the level of model misspecification. When Table 3 is examined, it has been understood that 95% coverage values related to main loadings parameters did not differ significantly according *to the number of omitted cross-loadings, factor loading values, sample size, number of factors, and number of indicators per factor*. In relation to factor correlation parameter, low coverage values were obtained when the models only with 3 factors and 9 indicators were tested at minor model misspecification level in the condition of n=500 (conditions 6 and 8). But coverage values very close to 95% were obtained in all other conditions. It has been seen that the estimations that were very close to the population value (0.30) for factor correlation parameter were obtained for the models with minor misspecification. For the models with major misspecification, 95% coverage values for factor correlation parameter changed between the conditions tested in the simulation study. There was no systematic change for coverage values for this parameter according to the sample size. However, at the multiple cross-loadings conditions in which models tested with large samples (n=500) lower coverage values were obtained. Also generally coverage values very close to 95% were obtained for single cross-loading conditions related to factor correlations parameters whereas it has been observed that coverage values were low for multiple cross-loadings conditions. Especially, a considerably low

coverage value was obtained in the condition in which the most complex model (4 factors with 16 indicators) was tested with the strong indicators ($\lambda$=.7), n=200 people. This finding shows that biased parameter estimations related to factor correlations can be obtained when the more complex models are tested in small samples even if they have strong indicators. It has been generally observed that factor correlations were over-estimated in multiple cross-loading conditions at the level of major model misspecification. These findings reveal that when ML-CFA was conducted, minor model misspecification generally did not affect significantly the estimation of the correlations between factors. On the other hand, major model misspecification significantly affected the estimation of correlations between factors in accordance with the number of omitted cross-loading. It has been found that ppp-values (varied between 0.014 and 0.106) calculated for the models with minor misspecification at single cross-loading conditions indicate acceptable model fit in Bayes CFA. However, it has been observed that ppp-values (varied between 0.032 and 0.538) for models with minor misspecification tested at multiple cross-loading conditions were high in a way that would indicate better model fit. For models with major misspecification, higher ppp-values (varied between 0.106 and 1.00) were obtained. At both minor and major model misspecification levels it has been found that the performance of Bayes-CFA in terms of model fit changed according to the number of omitted cross-loadings. Average estimates and 95% coverage values, which were obtained when the models were tested through Bayes CFA, were given in Table 4 below and interpreted.

**Table 4.** Bayes CFA Results for Minor Misspecification and Major Misspecificaton Levels

| Condition | Misspec | Parameter | Estimate | 95% Coverage | %Sig.Coef | Condition | Misspec. | Parameter | Estimate | 95% Coverage | %Sig.Coef. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Minor | Main Loading | 0.402 | 0.954 | 1.000 | 17 | Minor | Main Loading | 0.391 | 0.936 | 1.000 |
|  |  | Factor correlation | 0.303 | 0.942 | 0.788 |  |  | Factor correlation | 0.311 | 0.960 | 0.772 |
|  | Major | Main Loading | 0.401 | 0.970 | 1.000 |  | Major | Main Loading | 0.389 | 0.972 | 1.000 |
|  |  | Factor correlation | 0.282 | 0.932 | 0.864 |  |  | Factor correlation | 0.293 | 0.954 | 0.898 |
| 2 | Minor | Main Loading | 0.404 | 0.950 | 1.000 | 18 | Minor | Main Loading | 0.391 | 0.940 | 1.000 |
|  |  | Factor correlation | 0.366 | 0.930 | 0.944 |  |  | Factor correlation | 0.379 | 0.910 | 0.928 |
|  | Major | Main Loading | 0.400 | 0.958 | 1.000 |  | Major | Main Loading | 0.390 | 0.966 | 1.000 |
|  |  | Factor correlation | 0.457 | 0.722 | 0.988 |  |  | Factor correlation | 0.475 | 0.682 | 0.998 |
| 3 | Minor | Main Loading | 0.703 | 0.946 | 1.000 | 19 | Minor | Main Loading | 0.697 | 0.944 | 1.000 |
|  |  | Factor correlation | 0.310 | 0.950 | 0.960 |  |  | Factor correlation | 0.312 | 0.962 | 0.958 |
|  | Major | Main Loading | 0.702 | 0.962 | 1.000 |  | Major | Main Loading | 0.695 | 0.958 | 1.000 |
|  |  | Factor correlation | 0.318 | 0.946 | 0.980 |  |  | Factor correlation | 0.322 | 0.948 | 0.976 |
| 4 | Minor | Main Loading | 0.703 | 0.946 | 1.000 | 20 | Minor | Main Loading | 0.699 | 0.948 | 1.000 |
|  |  | Factor correlation | 0.356 | 0.902 | 0.986 |  |  | Factor correlation | 0.357 | 0.914 | 0.984 |
|  | Major | Main Loading | 0.703 | 0.968 | 1.000 |  | Major | Main Loading | 0.695 | 0.952 | 1.000 |
|  |  | Factor correlation | 0.442 | 0.618 | 1.000 |  |  | Factor correlation | 0.449 | 0.600 | 1.000 |
| 5 | Minor | Main Loading | 0.402 | 0.950 | 1.000 | 21 | Minor | Main Loading | 0.399 | 0.960 | 1.000 |
|  |  | Factor correlation | 0.305 | 0.954 | 0.994 |  |  | Factor correlation | 0.311 | 0.954 | 0.988 |
|  | Major | Main Loading | 0.401 | 0.972 | 1.000 |  | Major | Main Loading | 0.397 | 0.966 | 1.000 |
|  |  | Factor correlation | 0.270 | 0.890 | 0.996 |  |  | Factor correlation | 0.283 | 0.920 | 1.000 |
| 6 | Minor | Main Loading | 0.401 | 0.952 | 1.000 | 22 | Minor | Main Loading | 0.399 | 0.948 | 1.000 |
|  |  | Factor correlation | 0.370 | 0.854 | 1.000 |  |  | Factor correlation | 0.379 | 0.848 | 1.000 |
|  | Major | Main Loading | 0.402 | 0.964 | 1.000 |  | Major | Main Loading | 0.398 | 0.962 | 1.000 |
|  |  | Factor correlation | 0.450 | 0.446 | 1.000 |  |  | Factor correlation | 0.462 | 0.374 | 1.000 |
| 7 | Minor | Main Loading | 0.702 | 0.954 | 1.000 | 23 | Minor | Main Loading | 0.700 | 0.950 | 1.000 |
|  |  | Factor correlation | 0.309 | 0.964 | 1.000 |  |  | Factor correlation | 0.314 | 0.952 | 1.000 |
|  | Major | Main Loading | 0.701 | 0.964 | 1.000 |  | Major | Main Loading | 0.700 | 0.970 | 1.000 |
|  |  | Factor correlation | 0.318 | 0.950 | 1.000 |  |  | Factor correlation | 0.323 | 0.952 | 1.000 |
| 8 | Minor | Main Loading | 0.702 | 0.946 | 1.000 | 24 | Minor | Main Loading | 0.700 | 0.956 | 1.000 |
|  |  | Factor correlation | 0.351 | 0.872 | 1.000 |  |  | Factor correlation | 0.356 | 0.850 | 1.000 |
|  | Major | Main Loading | 0.702 | 0.962 | 1.000 |  | Major | Main Loading | 0.700 | 0.962 | 1.000 |
|  |  | Factor correlation | 0.441 | 0.260 | 1.000 |  |  | Factor correlation | 0.447 | 0.242 | 1.000 |
| 9 | Minor | Main Loading | 0.410 | 0.950 | 1.000 | 25 | Minor | Main Loading | 0.399 | 0.946 | 1.000 |
|  |  | Factor correlation | 0.301 | 0.934 | 0.870 |  |  | Factor correlation | 0.315 | 0.948 | 0.848 |
|  | Major | Main Loading | 0.411 | 0.956 | 1.000 |  | Major | Main Loading | 0.397 | 0.954 | 1.000 |
|  |  | Factor correlation | 0.305 | 0.926 | 0.922 |  |  | Factor correlation | 0.323 | 0.948 | 0.906 |
| 10 |  | Main Loading | 0.411 | 0.944 | 1.000 | 26 |  | Main Loading | 0.398 | 0.948 | 1.000 |
|  | Minor | Factor correlation | 0.355 | 0.916 | 0.950 |  | Minor | Factor correlation | 0.371 | 0.918 | 0.934 |
|  | Major | Main Loading | 0.409 | 0.962 | 1.000 |  | Major | Main Loading | 0.398 | 0.956 | 1.000 |
|  |  | Factor correlation | 0.455 | 0.716 | 1.000 |  |  | Factor correlation | 0.476 | 0.648 | 1.000 |
| 11 | Minor | Main Loading | 0.707 | 0.944 | 1.000 | 27 | Minor | Main Loading | 0.699 | 0.942 | 1.000 |
|  |  | Factor correlation | 0.311 | 0.946 | 0.966 |  |  | Factor correlation | 0.317 | 0.960 | 0.958 |
|  | Major | Main Loading | 0.706 | 0.962 | 1.000 |  | Major | Main Loading | 0.698 | 0.960 | 1.000 |
|  |  | Factor correlation | 0.324 | 0.944 | 0.986 |  |  | Factor correlation | 0.330 | 0.948 | 0.974 |
| 12 | Minor | Main Loading | 0.706 | 0.960 | 1.000 | 28 | Minor | Main Loading | 0.698 | 0.948 | 1.000 |
|  |  | Factor correlation | 0.344 | 0.924 | 0.984 |  |  | Factor correlation | 0.350 | 0.934 | 0.986 |
|  | Major | Main Loading | 0.707 | 0.952 | 1.000 |  | Major | Main Loading | 0.698 | 0.964 | 1.000 |
|  |  | Factor correlation | 0.417 | 0.732 | 1.000 |  |  | Factor correlation | 0.173 | 0.389 | 0.988 |
| 13 | Minor | Main Loading | 0.400 | 0.942 | 1.000 | 29 | Minor | Main Loading | 0.400 | 0.954 | 1.000 |
|  |  | Factor correlation | 0.309 | 0.950 | 0.998 |  |  | Factor correlation | 0.310 | 0.948 | 1.000 |
|  | Major | Main Loading | 0.399 | 0.934 | 1.000 |  | Major | Main Loading | 0.398 | 0.964 | 1.000 |
|  |  | Factor correlation | 0.312 | 0.928 | 1.000 |  |  | Factor correlation | 0.316 | 0.944 | 1.000 |
| 14 | Minor | Main Loading | 0.400 | 0.940 | 1.000 | 30 | Minor | Main Loading | 0.399 | 0.956 | 1.000 |
|  |  | Factor correlation | 0.366 | 0.854 | 1.000 |  |  | Factor correlation | 0.370 | 0.840 | 1.000 |
|  | Major | Main Loading | 0.391 | 0.946 | 1.000 |  | Major | Main Loading | 0.399 | 0.962 | 1.000 |
|  |  | Factor correlation | 0.467 | 0.316 | 1.000 |  |  | Factor correlation | 0.474 | 0.304 | 1.000 |
| 15 | Minor | Main Loading | 0.699 | 0.942 | 1.000 | 31 | Minor | Main Loading | 0.699 | 0.956 | 1.000 |
|  |  | Factor correlation | 0.310 | 0.946 | 1.000 |  |  | Factor correlation | 0.309 | 0.952 | 1.000 |
|  | Major | Main Loading | 0.699 | 0.948 | 1.000 |  | Major | Main Loading | 0.700 | 0.958 | 1.000 |
|  |  | Factor correlation | 0.322 | 0.928 | 1.000 |  |  | Factor correlation | 0.321 | 0.932 | 1.000 |
| 16 | Minor | Main Loading | 0.699 | 0.942 | 1.000 | 32 | Minor | Main Loading | 0.700 | 0.958 | 1.000 |
|  |  | Factor correlation | 0.342 | 0.874 | 1.000 |  |  | Factor correlation | 0.341 | 0.894 | 1.000 |
|  | Major | Main Loading | 0.699 | 0.946 | 1.000 |  | Major | Main Loading | 0.700 | 0.962 | 1.000 |
|  |  | Factor correlation | 0.416 | 0.460 | 1.000 |  |  | Factor correlation | 0.415 | 0.482 | 1.000 |

