Article

Collinear Latent Variables in Multilevel Confirmatory Factor Analysis: A Comparison of Maximum Likelihood and Bayesian Estimations

**Seda Can[1], Rens van de Schoot[2], and
Joop Hox[2]**

Abstract

Because variables may be correlated in the social and behavioral sciences, multicollinearity might be problematic. This study investigates the effect of collinearity manipulated in within and between levels of a two-level confirmatory factor analysis by Monte Carlo simulation. Furthermore, the influence of the size of the intraclass correlation coefficient (ICC) and estimation method; maximum likelihood estimation with robust chi-squares and standard errors and Bayesian estimation, on the convergence rate are investigated. The other variables of interest were rate of inadmissible solutions and the relative parameter and standard error bias on the between level. The results showed that inadmissible solutions were obtained when there was between level collinearity and the estimation method was maximum likelihood. In the within level multicollinearity condition, all of the solutions were admissible but the bias values were higher compared with the between level collinearity condition. Bayesian estimation appeared to be robust in obtaining admissible parameters but the relative bias was higher than for maximum likelihood estimation. Finally, as expected, high ICC produced less biased results compared to medium ICC conditions.

[1]İzmir University of Economics, İzmir, Turkey
[2]Utrecht University, Utrecht, Netherlands

Corresponding Author:
Seda Can, İzmir University of Economics, Sakarya cad No:156, 35330 Balçova, Izmir, Turkey.
Email: seda.can@ieu.edu.tr

Keywords

Multicollinearity has been an important issue both in multiple regression and generalized linear models and in structural equation modeling (SEM) or multilevel structural equation modeling (MSEM). While multicollinearity has been studied extensively in multiple regression and generalized linear models (cf., Aiken & West, 1991; Field, 2013; Stevens, 2009; Tabachnick & Fidell, 2006), it has been studied less in the general SEM framework (for exceptions, see Grewal, Cote, & Baumgartner, 2004; Jagpal, 1982; Kaplan, 1994; Marsh, Dowson, Pietsch, & Walker, 2004). It has received some attention in multilevel regression modeling (i.e., Kreft & de Leeuw, 1998; Kubitschek & Hallinan, 1999; Shieh & Fouladi, 2003) but, has to our knowledge, never been studied in MSEM.

From the 1960s, multicollinearity has been an issue of interest to researchers studying multiple regression, especially in economics and statistics (e.g., Fabrycy, 1975; Farrar & Glauber, 1967; Gordon, 1968; Harvey, 1977; Johnston, 1963; Malinvaud, 1966; Pedhazur, 1982). Because independent variables are usually correlated in the social and behavioral sciences, multicollinearity has also become a field of interest to researchers who use multiple regression in the analysis of social science data (e.g., Morrow-Howell, 1994). The terms multicollinearity, collinearity and 'ill conditioning' are all used interchangeably, and without a firmly established definition in the literature (Belsley, 1991). Multicollinearity refers to a situation consisting of linearly or near linearly related predictors, independent or explanatory variables. In this case the correlation matrix among the set of predictors is ill conditioned, resulting in at least one eigenvalue that is nonpositive (Shieh & Fouladi, 2003). In such cases, there is no unique mathematical solution for estimating the regression coefficients.

The consequences of multicollinearity can be discussed in terms of standard errors and the regression coefficients. First, when multicollinearity exists, the standard errors for the coefficients become large and this increases the probability of Type II error (Cohen, Cohen, West, & Aiken, 2003; Mason & Perrault, 1991; Pedhazur, 1997). As a result, the estimates with large standard errors are more likely to vary between samples causing unstable estimates across samples (Hays, 1981; Neter, Wasserman, & Whitmore, 1978). Second, regression coefficients may decrease in magnitude or change in sign, resulting in interpretation problems (Cohen & Cohen, 1983; Kleinbaum, Kupper, & Muller, 1988). Pedhazur (1997) showed that the parameters estimated from two samples are quite different if the predictors are highly correlated.

The current study uses Monte Carlo simulation to investigate the effect of multicollinearity manipulated in the within and between levels of a two-level confirmatory factor model, on the convergence rate, the rate of inadmissible solutions, relative

parameter and standard error bias. In addition to the magnitude of multicollinearity, the effect of the magnitude of the intraclass correlation coefficient (ICC) and the estimation method; maximum likelihood (ML) estimation versus Bayesian estimation, are investigated.

In what follows, we first present an overview of previous studies that investigated multicollinearity in multilevel regression and SEM frameworks. Then we end the introduction section with a comparison between ML and Bayesian estimation.

Multilevel Regression, Structural Equation Modeling, and Multicollinearity

Few studies have investigated multicollinearity in the context of multilevel modeling. Kreft and de Leeuw (1998) and Kubitschek and Hallinan (1999) used National Educational Longitudinal Survey data to demonstrate the effect of collinearity. Kreft and de Leeuw developed different multilevel models by using different predictor variables in the models. The results showed that multicollinearity can make the interpretation of the regression coefficients difficult especially when cross-level interaction exists. Kubitschek and Hallinan used the same data set and found that owing to the collinearity, parameter estimates had large standard errors and the effect of the predictor variable varied strongly from sample to sample. Shieh and Fouladi (2003) investigated the effect of multicollinearity, operationalized as the correlation between Level 1 predictors, on multilevel model parameters and standard errors when cross-level interactions exist in the data by using Monte Carlo simulation. Sample size and intraclass correlation were manipulated in the research design as well. Sample size was varied by both number of groups and the number of cases in each group. The number of groups was 10, 20, and 40 and the number of cases per group was also 10, 20, and 40. The second independent variable ICC was manipulated in three levels including .25, .50, and .75. Finally, the varied Level 1 predictor correlations were .0, .10, .30, .50, .70, and .90. The results for the convergence rate showed that it improved as the number of groups increased, as the number of cases in each group increased, as the ICC is decreased and as the correlation between Level 1 predictors decreased. Their study showed that the fixed-effect parameter estimates produced relatively unbiased estimates. However, the variance and covariance component estimates produced negatively biased values (except for Level 1 variance) and the relative bias of the standard errors of the parameters increased when the correlation between the Level 1 predictors increased.

The effect of multicollinearity in the SEM framework is unclear. Measurement error can be eliminated in structural models, and the subsequent increase of the explained variance may lessen the effect of multicollinearity (Bollen, 1989). On the other hand, the elimination of error may also increase the size of the estimated correlations among latent variables (Grewal et al., 2004). To illustrate the interpretation problems when multicollinearity exists in a data, a study was conducted by Marsh et al. (2004). They showed that constraining paths to be equal in a model with collinear latent constructs made the fit better and reduced the standard errors.

Two of the few studies investigating multicollinearity in the SEM framework, especially on its detection, were carried out by Schmidt and Muller (1978) and Kaplan (1994). In addition to the tools suggested to detect multicollinearity in multiple regression, Kaplan proposed inspecting the determinant of the correlation matrix of the predictor variables to evaluate the multicollinearity problem. Schmidt and Muller (1978) suggest using the Haitovsky test (Haitovsky, 1969) to assess the amount of singularity of the correlation matrix of the predictors, together with determinant of correlation matrices in detecting multicollinearity.

To investigate to what degree multicollinearity may be problematic in a structural equation model; two Monte Carlo simulation studies were conducted by Grewal et al. (2004). They manipulated the level of multicollinearity, measurement error, amount of explained variance, relative importance of exogenous variables and sample size. The accuracy of coefficient estimates and standard error estimates were studied in addition to studying under which conditions multicollinearity and measurement error in the structural model caused misleading tests of theory. The results showed that in extreme multicollinearity, Type II error rates were more than 80%, which is unacceptably high. When multicollinearity was between 0.60 and 0.80 Type II error rates were greater than 50% and frequently more than 80% if composite reliability was weak, explained variance was low and sample size was relatively small. Moreover when multicollinearity was between 0.40 and 0.50, Type II error rates were small unless the reliability was weak, R^2 was low and sample size was small but the error rates were still high.

To sum up, both in multilevel regression and SEM framework, when multicollinearity exists, undesirable results may occur. For example, the signs of the regression parameters might be different from the expected direction, and the standard errors of the regression parameters might become extremely large which makes the parameter estimates statistically nonsignificant and results in high Type II error rates. These effects have not yet been studied in an MSEM framework and this is exactly what we will do in the current study. The purpose of this study to examine how multicollinearity, which is operationally defined as the correlation between the factors in a multilevel confirmatory factor analysis model, can affect the parameter estimates obtained from the model. We expect a more severe bias caused by multicollinearity on the between level of the model because of the smaller sample size on this level when compared with the within level.

## Maximum Likelihood and Bayesian Estimation

ML estimation is generally being used as the standard estimation method for the parameters of statistical models. For the ML equations and their implementation in the MSEM framework we refer to B. O. Muthén (1990) and Mehta and Neale (2005). We would like to briefly discuss ML with robust chi-squares and standard errors, which is denoted as MLR. MLR produces the same parameter estimates as ML, but the chi-squares for the model test and the standard errors for the parameters are

calculated differently (L. K. Muthén & Muthén, 2014). MLR is assumed to be robust against moderate violations of assumptions, including unmodeled heterogeneity. With multilevel data, robust chi-squares and standard errors are assumed to provide some protection against mis-specifying the group level model or by omitting a variable resulting in unmodeled heterogeneity (Hox, 2010). Owing to these mentioned advantages of MLR for multilevel data, in the current study we will use MLR instead of ML estimation.