When Table 4 is examined, it has been understood that coverage values that were very close to 95% were obtained for the estimations related to main loadings parameter in all the conditions on the basis of Bayes CFA. Similar to the table emerged when ML-CFA was conducted, 95% coverage values here did not differ significantly *according to the level of model misspecification, number of omitted cross-loadings, factor loading value, sample size, number of factors, and number of indicators per factor*. It has been understood that 95% coverage values related to factor correlation parameter were lower in the multiple cross-loadings conditions in which the models with minor misspecification were tested with generally large samples (n=500). On the other hand, they were high in other conditions. For the models with major misspecification, coverage values which were generally very close to 95% were obtained for single cross loading conditions. But low coverage values related to factor correlation parameter were obtained for all multiple cross loadings conditions. It has been found that the lowest coverage value related to factor correlation parameter estimated on the basis of ML-CFA was obtained in the condition (condition 28) in which the most complex model was tested with strong indicators and small sample. And as for model testing with Bayes CFA, the lowest coverage value for the relevant parameter was obtained in the condition (condition 24) in which the model with 4 factors and 12 indicators was tested with strong indicators and large sample. For the models with minor misspecification, similar to the results of ML-CFA, average estimations which were close to the population value (0.30) for factor correlation parameter were obtained in all conditions in Bayes CFA. For the models with major misspecification, the findings, which were similar to those in ML-CFA again, were obtained, and factor correlations resulted in over-estimation in multiple cross-loading conditions (except condition 28) in Bayes CFA.

It has been observed that ppp-values calculated on the basis of BSEM approach (ppp<.01) indicated poor fit except for three conditions: For the conditions in which the model with 3 factors and 12 indicators tested with small samples (condition 10 for cl=0.3, condition 12 for cl=0.1 and cl=0.3), obtained ppp-values (varied between 0.000 and 0.014) indicated acceptable model fit. In order to compare BSEM approach with ML-CFA and Bayes CFA in terms of parameter estimations and parameter coverage, average estimates and 95% coverage values for main loading and cross-loading parameters were given in Table 5 below. Because of the space limitation again, the values related to only y3 indicator for main loading and y1 indicator for cross-loading were presented here. It has been observed that factor correlation parameters were over-estimated only in the situations in which the model with 3 factors and 9 indicators was tested in multiple cross-loading conditions with cl=0.3 (for conditions 2,4 and 6 average estimates for factor correlation varied between 0.443 and 0.474). Except for these conditions, it has been observed that average estimates (varied between 0.244 and 0.380), which were close to the ones (0.30) that used to generate data related to factor correlation parameter, and high coverage values (varied between 0.920 and 0.988) were obtained.