Bayesian estimation has started to be used in structural equation modeling of hierarchical data (e.g., Jedidi & Ansari, 2001). In the Bayesian approach, posterior distributions for the parameters are acquired from the prior distribution and from the likelihood of the data. For an introduction to Bayesian estimation see Gelman, Carlin, Stern, and Rubin (2004). Bayesian estimation offers several potential advantages over MLR estimation. First, Bayesian methods require specifying prior information about all the parameters in the model. By a judicious choice of the prior distribution, we can ensure that parameter estimates are confined to the permissible parameter space. That is, if the priors for (residual) variance terms are set in such a way that they can only obtain positive values, the obtained posterior distribution can never be negative. For this reason, inadmissible values of residual variance terms cannot occur. If a different prior would be specified, for example, normal priors, negative values would be possible. As said before, the crucial part of a Bayesian analysis is to choose the prior distributions wisely. Furthermore, Bayesian procedures do not rely on asymptotic inference, and as such are valid for small sample sizes. Third, central credibility intervals (CCIs), the Bayesian equivalent to a confidence interval (CI), are actually the probability that a certain parameter is in between two numbers, which is not the definition of a confidence interval. For detailed information about these advantages, see also Howard, Maxwell, and Fleming (2000), Lee and Wagenmakers (2005), van de Schoot et al. (2011), and Walker, Gustafson, and Hennig (2001).

Bayesian estimation uses Markov chain Monte Carlo (MCMC) algorithms to create approximations to the posterior distributions by iteratively taking random draws in the MCMC (Gelman et al., 2004; L. K. Muthén & Muthén, 2014). We refer to Lynch (2007) for an introduction to the Bayesian approach, and for more technical details in Bayesian estimation to Gelman et al. (2004). For Bayesian SEM, see Jedidi and Ansari (2001) and Lee (2007). In our simulation, we compare the MLR results to the results obtained with Bayesian estimation as a possible solution to multicollinearity problems because it provides a convenient way to avoid estimation problems due to inadmissible estimates. Also, Bayesian estimation does not rely on asymptotic inference, and as such we expect less biased results on the between level.

In conclusion, we hypothesize that multicollinearity at the between level will cause more problems such as uninterpretable or unexpected results (e.g., high standard error values) when compared with multicollinearity manipulated at the within level. The main reason is that the sample sizes tend to be larger at the within level than at the between level, while correlations tend to be higher at the between level
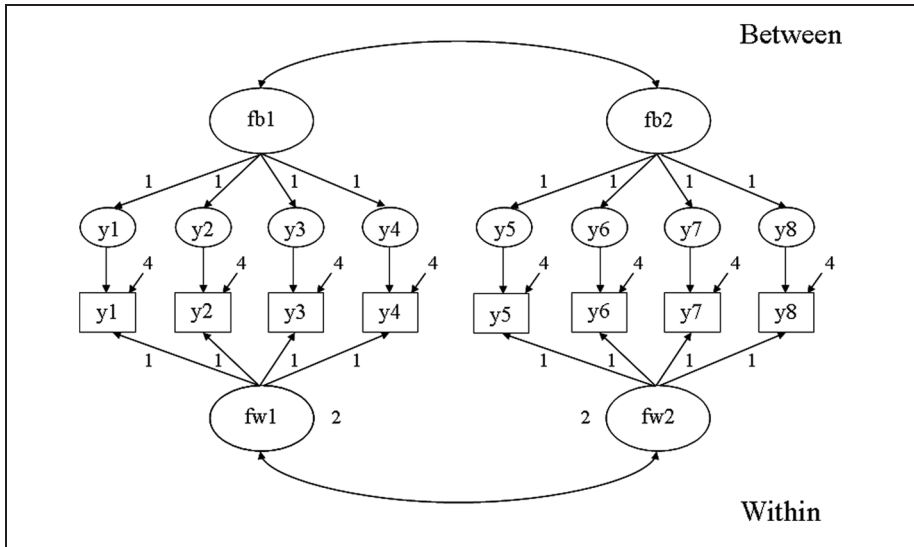
Figure 1. The path diagram of multilevel confirmatory factor analysis model tested in the study.

than at the within level. In addition to the multicollinearity, the effect of the size of the ICC and the estimation method is also investigated.

Method

The steps suggested by Skrondal (2000) and Paxton, Curran, Bollen, Kirby, and Fen (2001) are used in the current Monte Carlo simulation study.

Simulated Model

A two-factor model with four indicators, two factors both in the within and between levels is used in this study. We use a confirmatory factor model to represent a research problem often used in practice. In determining the population parameter values, especially the ICC conditions, we benefited from the values that Julian (2001) used in his study. ICC levels are manipulated by changing the residual variances and factor variances of the between level, while keeping the within model the same. The path diagram of the model tested in the study and common population parameters used in each simulation are shown in Figure 1 and the other parameter values for each simulation are presented in Tables 1 and 2 separately for different values of within- and between-level collinearity. For within-level collinearity simulations, within-level covariances, between-level residual variances, between-level factor variances, and between-level covariance values are presented in terms of ICC conditions

Table I. Population Parameter Values Used in Within-Level Multicollinearity Simulations.