When Table 5 is examined, it has been understood that coverage values, which were close to 95%, for all conditions related to main loading parameter were obtained. It has been observed that, for cross-loading parameter, high coverage values (very close to 95% or higher) were obtained in all conditions (when other factors are fixed) with cross-loading=0.1. Also, these values did not differ significantly between the conditions. For all conditions with cross-loading=0.3, lower coverage values related to cross-loading parameter were obtained. It was understood that there was not a considerable change in coverage values for cross-loading parameter in terms of number of cross-loading (single cross-loading vs. multiple cross-loadings). When the magnitude of main loadings and sample size increased, the coverage values related to factor correlation parameter also increased. It has been generally observed that there is an amount of increase in coverage values related to cross-loading parameter along with the increase in the number of factors for large samples (n=500) and conditions with cl=0.3. When the comparisons were made between the conditions, it has been generally understood that coverage values related to cross-loading parameter were higher in the conditions in which the models including a lot of indicators (4 indicators per factor) were tested with large samples.

**Table 5.** Results of Bayesian analyses with informative priors

| Condition | Cross-loading | Parameter | Estimate | 95% Coverage | %Sig.Coef. | Condition | Cross-loading | Parameter | Estimate | 95% Coverage | %Sig.Coef. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.1 | Main Loading | 0.432 | 0.964 | 0.998 | 17 | 0.1 | Main Loading | 0.429 | 0.966 | 1.000 |
| | | CrossLoading | 0.049 | 0.994 | 0.002 | | | CrossLoading | 0.051 | 0.996 | 0.006 |
| | 0.3 | Main Loading | 0.424 | 0.976 | 0.998 | | 0.3 | Main Loading | 0.422 | 0.982 | 1.000 |
| | | CrossLoading | 0.117 | 0.242 | 0.140 | | | CrossLoading | 0.120 | 0.264 | 0.148 |
| 2 | 0.1 | Main Loading | 0.427 | 0.976 | 0.998 | 18 | 0.1 | Main Loading | 0.423 | 0.972 | 1.000 |
| | | CrossLoading | 0.049 | 1.000 | 0.008 | | | CrossLoading | 0.060 | 0.992 | 0.006 |
| | 0.3 | Main Loading | 0.387 | 0.982 | 0.998 | | 0.3 | Main Loading | 0.385 | 0.978 | 0.998 |
| | | CrossLoading | 0.133 | 0.296 | 0.284 | | | CrossLoading | 0.134 | 0.294 | 0.258 |
| 3 | 0.1 | Main Loading | 0.716 | 0.954 | 1.000 | 19 | 0.1 | Main Loading | 0.714 | 0.964 | 1.000 |
| | | CrossLoading | 0.055 | 0.992 | 0.010 | | | CrossLoading | 0.055 | 0.998 | 0.014 |
| | 0.3 | Main Loading | 0.715 | 0.954 | 1.000 | | 0.3 | Main Loading | 0.714 | 0.974 | 1.000 |
| | | CrossLoading | 0.154 | 0.406 | 0.624 | | | CrossLoading | 0.156 | 0.436 | 0.624 |
| 4 | 0.1 | Main Loading | 0.714 | 0.946 | 1.000 | 20 | 0.1 | Main Loading | 0.711 | 0.966 | 1.000 |
| | | CrossLoading | 0.055 | 0.996 | 0.016 | | | CrossLoading | 0.056 | 0.994 | 0.014 |
| | 0.3 | Main Loading | 0.726 | 0.950 | 1.000 | | 0.3 | Main Loading | 0.693 | 0.956 | 1.000 |
| | | CrossLoading | 0.158 | 0.396 | 0.686 | | | CrossLoading | 0.160 | 0.440 | 0.706 |
| 5 | 0.1 | Main Loading | 0.418 | 0.948 | 1.000 | 21 | 0.1 | Main Loading | 0.426 | 0.950 | 1.000 |
| | | CrossLoading | 0.060 | 0.944 | 0.030 | | | CrossLoading | 0.063 | 0.998 | 0.014 |
| | 0.3 | Main Loading | 0.419 | 0.970 | 1.000 | | 0.3 | Main Loading | 0.425 | 0.958 | 1.000 |
| | | CrossLoading | 0.161 | 0.432 | 0.624 | | | CrossLoading | 0.165 | 0.530 | 0.660 |
| 6 | 0.1 | Main Loading | 0.416 | 0.960 | 1.000 | 22 | 0.1 | Main Loading | 0.424 | 0.936 | 1.000 |
| | | CrossLoading | 0.060 | 1.000 | 0.028 | | | CrossLoading | 0.063 | 0.996 | 0.028 |
| | 0.3 | Main Loading | 0.389 | 0.954 | 1.000 | | 0.3 | Main Loading | 0.394 | 0.964 | 1.000 |
| | | CrossLoading | 0.173 | 0.492 | 0.838 | | | CrossLoading | 0.176 | 0.572 | 0.844 |
| 7 | 0.1 | Main Loading | 0.706 | 0.958 | 1.000 | 23 | 0.1 | Main Loading | 0.709 | 0.948 | 1.000 |
| | | CrossLoading | 0.062 | 1.000 | 0.042 | | | CrossLoading | 0.064 | 0.998 | 0.012 |
| | 0.3 | Main Loading | 0.706 | 0.956 | 1.000 | | 0.3 | Main Loading | 0.708 | 0.948 | 1.000 |
| | | CrossLoading | 0.180 | 0.544 | 0.962 | | | CrossLoading | 0.182 | 0.582 | 0.974 |
| 8 | 0.1 | Main Loading | 0.706 | 0.952 | 1.000 | 24 | 0.1 | Main Loading | 0.708 | 0.950 | 1.000 |
| | | CrossLoading | 0.062 | 0.996 | 0.022 | | | CrossLoading | 0.062 | 1.000 | 0.024 |
| | 0.3 | Main Loading | 0.700 | 0.950 | 1.000 | | 0.3 | Main Loading | 0.699 | 0.954 | 1.000 |
| | | CrossLoading | 0.183 | 0.560 | 0.974 | | | CrossLoading | 0.185 | 0.624 | 0.972 |
| 9 | 0.1 | Main Loading | 0.454 | 0.940 | 1.000 | 25 | 0.1 | Main Loading | 0.459 | 0.944 | 1.000 |
| | | CrossLoading | 0.058 | 1.000 | 0.018 | | | CrossLoading | 0.060 | 0.998 | 0.016 |
| | 0.3 | Main Loading | 0.458 | 0.928 | 1.000 | | 0.3 | Main Loading | 0.456 | 0.948 | 1.000 |
| | | CrossLoading | 0.157 | 0.460 | 0.567 | | | CrossLoading | 0.156 | 0.500 | 0.548 |
| 10 | 0.1 | Main Loading | 0.454 | 0.932 | 1.000 | 26 | 0.1 | Main Loading | 0.454 | 0.940 | 1.000 |
| | | CrossLoading | 0.059 | 0.998 | 0.028 | | | CrossLoading | 0.059 | 1.000 | 0.024 |
| | 0.3 | Main Loading | 0.412 | 0.952 | 1.000 | | 0.3 | Main Loading | 0.410 | 0.968 | 1.000 |
| | | CrossLoading | 0.163 | 0.498 | 0.652 | | | CrossLoading | 0.163 | 0.530 | 0.670 |
| 11 | 0.1 | Main Loading | 0.726 | 0.952 | 1.000 | 27 | 0.1 | Main Loading | 0.722 | 0.958 | 1.000 |
| | | CrossLoading | 0.062 | 0.994 | 0.036 | | | CrossLoading | 0.062 | 0.990 | 0.040 |
| | 0.3 | Main Loading | 0.727 | 0.950 | 1.000 | | 0.3 | Main Loading | 0.723 | 0.954 | 1.000 |
| | | CrossLoading | 0.181 | 0.590 | 0.908 | | | CrossLoading | 0.181 | 0.574 | 0.900 |
| 12 | 0.1 | Main Loading | 0.724 | 0.958 | 1.000 | 28 | 0.1 | Main Loading | 0.720 | 0.962 | 1.000 |
| | | CrossLoading | 0.061 | 0.990 | 0.040 | | | CrossLoading | 0.063 | 0.996 | 0.048 |
| | 0.3 | Main Loading | 0.709 | 0.994 | 1.000 | | 0.3 | Main Loading | 0.410 | 0.968 | 1.000 |
| | | CrossLoading | 0.183 | 0.572 | 0.934 | | | CrossLoading | 0.163 | 0.530 | 0.670 |
| 13 | 0.1 | Main Loading | 0.426 | 0.932 | 1.000 | 29 | 0.1 | Main Loading | 0.433 | 0.946 | 1.000 |
| | | CrossLoading | 0.071 | 0.998 | 0.076 | | | CrossLoading | 0.070 | 0.996 | 0.056 |
| | 0.3 | Main Loading | 0.427 | 0.936 | 1.000 | | 0.3 | Main Loading | 0.435 | 0.952 | 1.000 |
| | | CrossLoading | 0.197 | 0.672 | 0.974 | | | CrossLoading | 0.198 | 0.740 | 0.962 |
| 14 | 0.1 | Main Loading | 0.424 | 0.942 | 1.000 | 30 | 0.1 | Main Loading | 0.430 | 0.942 | 1.000 |
| | | CrossLoading | 0.069 | 0.996 | 0.078 | | | CrossLoading | 0.070 | 0.998 | 0.058 |
| | 0.3 | Main Loading | 0.402 | 0.968 | 1.000 | | 0.3 | Main Loading | 0.407 | 0.956 | 1.000 |
| | | CrossLoading | 0.202 | 0.706 | 0.992 | | | CrossLoading | 0.202 | 0.718 | 0.990 |
| 15 | 0.1 | Main Loading | 0.709 | 0.942 | 1.000 | 31 | 0.1 | Main Loading | 0.711 | 0.954 | 1.000 |
| | | CrossLoading | 0.070 | 1.000 | 0.1000 | | | CrossLoading | 0.068 | 0.996 | 0.066 |
| | 0.3 | Main Loading | 0.710 | 0.950 | 1.000 | | 0.3 | Main Loading | 0.712 | 0.956 | 1.000 |
| | | CrossLoading | 0.208 | 0.706 | 0.996 | | | CrossLoading | 0.206 | 0.718 | 0.998 |
| 16 | 0.1 | Main Loading | 0.708 | 0.952 | 1.000 | 32 | 0.1 | Main Loading | 0.710 | 0.958 | 1.000 |
| | | CrossLoading | 0.070 | 0.998 | 0.092 | | | CrossLoading | 0.070 | 0.998 | 0.076 |
| | 0.3 | Main Loading | 0.701 | 0.958 | 1.000 | | 0.3 | Main Loading | 0.702 | 0.962 | 1.000 |
| | | CrossLoading | 0.208 | 0.690 | 1.000 | | | CrossLoading | 0.209 | 0.760 | 0.998 |

## 3.3. Detection of Model Misspecification

The power to detect model misspecification of CFA based on Frequentist approach and CFA based on Bayesian approach has been examined in terms of rejection of incorrect model. In this context, ML LRT rejection rates and ppp-rejection rates have been taken into consideration in order to compare ML-CFA and Bayes CFA (with non-informative priors), and these are presented in Table 6 below.

**Table 6.** Rejection rates for Bayes CFA and ML-CFA for minor misspecification and major misspecification levels

| Condition | Misspecification level | Rejection rate for ML LRT | Bayes PPP rejection rate | Condition | Misspecification level | Rejection rate for ML LRT | Bayes PPP rejection rate |
|---|---|---|---|---|---|---|---|
| 1 | Minor | 0.340 | 0.028 | 17 | Minor | 0.363 | 0.014 |
|  | Major | 0.073 | 0.142 |  | Major | 0.123 | 0.106 |
| 2 | Minor | 0.182 | 0.046 | 18 | Minor | 0.246 | 0.032 |
|  | Major | 0.000 | 0.828 |  | Major | 0.001 | 0.706 |
| 3 | Minor | 0.435 | 0.038 | 19 | Minor | 0.316 | 0.026 |
|  | Major | 0.003 | 0.524 |  | Major | 0.014 | 0.376 |
| 4 | Minor | 0.089 | 0.134 | 20 | Minor | 0.153 | 0.082 |
|  | Major | 0.000 | 1.000 |  | Major | 0.000 | 0.988 |
| 5 | Minor | 0.265 | 0.054 | 21 | Minor | 0.293 | 0.032 |
|  | Major | 0.002 | 0.608 |  | Major | 0.008 | 0.492 |
| 6 | Minor | 0.040 | 0.236 | 22 | Minor | 0.088 | 0.162 |
|  | Major | 0.000 | 1.000 |  | Major | 0.000 | 1.000 |
| 7 | Minor | 0.157 | 0.098 | 23 | Minor | 0.212 | 0.074 |
|  | Major | 0.000 | 0.992 |  | Major | 0.000 | 0.970 |
| 8 | Minor | 0.004 | 0.538 | 24 | Minor | 0.020 | 0.380 |
|  | Major | 0.000 | 1.000 |  | Major | 0.000 | 1.000 |
| 9 | Minor | 0.321 | 0.020 | 25 | Minor | 0.307 | 0.034 |
|  | Major | 0.034 | 0.276 |  | Major | 0.061 | 0.180 |
| 10 | Minor | 0.168 | 0.088 | 26 | Minor | 0.194 | 0.064 |
|  | Major | 0.000 | 0.920 |  | Major | 0.001 | 0.856 |
| 11 | Minor | 0.285 | 0.042 | 27 | Minor | 0.280 | 0.052 |
|  | Major | 0.004 | 0.578 |  | Major | 0.013 | 0.398 |
| 12 | Minor | 0.108 | 0.156 | 28 | Minor | 0.141 | 0.108 |
|  | Major | 0.000 | 1.000 |  | Major | 0.000 | 1.000 |
| 13 | Minor | 0.249 | 0.064 | 29 | Minor | 0.261 | 0.040 |
|  | Major | 0.001 | 0.872 |  | Major | 0.001 | 0.720 |
| 14 | Minor | 0.034 | 0.308 | 30 | Minor | 0.066 | 0.190 |
|  | Major | 0.000 | 1.000 |  | Major | 0.000 | 1.000 |
| 15 | Minor | 0.181 | 0.106 | 31 | Minor | 0.210 | 0.064 |
|  | Major | 0.000 | 0.990 |  | Major | 0.000 | 0.972 |
| 16 | Minor | 0.008 | 0.526 | 32 | Minor | 0.024 | 0.358 |
|  | Major | 0.000 | 1.000 |  | Major | 0.000 | 1.000 |