Within-level multicollinearity		Intraclass correlation coefficient	Within-level covariance	Between-level residual variances	Between-level factor variances	Between-level covariance
Low (.30)		Medium (.15)	0.65	0.71	0.35	0.130
		High (.25)	0.65	1.00	0.90	0.295
Medium (.50)		Medium (.15)	1.04	0.71	0.35	0.130
		High (.25)	1.04	1.00	0.90	0.295
High	.80	Medium (.15)	1.64	0.71	0.35	0.130
		High (.25)	1.64	1.00	0.90	0.295
	.90	Medium (.15)	1.84	0.71	0.35	0.130
		High (.25)	1.84	1.00	0.90	0.295
	.95	Medium (.15)	1.88	0.71	0.35	0.130
		High (.25)	1.88	1.00	0.90	0.295

Table 2. Population Parameter Values Used in Between-Level Multicollinearity Simulations.

Between-level multicollinearity		Intraclass correlation coefficient	Between-level residual variances	Between-level factor variances	Between-level covariance
Low (.30)		Medium (.15)	0.71	0.35	0.130
		High (.25)	1.00	0.90	0.300
Medium (.50)		Medium (.15)	0.71	0.35	0.200
		High (.25)	1.00	0.90	0.500
High	.80	Medium (.15)	0.71	0.35	0.310
		High (.25)	1.00	0.90	0.760
	.90	Medium (.15)	0.71	0.35	0.342
		High (.25)	1.00	0.90	0.850
	.95	Medium (.15)	0.71	0.35	0.345
		High (.25)	1.00	0.90	0.890

Note. The population parameter value for within level covariance is 0.65 in all between level multicollinearity simulations.

in Table 1. Between-level residual variances, between-level factor variances, and between-level covariance values are presented in Table 2 for simulating between-level multicollinearity.

A sample size about 200 cases is accepted as reasonable with a good model and normal data in single level SEM framework in order to reach accurate maximum likelihood estimates (Boomsma, 1983). This finding was also supported by Hox and Maas (2001) in an MSEM simulation study. To ensure that potential estimation problems are not the result of insufficient sample sizes, we used 200 groups in the between level with 25 individuals nested in each group.

Design Factors

Three variables with 5 (multicollinearity) $\times$ 2 (ICC) $\times$ 2 (estimation method) $\times$ 2 (level on which the multicollinearity is manipulated) = 40 conditions are used in the simulation. Multicollinearity is manipulated in the within and between levels of the model separately with low, medium, and high multicollinearity conditions. To cover high levels of correlations between the factors in the simulations we investigated three sublevels: high, extremely high, and almost perfect correlation in the high multicollinearity condition. In total, the population correlation values between the two factors in the model are .30, .50, .80, .90, and .95. The ICC is manipulated with medium (.15) and high (.25) conditions. And finally, type of estimation is used as the third variable consisting of MLR and Bayesian estimation.

Mplus Version 7.11 (L. K. Muthén & Muthén, 2013) was used to carry out the Monte Carlo simulations. The simulation syntax is given in Appendix A. For MLR estimation, we relied on the Mplus defaults, which include a robust chi-square (Yuan–Bentler correction) and robust standard errors (sandwich estimators).

Bayesian estimation became available in Mplus since version 6.11 (L. K. Muthén & Muthén, 2011). For a technical description of Bayesian estimation in Mplus, see Asparouhov and Muthén (2010a, 2010b). We used the default Mplus settings in Bayesian estimation. The prior specifications are by default set to prior distributions that are noninformative, and constrain the parameter values to admissible values.[1] From the posterior distribution a point estimate of mean, median or mode is provided, we used the median which is the default. To monitor the convergence of the Bayesian estimation procedure, Mplus uses the Gelman–Rubin convergence criterion which considers the variability within and between different estimation chains (Gelman et al., 2004, pp. 296-298). This process proceeds by default with two chains of the Gibbs sampler in Mplus running in a parallel process.

Dependent Variables

The convergence rate is the first dependent variable in the study, and it is computed as the rate of completed number of replications over the requested number of replications. In addition, we present the percentage of the number of inadmissible solutions, computed by dividing the number of inadmissible solutions by the number of replications. To compare the degree of bias in the estimation of parameter values (between-level factor loadings and between-level covariance) of different magnitude, we use the percentage relative bias. The percentage relative bias is computed as follows:

$$RelativeBias = ((\hat{\theta} - \theta)/\theta) \times 100 \tag{1}$$

where $\hat{\theta}$ equals the mean of the parameter estimates across the replications and θ equals the population parameter value. Since there are six estimated factor loadings (for each factor one is used for scaling) we calculated the mean relative bias across the between level factor loadings. The percentage relative bias for the standard errors

are also computed with Equation 1; but $\hat{\theta}$ is the mean of the standard error estimates of the corresponding parameter estimates obtained from the replications and θ is the population standard deviation in this case. When the number of replication is large, the standard deviation obtained from the simulations is considered to be a good estimate of the population standard error (L. K. Muthén & Muthén, 2010). These bias values are also averaged to obtain the mean relative bias for the standard error estimates. To assess whether the estimated parameters and their standard errors are biased, a $\pm 10\%$ criterion value is used, which is suggested by Hoogland and Boomsma (1998).

For each cell of the research design, 1000 normally distributed data sets are generated and analyzed with Mplus 7.11 (L. K. Muthén & Muthén, 1998-2013). We have included Appendix A consisting of the Mplus syntax for 0.80 between level collinearity, medium ICC condition with MLR and Bayesian estimation. Other conditions can be specified by changing the population parameter values presented in Tables 1 and 2.

Results

Convergence Rate and Inadmissible Solutions

There were no nonconvergent solutions across all simulated data sets. All the solutions are admissible for Bayesian estimation. However, inadmissible solutions do occur in MLR estimation when the multicollinearity is manipulated at the between level. As can be seen in Figure 2, there were no inadmissible solutions for the low and medium collinearity in both ICC conditions. On the other hand, the percentages of inadmissible solutions increase as the multicollinearity increases for the high collinearity conditions. When the ICC is medium, the percentages of inadmissible solutions are 12%, 42%, and 52%, respectively, for the high, extremely high, and almost perfect collinearity. In the high ICC condition, these percentages go down to 0.2%, 11.7%, and 42.4%, respectively, as the multicollinearity increases. The results in the next section are based only on admissible solutions; however, the estimates including the inadmissible solutions lead to highly similar results.

Parameter Estimates

In the within part of the model (where the effective sample size is the largest), even high degrees of multicollinearity do not present a problem in any of the simulated conditions. Therefore, we concentrate on a discussion of the results at the between level, where multicollinearity manipulated at either the within or the between level does cause problems.

The mean relative parameter bias of between-level factor loadings and relative bias for the covariance between the latent variables obtained through MLR estimation when the multicollinearity is manipulated in the within level of the model are presented in Table 3. The relative bias values of factor loadings are nearly the same with small differences in the third decimal places for each multicollinearity condition in
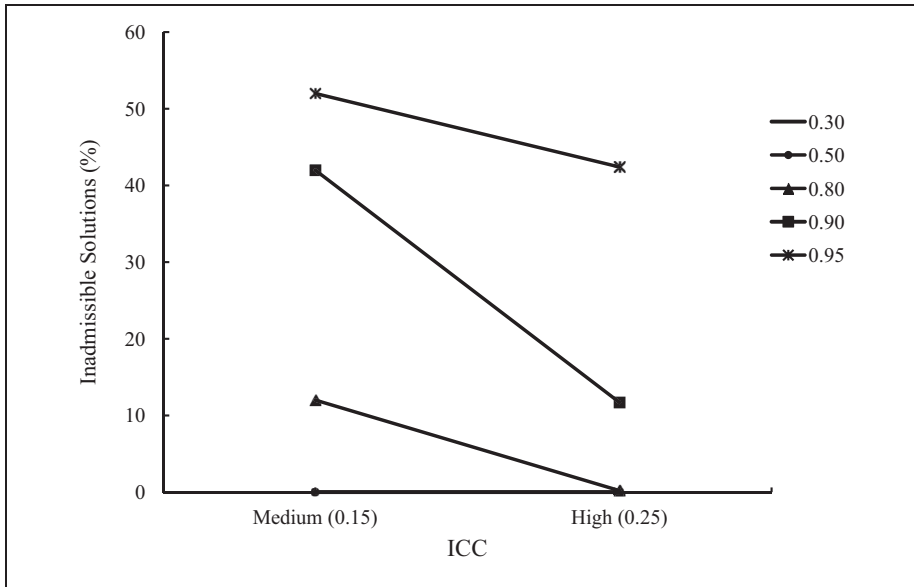
Figure 2. Percentages of inadmissible solutions in between level multicollinearity conditions across intraclass correlations (ICCs).

terms of ICC levels. For medium ICC, the values are nearly 2.95% and for high ICC conditions, the mean relative bias values for factor loadings are about 0.64% regardless of the multicollinearity levels. The relative bias percentages for the covariance between the latent variables are lower than the bias for the factor loadings and the values are both negative showing underestimated covariances. The relative bias percentages of the covariance are between −0.20% and −0.46%, with small differences across the conditions. The values for the low and medium collinearity conditions are −0.31% and −0.39%, for the high collinearity conditions, the values are same for medium ICC condition, which is −0.46%. For high ICC condition, the percentages are lower with respect to other multicollinearity conditions with very small differences in low and medium multicollinearity conditions. The relative bias values for low and medium multicollinearity conditions are −0.20% and −0.24%; for the 0.80, 0.90, and 0.95 collinearity conditions relative bias values are found to be −0.31%, −0.34%, and −0.34%, respectively.