Rejection rates for ML-LRT presented at Table 6 indicate that the power of ML-CFA to detect minor model misspecification is high, whereas it's power to detect major model misspecification is so much lower. However, it has been observed that at the minor model misspecification, in single cross-loading conditions rejection rates are high and generally in multiple cross-loading conditions these rejection rates are lower. At minor model misspecification level, ML-CFA rejected approximate 44% of the models in the single cross-loading conditions in which the simplest factor structure with strong indicators was tested with small samples. At major model misspecification level, in general, the rejection rates are low. It has been understood that there is no considerable change according to the number of omitted cross-loadings. However, Bayes ppp-rejection rates indicate that the power of Bayes CFA to detect minor model misspecification is low, but it's power to detect major model misspecification is so much higher. For both minor and major misspecified models, at the single cross-loading conditions ppp-rejection rates were lower, whereas at multiple cross-loading conditions these rates were slightly higher. It has been found that the power to detect major model misspecification of Bayes CFA was low at the four conditions in which the related models were tested with weak indicators and small samples (conditions 1,9,17 and 25). However, at the all other conditions the power of Bayes CFA to detect major model misspecification was so much higher. The striking finding

is that at the major model misppecification level for 11 conditions, ML-CFA does not reject any of the incorrect models (ML-LRT=0.000). But for these same conditions Bayes CFA rejects all of the incorrect models (ppp-rejection rate=1.000). Besides this, in case of the major model misspecification, ML-CFA does not reject any of the incorrect models in another seven conditions. These findings reveal that ML-CFA is so sensitive to the minor model misspecification, especially in single cross-loading conditions and Bayes CFA is so senstitive to the major model misspecification, especially in multiple cross-loading conditions.

When the models were tested with ML-CFA, *according to magnitude of main loadings*, a systematic difference has not been observed for single cross-loading conditions at minor model misspecification level and for multiple cross-loading conditions at major model misspecification level in terms of the rejection rates. However, it could be said that the rejection rates decreased when the magnitude of factor loadings ($\lambda=.7$) increased for multiple cross-loading conditions at minor model misspecification level and single cross-loading conditions at major model misspecification level. In all conditions in Bayes CFA, it has been observed that there was an increase in ppp rejection rates along with the increase in magnitude of factor loadings. *According to the increase in sample size,* power of ML-CFA to reject the incorrect model and power of Bayes CFA to reject the incorrect model did not differ

only at major model misspecification level for multiple cross-loading conditions. Whereas the rejection rates showed a decrease in all other conditions in model testing with ML-CFA for large samples (n=500), ppp rejection rates showed an increase in model testing with Bayes CFA. It has been found that this change was much more for the conditions with strong indicators ($\lambda$=.7) for both ML-CFA and Bayes CFA.

*In accordance with the increase in the number of factors*, power of ML-CFA to reject the incorrect model showed an amount of increase only at minor model misspecification level for multiple cross-loadings conditions .But power of Bayes CFA to reject the incorrect model showed an amount of decrease at both minor and major model misspecification levels. For other conditions, *according to the number factors,* a considerable difference has not been observed in terms of both two approaches' power to reject the incorrect model. Finally, a systematic and considerable change has not been found *according to the number of indicators per factor* in terms of both ML-CFA' power to reject the incorrect model and Bayes CFA's power to reject the incorrect model.

### 3.4. Power to Detect Cross-loadings in BSEM

Power of BSEM approach to detect cross-loading has been evaluated on the basis of %sig coeff values given in Table 5. It has been understood that the power to detect cross-loadings was considerably low when informative prior with low variance was specified for cross-loading parameter, and it was cross-loading=0.1. However, it has been observed that BSEM approach had sufficient power (power higher than 0.8) to detect cross-loading for 17 conditions when it was cross-loading=0.3 [11]. The analysis results show that when it was cross-loading=0.3, power of BSEM approach to detect cross-loading was higher for some conditions: For the conditions in which each factor was generally represented by more indicators (conditions 11-16 and conditions 27, 29, 30, 31, 32) and six conditions including large samples (conditions 6-8 and 22-24). Since power of BSEM approach to detect cross-loading was too low when it was cl=0.1, the comparisons between the conditions were made only for the conditions in which it was cl=0.3. It has been observed that when it was the same in terms of other factors, power of BSEM approach to detect cross-loading was generally lower for single cross-loading conditions. However, the power showed an amount of increase for multiple cross-loading conditions. When the change in the power to detect cross-loading is examined according to magnitude of main loading, it has been understood that the power increased when there were strong indicators.

## 4. Discussion and Conclusions

Since most psychological constructs are complex in nature, the indicators, which are chosen to represent those, may be mostly related to more than one factor of the construct [35]. Therefore, when the relationships between the constructs and their indicators are modelled by considering that each indicator is only related to one factor of the construct but not certainly related to other factors, this leads the researchers to misspecified measurement models. Although biased parameter estimations are obtained as the consequence of testing a misspecified measurement model, fit indexes, which will indicate that this model is a "valid" model, can be obtained. In fact, this case means the specifications of the measurement models which do not reflect the theory well. But unfortunately this may cause a problem in the researchers' theory development process since it may hinder the accuracy of the inferences made on the basis of these models in further studies [12, 32].

There are various ways of dealing with the misspecification of the measurement model. In line with the theory, some of these ways are: (a) the specification of bifactor model in relation to the construct examined, (b) the specification of higher order factor model in an appropriate condition, and (c) allowing an indicator to load into more than one factor or the specification of the relationships in the model among the specific variances related to indicators [10, 32]. Allowing a specific indicator to load into more than one factor in the measurement model means the specification and estimation of cross-loading in the model in fact. In Frequentist approach, in specification of CFA models, an indicator's loadings for other factors are fixed at "0"; in other words, cross-loadings are not allowed. The departure point of BSEM approach is the idea that underlying theory in educational and psychological measures is reflected better when the specifications as the absolute "0" in CFA models in Frequentist approach are replaced by the approximate "0" by using an informative prior with small variance. In BSEM approach, it is allowed to both cross-loadings and use of prior information in addition to the information in data [22,34]. Accordingly, a simulation study has been conducted to examine the power of BSEM approach to detect cross-loading and to compare the performances of Frequentist and Bayesian approaches in the existence of model misspecification. In this study, model misspecification was considered in terms of omitted-cross loading. In the study, firstly the performance of ML-CFA and the performance of Bayes CFA were compared for the models with no cross-loading (correctly specified models). As in Muthen and Asparouhov's study [11], this study has found that Frequentist approach showed a better performance in comparison with Bayesian approach with non-informative priors in terms of the model fit for CFA models with no cross-loading. However, in reality, an idea that the indicator of a psychological construct is related to only a factor but not certainly related to other factors will reflect an imaginative case. Hence, it will be a more realistic approach to evaluate both these

approaches' performance in the existence of cross-loading. For that reason, this research has been needed and carried out accordingly.

For misspesified models, expectedly, it has been found that the levels of model fit differed in accordance with the level of misspecification as a consequence of the analyses conducted on the basis of both two approaches. At minor misspecification level, both two approaches lead the researchers to a decision pointing out that these models show a good/acceptable fit to the data. When the misspecification was ignorable, it has not been observed that there was a considerable difference between conditions in terms of the model fit levels in Frequentist approach. But in Bayesian approach (with non-informative priors), it has been observed that the levels of model fit changed according to the number of omitted cross-loading. From the perspective of parameter estimations, Frequentist approach was resulted in the insubstantial backing of factor correlation parameters in the condition in which the simplest factor structure examined in this study was tested with large sample (just for multiple cross-loadings conditions). Bayesian approach (with non-informative priors) approach was resulted in the insubstantial backing of factor correlation parameters in generally large sample conditions (just for multiple cross-loadings conditions). It is considered that the fact that the average standard errors in relation to factor correlation parameter estimations in the relevant conditions were higher in both two approaches than those in other conditions might cause these results. When the findings related to model fit assessment and parameter estimates are evaluated together, it has been understood that at low misspecification level, the performance of Bayes CFA and the performance of CFA within Frequentist framework are comparable in terms of the model fit and parameter estimation.