The Bayesian results when there is within-level multicollinearity in the model are shown in Table 4. It can be seen that the percentages of the mean relative bias for the factor loadings are higher than the MLR results. The percentages are, respectively, 8.38%, 8.01%, 8.00%, 8.55%, and 9.28% for low, medium, and high collinearity conditions in medium ICC. In high ICC condition, the corresponding values are 2.39%, 2.31%, 2.37%, 2.39%, and 2.52%. When we look at the bias for the covariance, it can be said that the covariances are negatively biased in medium ICC conditions and

Table 3. MLR Results of Between-Level Parameters From Within-Level Multicollinearity Simulations.

Multicollinearity		Intraclass correlation coefficient	Mean bias λ	Mean bias SE λ	Bias covariance Ψ_{12}	SE covariance Ψ_{12}
Low (.30)		Medium (.15)	2.93	0.10	−0.31	2.47
		High (.25)	0.65	−0.46	−0.20	0.79
Medium (.50)		Medium (.15)	2.94	0.23	−0.39	2.77
		High (.25)	0.65	−0.35	−0.24	0.59
High	.80	Medium (.15)	2.96	0.55	−0.46	2.26
		High (.25)	0.64	−0.17	−0.31	0.49
	.90	Medium (.15)	2.95	0.71	−0.46	2.25
		High (.25)	0.64	−0.05	−0.34	0.59
	.95	Medium (.15)	2.95	0.75	−0.54	2.25
		High (.25)	0.64	−0.03	−0.34	0.59

Note. MLR = maximum likelihood with robust chi-squares and standard errors.

Table 4. Bayesian Results of Between-Level Parameters From Within-Level Multicollinearity Simulations.

Multicollinearity		Intraclass correlation coefficient	Mean bias λ	Mean bias SE λ	Bias covariance Ψ_{12}	SE covariance Ψ_{12}
Low (.30)		Medium (.15)	8.38	−7.91	−13.39	−3.30
		High (.25)	2.39	−7.78	−3.39	0.69
Medium (.50)		Medium (.15)	8.01	−2.94	−13.69	−4.15
		High (.25)	2.31	−2.03	−3.36	0.79
High	.80	Medium (.15)	8.00	−2.28	−13.92	−4.46
		High (.25)	2.37	−2.43	−3.59	−0.39
	.90	Medium (.15)	8.55	−9.39	−14.69	−3.24
		High (.25)	2.39	−2.07	−3.93	0.68
	.95	Medium (.15)	9.28	−11.48	−15.31	−3.62
		High (.25)	2.52	−2.54	−4.17	1.28

the percentages of the bias increase as the multicollinearity increases. The values in this ICC condition are −13.39%, −13.69%, −13.92%, −14.69%, and −15.31-% for low, medium, and 0.80, 0.90, and 0.95 multicollinearity conditions showing a large bias exceeding the 10% criterion suggested by Hoogland and Boomsma (1998) for "reasonable" accuracy. In the high ICC condition, the percentages are very close to each other with the values of −3.39%, −3.39%, −3.59%, −3.93%, and −3.53%.

When multicollinearity is manipulated in the between level, the mean of between-level factor loadings from MLR estimation have a relative bias of 2.93%, 2.64%, 2.00%, 1.79%, and 1.77% for medium ICC and 0.65%, 0.63%, 0.55%, 0.51%, and

Table 5. MLR Results of Between-Level Parameters From Between-Level Multicollinearity Simulations.

Multicollinearity		Intraclass correlation coefficient	Mean bias λ	Mean bias SE λ	Bias covariance Ψ_{12}	SE covariance Ψ_{12}
Low (.30)		Medium (.15)	2.93	0.09	−0.31	2.47
		High (.25)	0.65	−0.46	−0.31	0.79
Medium (.50)		Medium (.15)	2.64	0.73	−0.40	2.34
		High (.25)	0.63	−0.26	−0.18	0.86
High	.80	Medium (.15)	2.00	0.89	−0.52	1.84
		High (.25)	0.55	0.04	−0.26	0.87
	.90	Medium (.15)	1.79	0.63	−0.53	1.77
		High (.25)	0.51	−0.03	−0.29	0.76
	.95	Medium (.15)	1.77	0.59	−0.52	1.89
		High (.25)	0.49	−0.12	−0.30	0.61

Note. MLR = maximum likelihood with robust chi-squares and standard errors.

Table 6. Bayesian Results of Between-Level Parameters From Between-Level Multicollinearity Simulations.

Multicollinearity		Intraclass correlation coefficient	Mean bias λ	Mean bias SE λ	Bias covariance Ψ_{12}	SE covariance Ψ_{12}
Low (.30)		Medium (.15)	8.38	−7.91	−13.39	−3.30
		High (.25)	2.39	−2.87	−3.37	−0.39
Medium (.50)		Medium (.15)	6.79	0.12	−11.2	−5.40
		High (.25)	2.18	−1.68	−2.82	−1.34
High	.80	Medium (.15)	5.33	−3.51	−8.94	−2.35
		High (.25)	1.93	−1.96	−1.84	0.07
	.90	Medium (.15)	5.00	−5.38	−9.21	−3.15
		High (.25)	1.88	−4.34	−2.00	1.23
	.95	Medium (.15)	4.77	−4.93	−9.33	−1.65
		High (.25)	1.66	−4.26	−2.28	3.50

0.49% for high ICC condition (see Table 5). Although these values are negligible, the values do decrease as the multicollinearity increase in each ICC condition. The relative bias values for the covariance between the factors are small and negative with the values of −0.31%, −0.40%, −0.52%, −0.53%, and −0.52% for medium ICC and −0.31%, −0.18%, −0.26%, −0.29%, and −0.30% for high ICC in terms of low, medium, and high multicollinearity conditions.

The Bayesian results presented in Table 6 show that the relative bias values for the factor loadings in 0.30, 0.50, 0.80, 0.90, and 0.95 collinearity conditions are 8.38%,

6.79%, 5.33%, 5.00%, and 4.77% in the medium ICC condition and 2.39%, 2.18%, 1.93%, 1.88%, and 1.66% in the high ICC condition. The relative bias for the covariance are negative and biased for some of the conditions when the collinearity is manipulated in the between level. They are underestimated and biased for the low and medium multicollinearity conditions with the values of −13.39% and −11.20%, for the other collinearity conditions the values are very close to critical value with the values of −8.94%, −9.21%, and −9.33% in medium ICC showing an increase as the collinearity increases and they drop to −3.37%, −2.82%, −1.84%, −2.00%, and −2.28% in the high ICC condition.

Standard Error Estimates

The between level standard error bias values of the factor loadings for the within-level collinearity produced from MLR estimation are very small changing from −0.46% to 0.75%. As can be seen in Table 3, the values are slightly lower in the high ICC than in the medium ICC condition. When the relative standard errors for the covariance between the factors are investigated, they are very close to each other ranging from 2.25% to 2.77% in the medium ICC condition. In the high ICC condition, the relative bias values of standard errors of the covariances for the 0.30, 0.50, 0.80, 0.90, and 0.95 collinearity conditions are 0.79%, 0.59%, 0.49%, 0.59%, and 0.49 with small changes in the third decimal places that makes these values appear the same in the table. The percentages obtained through Bayesian estimation are presented in Table 4. Most of the standard error bias values of factor loadings are higher than the ones obtained from MLR estimation and they are negatively biased. The values for the low and medium collinearity conditions are −7.91% and −2.94% in medium ICC. For the 0.80 collinearity condition, the bias value is found to be −2.28%, and in the 0.90 collinearity condition, the value is very close to the critical value with the value of 9.39%. And in the 0.95 condition, there is found to be a negatively biased value, −11.48%, in the same ICC condition. Moreover, all the values in the high ICC condition are underestimated as in medium ICC but the values are smaller in comparison. When we look at the standard error bias for the covariance in detail, the values are negative with values of −3.33%, −4.15%, −4.46%, −3.24%, and −3.62%. In high ICC, there are small positive biases in all collinearity conditions.

The relative standard error bias of factor loadings are all less than 1% in each cell of the research design (Table 5) when the multicollinearity is manipulated in the between level of the model with MLR estimation. The same holds for the standard errors of the covariance except the medium ICC condition. In medium ICC, the values are 2.47%, 2.34%, 1.83%, 1.77%, and 1.89%, respectively, for 0.30, 0.50, 0.80, 0.90, and 0.95 multicollinearity conditions.

Bayesian estimation underestimates the standard errors for the factor loadings with values ranging from −7.91% to −1.68% with one exception in the medium ICC and medium multicollinearity condition, which is 0.12%, as shown in Table 6. The values are also underestimated for the covariance between the latent variables in medium

ICC condition. The bias values of the covariance are underestimated for each collinearity condition in medium ICC. On the other hand, the values are found to be overestimated in high ICC.

Although we focus on the effect of multicollinearity on the between-level parameters in the current study, we also examined the within-level part of the model. The within-level parameters obtained are shown in Appendix B (Tables B.1-B.4). As can be seen in the tables provided, there were no biased values but especially for the standard errors Bayesian estimation produced negative bias values around −5.00% both for within- and between-level collinearity conditions in medium and high ICC.