At major model misspecification, in both two approaches, the levels of model fit differed in accordance with the number of omitted cross-loading. When there is a model misspecification at this level, both Frequentist and Bayesian approaches will lead to a decision pointing out that the model is accepted as "a valid model" when there is only one omitted-cross loading. However, the value considered as the amount of omitted cross-loading (cl=0.3) at major model misspecification is taken into account as an evidence, which indicates that an indicator represents the underlying factor appropriately/sufficiently, in the literature [36]. Accordingly, the current simulation study findings implied that in both two approaches, the researchers should not relate an indicator to its underlying factor in an incorrect way. Because this may cause a problem related to the representation of the relevant factor. When more than one cross-loading are not modelled, Frequentist approach causes that most of the relevant models (except for condition 26) are rejected. On the other hand, Bayesian approach causes that all these models are incorrectly accepted as "valid model". In both two

approaches, major model misspecification caused the overestimation of factor correlation parameter only in the conditions with multiple omitted cross-loadings. Accordingly, it has been understood that the evaluation of the model fit on the basis of RMSEA and SRMR is appropriate within the framework of Frequentist approach in the conditions in which there are doubts (based on theory) indicating that more than one indicator may have a significant relationship with non-target underlying factors. However, this approach may lead the researchers to incorrect results within the context of the relationships between the factors.

When the literature [11,13,15,16,1737,38] is examined, it has been seen that the model fit has been evaluated on the basis of overall goodness of fit indexes for various types and levels of model misspecification under various conditions. Unlike the results in this study, in a study [16] it was stated that RMSEA was generally insensitive to multiple omitted cross-loadings and changed in accordance with the sample size when it was studied with strong indicators. Reference [37] reported that RMSEA changed in accordance with the magnitude of factor loadings when the misspecification was introduced as omitted cross-loading in CFA models. In another study [38], in which they examined the model misspecification within the context of omitted cross loading, that the decrease in RMESA was a function of the model size, and this decrease was higher for the smallest models.

However, in aforementioned studies and in the current study, the sensitivity of RMSEA to the model misspecification was examined under different conditions in terms of the magnitude and number of omitted cross-loading, sample size, number of factors and indicators per factor, magnitude of main loading and factor correlations. When the findings of all these studies and current study are evaluated together, it has been understood that RMSEA actually performed differently at different types and levels of model misspecification under more various model sizes and parameter values. In addition it could be concluded that the sensitivity of RMSEA changed according to the number of omitted cross-loading even if, in the current research, the sensitivity of RMSEA to model misspecification at both misspecification levels did not show a systematic change in terms of the factors examined in the research. In the literature [15] it is emphasized that goodness of fit index especially for misspecified models was affected by other features of model (number of factors and indicator and etc.) and data (like sample size and normality) in addition to model misspecification. Accordingly, a model with little misspecification may be rejected since a fit index is very sensitive for a type and level of model misspecification under specific conditions. Besides that, a major misspecified model may be accepted as "valid model" since a fit index is not sensitive enough for the relevant type and level of model misspecification. Therefore, it has been seen that when the model fit for

misspecified models is assessed on the basis of overall goodness of fit indexes, the accuracy levels of decisions to be made on the model validity change in accordance with the fit index utilized and the quality of model and data. From this point of view, it is suggested to examine the parameter estimations in addition to the utilization of multiple fit indexes in evaluating whether a model is a valid model or not. Accordingly, an examination was also conducted in the current research in terms of the parameter estimates and parameter coverage.

In this study, it has been found that whereas generally factor correlations are effectively estimated at ignorable model misspecification, factor correlations are overestimated in the conditions with multiple omitted cross-loading at major misspecification level. These findings indicate that the analyses based on Frequentist approach may lead the researchers to incorrect results about identifying the nature of underlying construct in educational and psychological measures correctly in the conditions with omitted multiple cross loading at major misspecification level. As in a study [11], similarly, it was observed that factor correlations had a tendency to be overestimated along with the increase in the magnitude of omitted cross-loading.

The results, which were similar to those in Muthen and Asparouhov's [11] study, were obtained for both two levels of misspecification on the basis of Bayesian approach (with non-informative priors) in terms of the model fit and parameter estimation. When it has been found that the models showed an acceptable/good fit to the data (based on ppp-values) at both two misspecification levels; the decrease in factor correlation parameter coverage according to the level of misspecification and number of omitted cross-loadings actually indicates the sensitivity of Bayesian approach to model misspecification in this context. Accordingly, it is considered that the fact that two approaches lead the researchers to different decisions within the context of the model fit for the important level of misspecification may stem from the difference in the nature of the criteria used in the evaluation of the model fit in these approaches. It is stated that RMSEA is sensitive to the model misspecification error especially when the model misspecification is related to factor loadings since it indicates the degree of difference between the population covariance matrix and covariance matrix generated by the model again [14,32]. However, the evaluation of the model fit according to ppp-values in Bayesian approach is based on the validity for further observations of the model [24]. As supporting this idea, it is stated in the literature [13,33] that Bayesian variant of RMSEA (BRMSEA) and Deviance information criteria (DIC) are more appropriate in the comparison of Frequentist and Bayesian approaches in terms of the evaluation of model fit in testing CFA models. Accordingly, in further studies, it can be suggested to examine the performance of DIC and the performance of BRMSEA for both correctly specified models and misspecified models in the evaluation of model fit for CFA models.

In BSEM approach, the fact that low ppp-values (indicate a poor fit) are obtained shows that the evaluation of ppp-values within the context of the model fit is not appropriate as mentioned above. The estimation of factor correlation parameter except for three conditions and main loading parameter efficiently on the basis of BSEM approach is qualified as supporting this idea. Similarly, Muthen and Asparouhov [11] reported in their study that BSEM performed well for misspecified models. Therefore, it can be also suggested in BSEM approach to carry out the studies about the evaluation of the model fit on the basis of aforementioned criteria.

Following the comparison of Frequentist and Bayesian approaches in terms of the model fit and parameter estimation, the power of these two approaches to detect model misspecification has been evaluated. It has been seen that in Frequentist approach, the power of detecting ignorable model misspecification was very high when there was only one omitted cross-loading. On the other hand, this power had a tendency to decrease when the number of omitted cross-loading increased. An inference that Bayesian approach with non-informative prior generated more reasonable results due to the fact that its sensitivity to this misspecification increased in accordance with number of omitted cross-loadings may be made. When both magnitudes of cross-loading and number of omitted cross-loading are low, misspecification is evaluated as ignorable misspecification at this level. Therefore it is considered that the fact that this misspecification leads to an inference in a way that model is invalid will not be appropriate. Also, it is understood that the misspecification at this level did not affect significantly the model fit levels and parameter estimations, and the model might be accepted as a valid model. In line with these findings, in a study [11] it was observed that whereas ML-CFA was highly sensitive to ignorable misspecification, Bayes CFA performed better in this context. In consistency with Muthen and Asparouhov's [11] study, the results of the current study (especially in the conditions with multiple cross-loadings) show that the power of Bayes CFA to detect misspecification was high. However, unlike the relevant study, the results of the current study put forward that the power of ML-CFA to detect misspecification was very low.

The most important thing is, for the deviations from the correct model at ignorable level (ignorable model misspecification) the acceptance of the models at and the rejection of the models at important level of misspecification [17]. Accordingly, it has been found that ML-CFA, which was conducted within Frequentist framework, was highly sensitive to minor model misspecification, and Bayesian approach performed better at this level of model misspecification. For the important

level of model misspecification, ML-CFA based on Frequentist approach was insufficient to detect model misspecification, and Bayes-CFA based on Bayesian approach was highly sensitive to misspecification. The most remarkable finding of this study is that for the important level of misspecification, none of incorrect models are rejected in eleven conditions on the basis of the analyses conducted within Frequentist framework. On the other hand, all incorrect models are rejected as the consequence of the analyses conducted within Bayesian framework under the same conditions. When these eleven conditions are examined in terms of the variables manipulated in the research, only one common point has been found: *All of them are the multiple cross-loadings conditions*. Therefore, it has been understood that Frequentist approach may lead the researchers to incorrect results in this subject when there are theoretical and empirical bases about the significant relationship between more than one indicator and non-target factors in the specification of the correct/appropriate model related to the factor structure of a measurement tool; Bayesian approach is robust in revealing the existence and level of model misspecification. It has been found that the power of both two approaches to detect model misspecification changed in accordance with the factor loading values, number of factors and sample size in addition to the number of omitted cross-loadings. It has been found that the power of Frequentist approach to detect model misspecification decreased for some conditions in the existence of strong indicators. But it has been observed that the power of Bayesian approach to detect model misspecification generally increased in the existence of strong indicators. This case indicates that it may be resulted in the fact that model misspecification is ignored when there are strong indicators in Frequentist approach.

Since the power of detecting model misspecification has been examined in this study by using chi-square test of exact fit in Frequentist approach and it has been evaluated on the basis of ppp-rejection rates in Bayesian approach, it is expected that this power shows change in both two approaches according to sample size [32]. In this study, it has been found that the power of ML-CFA to reject the incorrect model decreased in accordance with the increase in sample size as it was observed in a study [33]. However the power of Bayes CFA to reject the incorrect model increased as it is similar to Muthen and Asparouhov's [11] study. The high rejection rates which were found when the models with minor misspecification were tested with small samples in Frequentist approach (these are single cross-loading conditions) can be explained with the small sample bias of ML $\chi^2$ test (Muthen & Asparouhov, 2012). That the rates of rejecting incorrect models were found as high in the condition in which these models were tested with large samples in Bayesian approach (these are multiple cross-loadings conditions) can be explained with the fact that ppp-rejection rates were very sensitive to the

deviations from the correct model at ignorable level in large samples as it is stated in the literature [34]. It has been found that the number of factors in Frequentist approach affected the power of detecting model misspecification to some extent only when there were the deviations from the correct model at ignorable level (tendency to increase). The number of factors in Bayesian approach affected to some extent for each level of misspecification (tendency to decrease).

In this study, finally, the power of BSEM approach to detect cross-loading was examined. In ML-CFA conducted within Frequentist approach, sometimes the first model show a poor fit, and the researchers have doubts on the fact that various meaningful cross-loadings might be different from "0". In these situations it is required to carry out a large number of model modifications by releasing a cross-loading in each time, but this might cause a problem. On the contrary, it may be estimated freely and concurrently within BSEM framework with the addition of cross-loadings, considered as meaningful on the basis of theoretical and empirical findings in BSEM approach, to the model and without the requirement of model modification sequentially [34]. In line with these explanations, the current study results show that for the low value of cross-loading, BSEM was insufficient to reveal the existence of cross-loading required to be specified in the model. However, for the high level of cross-loading, it could effectively reveal the cross-loadings required to be specified in the model. The power of BSEM to detect cross-loading is especially higher when some indicators were affected by the factors, which are more than one, at the high level (cl=0.3), and the factors were represented by more indicators, and generally in the conditions in which the models were tested with large samples. Along with that, it has been understood that the power of BSEM to detect cross-loading increased for the models having strong indicators. However, in the literature [8,34] it is stated that both BSEM approach's power to detect cross-loading and its effectiveness in parameter estimations are quite sensitive to the variance of prior identified for cross-loading. In a study [8] it was examined the performance of BSEM approach (they called this approach as BSEM-RP) by identifying different priors with low variance for cross-loading. The researchers have reached the result that this approach performs better for low values of cross-loading when a prior with 0.001 variance is identified. However, in parallel with the results of the current study, it has been found that this approach performs better for high values of cross-loading when a prior with 0.01 variance is identified. When a variance with 0.01 prior is compared with a variance with 0.001 prior, it may cause that the posterior estimations related to cross-loading move far away from "0". Therefore, it is considered that the fact that for 0.1 value of cross-loading, the power of detecting cross-loading was found as lower may stem from the variance of prior identified in this study for cross-loading.

That the sensitivity analyses could not be carried out by identifying the priors including different variances for cross-loading becomes an important limitation of this study.

In light of these explanations and discussion, it is considered that RMSEA, which is used in Frequentist approach in the evaluation of the model fit for misspecified models (even if it leads to appropriate decisions in specifying "valid model" in this study), should not be used alone in the evaluation of the model fit since the contradictory results were obtained in this subject in the literature. It has been understood that in Bayesian approach, the evaluation of the model fit on the basis of ppp-values especially for misspecified models might lead the researchers to incorrect results. Within the context of rejecting the incorrect model, namely the detection of model misspecification, it has been found that Bayesian approach performed better in comparison with Frequentist approach. It has been seen that BSEM performed well especially for high level of cross-loading and in the conditions with multiple cross-loading. For that reason, especially within the context of omitted cross-loading, it is considered that it may be used effectively to avoid model misspecification.

In practice, a researcher using CFA in the Frequentist framework represents the underyling theory of the measures through imposing exclusion restrictions on Λ. But it is a rigid specicification and primarily, in general, it insinuates a loss of information in the sense of applying more exclusion restrictions than required for the model identification. Besides this specification, in which each item in the measurement instrument is related to just a certain dimension of the construct measured, may purport to omission errors and may lead to researchers to an unrealistic assumption that the items measure factorilly pure structure. In addition, such a specification may trigger the emergence of bias in the estimation of free parameters in CFA model. It is stated that in some way these matters could be related to widely-encountered case that the factor structures obtained through exploratory techniques are not confirmed by CFA (van Prooijen & van der Kloot, 2001). Accordingly it could be suggested to researchers and practitioners in the field of Behavioral Sciences to apply CFA based on BSEM approach along with EFA to examine the factor structure of the educational and psychological measures. As the priors chosen for the paramaters in CFA model affect the results in BSEM approach, one should be elaborately investigate related literature especialy with regard to the results of former exploratory factor analyses and expert opinions to specify priors for model parameters when applying CFA based on BSEM approach. By this way, it is believed that it could be avoided from making inferences and decisions based on misspecified model (as possible).

It is considered that the results of this research are important within the context of the evidences related to the validity of the measures since high-stake decisions are made on the basis of educational and psychological measures. In order to make correct decisions in educational settings on the basis of educational psychological measures, it is necessary to provide appropriate evidences on the psychometric qualities of these measures. In this context, model misspecification may lead the researchers to incorrect results in terms of the evidences on the validity of measures. Besides that, that the correct evidences on the validity of educational and psychological measures are obtained will support the use of these measures in view of educational accountability. In the literature [15, 26], it is pointed out that model misspecification will also have an impact on the results directly in testing invariance of measures between groups. Accordingly, it may lead the researchers to the incorrect inferences in the comparisons between the groups. Thus, when the measurement models are misspecified, further examinations, inferences and decisions based on these measures will be doubtful. For that reason, the effectiveness levels of Frequentist and Bayesian approaches were compared in this study, under more realistic and various conditions, in terms of the specification of the valid model related to educational and psychological measures and estimation of model parameters appropriately. It is considered that the results of this research will have the quality to guide the practitioners and researchers, who take various decisions on the basis of the educational and psychological measures, in the process of obtaining evidences to support these decisions.

# REFERENCES

[1] B. Byrne, Factor analytic models: Viewing the structure of an assessment instrument from three perspectives, J Pers Assess, Vol. 85, No.1, 17- 31, (2005).

[2] F. Lord, R. M. Novick, R, M, *S*tatistical Test Theories of Mental Test Scores. Addison Wesley Publishing Company, USA, 1968.

[3] J.C. Nunnally, I.H. Bernstein, Psychometric Theory (3rd ed.), McGraw-Hill, USA, 1994.

[4] K.A. Bollen, Structural Equations with Latent Variables, Wiley-Interscience Publication, USA, 1989.

[5] T. A. Brown, Confirmatory Factor Analysis for Applied Research, The Guilford Press, USA, 2006.

[6] R. Cudeck, Exploratory factor analysis. In H. E. A.Tinsley, S. D. Brown, (Eds.), Handbook of applied multivariate statistics and mathematical modeling (pp. 265-296), Academic Press, USA, 2000.

[7] G.W. Cheung, R. B. Rensvold, Evaluating goodness-of-fit indexes for testing measurement invariance, Structural Equation Modeling, 9, 233–255, (2002).

[8] Z. Lu, S. Chow, E. Loken. Bayesian factor analysis as a variable selection problem: Alternative priors and consequences, Multivariate Behav Res. Vol. 51, No.4,

519-539. (2016).

[9] K.G. Jöreskog, D. Sörbom, D. Lisrel 8: Structurel Equation Modelling with SIMPLIS™ Commanda Language, Scientific Software International, Inc, USA, 1993.

[10] R.B. Kline, Principals and Ppractice of structural equation modeling. The Guilford Press, USA, 1998.

[11] B. Muthén, T. Asparouhov. Bayesian Structural equation modeling: A more flexible representation of substantive theory, Psychological Methods, Vol.17, No.3, 313-335, (2012).

[12] R.C. MacCallum, Specification searches in covariance structure modeling, Psychological Bulletin, 100, 107-120, (1986).

[13] T. Asparouhov, B. Muthén, A.J.S. Morin, Bayesian structural equation modeling with cross-loadings and residual covariances: Comments on Stromeyer et al, Journal of Management, Vol.41, No.6, 1561-1577, (2015).

[14] L. Hu, P.M. Bentler, Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification, Psychological Methods, Vol. 3. No. 4,424-453, (1998).

[15] B.F. French, W.H. Finch. Model misspecification and invariance testing using confirmatory factor analytic procedures, The Journal of Experimental Education, 79, 404-428, (2011).

[16] V. Savalei, The relationship between root mean square error of approximation and model misspecification in confirmatory factor analysis models, Educational and Psychological Measurement, Vol.72, No.6, 910–932, (2012).

[17] W.E.Saris, A. Satorra, W.M Veld, Testing structural equation models or detection of misspecifications?, Structural Equation Modeling, 16, 561-582, (2009).

[18] J.R. Edwards, R.P. Bogazzi, On the nature and direction of relationships between constructs and measures, Psychological Methods, Vol. 5, No. 2, 155-174, (2000).

[19] R.C.McCallum, M.W. Browne, The use of causal indicators in covariance structure models: Some practical issues, Psychological Bulletin Vol.114, No.3, 533-41, (1993).

[20] Z. Lu, S. Chow & E. Loken. A comparison of bayesian and frequentist model selection methods for factor analysis models, Psychological Methods, Vol.22, No.2, 361-381, (2017).

[21] R. Cudeck, S. H.C. Du Toit, D. Sörbom, (Eds). Structural Equation Modeling: Present and Future, Scientific Software International, USA, 2001.

[22] J. Pan, E.H. Ip, L. Dube, An alternative to post hoc model modification in confirmatory factor analysis: The Bayesian Lasso, Psychological Methods, Vol.22, No.4, 687-704, (2017).

[23] L. Hu, P.M. Bentler, Evaluating Model Fit. In R. H. Hoyle (Ed.), Structural Equation Modeling: Concepts, Issues, and Applications (pp. 76-99). Sage Publications, Inc., USA, 1995.

[24] H. Hoijtink. (2008). Bayesian data analysis Online available:

http://www.researchgate.net/publication/46714374_Bayesian_data_analysis.

[25] D.A. Sass, J. Lopes, C. Oliveira, N. K. Martin, An evaluation of the behavior and instructional management scale's psychometric properties using Portuguese teachers, Teaching and Teacher Education, 55 279-290, (2016).

[26] D. Shi, H. Song, X. Liao, R.Terry, L.A. Snyder. Bayesian SEM for specification search problems in testing factorial invariance, Multivariate Behavioral Research, Vol. 52, No.4, 430-444, (2017).

[27] S.M. Lynch, Introduction to applied Bayesian statistics and estimation for social sciences, Springer Science+Business Media LLC, USA, 2007.

[28] A. Gelman, D.B. Rubin, Inference from iterative simulation using multiple sequences, Statistical Science, 7, 457–472, (1992).

[29] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Rubin, Bayesian Data Analysis (2nd ed.), CRC Press, USA, 2004.

[30] R. Levy, R.J. Mislevy, Bayesian Psychometric Modeling, CRC Press Taylor & Francis Group, USA, 2016.

[31] P. Golay, I. Reverte, J. Rossier, N. Favez, T. Lecerf, Further insights on the French WISC-IV factor structure through Bayesian structural equation modeling. Psychological Assessment. Advance online publication (2012).

[32] M. Rhemtulla, Population performance of SEM parceling strategies under measurement and structural model misspecification, Psychological Methods,Vol. 21, No. 3, 348–368, (2016).

[33] H. Hoofs, R. Schoot, N.W.H. Jansen, I. Kant. Evaluating model fit in bayesian confirmatory factor analysis with large samples: Simulation study introducing BRMSEA, Educational and Psychological Measurement, Vol.78, No. 4, 537–568, (2018).

[34] B. Muthén, T. Asparouhov. (2010). Bayesian SEM: A more flexible representation of substantive theory. Online available: http://www.statmodel.com/download/BSEMv4.pdf.

[35] R. B. Cattell, The Scientific Use of Factor Analysis, Plenum Press, USA, 1978.

[36] Tabachnick, B. G. and Fidell, L. (6th ed.), Using multivariate statistics. Pearson, USA, 2014.

[37] A.Beauducel, W.W. Wittman, Simulation Study on Fit Indexes in CFA Based on Data With Slightly Distorted Simple Structure Structural Equation Modeling A Multidisciplinary Journal, Vo.12, No.1, 41-75, (2005).

[38] E. Breivik, U. H. Olsson, Adding Variables to Improve Model Fit: The Effect of Model Size on Fit Assessment in LISREL. In R. Cudeck, S. Du Toit, & D. Sörbom (Eds.), Structural Equation Modeling: Present and Future (pp. 169-194), Scientific Software International, USA, 2001.

[39] van Prooijen, J. W., & van der Kloot, W. A. (2001). Confirrmatory Analysis of Exploratively Obtained Factor Structures. Educational and Psychological Measurement, 61, 777-792.