Discussion

The purpose of this study was to investigate the effect of multicollinearity on multilevel confirmatory factor analysis and to compare ML estimation and Bayesian estimation when there are collinear latent variables in the within or between level. The results of the current study suggests that when there was multicollinearity in the between level of the data, the number of inadmissible solutions increased as the multicollinearity increases for MLR estimation. Grewal et al. (2004), using SEM, also found an increase in the number of inadmissible solutions with higher levels of multicollinearity. There were no inadmissible solutions for the low and medium collinearity in both ICC conditions. Additionally, we found an interaction effect of multicollinearity and ICC on the percentage of the inadmissible solutions for the high levels of multicollinearity. In the high ICC condition in which there is more explained between-level variance than in the medium ICC condition, the number of inadmissible solutions was lower in high conditions of multicollinearity. There was no problem in obtaining admissible solutions when there was within-level multicollinearity in the model. On the other hand, the relative bias values were higher compared with between-level collinearity conditions. We expected a more severe bias caused by multicollinearity on the between-level model because of the smaller sample size on this level when compared with the within-level model. Nevertheless, we also found higher bias values in the between level when collinearity was manipulated in the within level. This may be the result of the way how the between-groups structure is estimated. The between level structure is obtained both from the within-level and group-level covariance matrices. Thus, the manipulation that we did in the within level might have produced higher bias values in the between level.

The results of the current study also indicate that all the parameters were admissible in all replications using Bayesian estimation. In Mplus (also in MLWIN and AMOS) the prior distributions are set to be noninformative by default but they reflect the admissible parameter space. This is done by using an inverse chi-square or a gamma distribution for variance terms such that these can only be positive. Because of this default setting, the variance terms are all larger than zero in Bayesian estimation. However, in MLR estimation, negative variance or residual variance values or correlations larger than ± 1 may be obtained, as we found in our estimation. On the

whole, Bayesian estimation was robust in obtaining admissible parameters but not in estimating parameter estimates of the model. In general, the mean relative bias values in estimates obtained by Bayesian estimation were higher than the ones obtained by MLR estimation. Therefore, it can be concluded that our hypothesis—Bayesian estimation is more robust compared with MLR estimation—was only partially supported in the current study.

In MLR estimation, inadmissible solutions for the parameters can be an indication of multicollinearity. Although there were relative bias values larger than the $\pm 10\%$ criterion (obtained especially for the covariances), we conclude that Bayesian estimation solves the issue of obtaining always admissible parameters at the cost of obtaining parameter estimates with higher relative bias values when compared with MLR estimation. On the basis of our results, we recommend researchers switching to the Bayesian estimation if you are interested in the structural part of the model (the relationships between the latent variables). But we also recommend caution in relying on estimates obtained from Bayesian estimation even if low and medium collinearities exist in models, because this affects other parameters too, as was demonstrated with the higher relative bias of the factor loadings.

It is important to remember that we defined the collinearity as the correlation between the latent factors in a confirmatory factor analysis model. High levels of correlation between the factors may also be viewed as indicating a misspecification of the model. Maybe the best solution for the multicollinearity in our model would be to collapse both latent factors on the between level. On the other hand, such highly correlated models may be retained in the model because of the corresponding constructs are viewed as conceptually different as long as there is not a correlation of 1 between them. Either you have misspecification or have highly collinear factors in your model, the results of this study may be crucial for these two cases. When MLR is used in these type of data, the inadmissible solutions warn researchers for model misspecification, but Bayesian does not. And if you could obtain admissible solutions our results also showed that the parameters and the standard errors may not be biased in MLR estimation. Additionally, the researchers always have to be aware of model misspecification more when they are using Bayesian estimation.

Another issue is that the results of our study based on a relatively large sample size in the group level. We intentionally used a high sample size to analyze only the effect of multicollinearity without having power or estimation issues. Future research should manipulate the sample size by changing both the number of groups and the number of cases in groups to investigate the interaction of power issues with multicollinearity.

While our study takes a first step toward investigating the effect of multicollinearity in multilevel confirmatory factor analysis, this effect can also be investigated in the structural modeling of hierarchical data consisting of more complex models.

Appendix A

Simulation Commands

A. MLR syntax

MONTECARLO:

names are y1-y8;
nobservations = 5000;
ncsizes = 1;
csizes = 200 (25);
seed = 50277;
nreps = 1000;
save = M2ICC2E1cond.dat;

ANALYSIS: TYPE = TWOLEVEL;
MODEL POPULATION:
%WITHIN%
fw1 BY y1-y4@1;
fw2 BY y5-y8@1;
y1-y8*4;
fw1-fw2*2;
fw1 WITH fw2*0.65;
%BETWEEN%
fb1 BY y1-y4@1;
fb2 BY y5-y8@1;
y1-y8@.71;
fb1-fb2*.35;
fb1 WITH fb2*.31;
MODEL:
%WITHIN%
fw1 BY y1@1 y2-y4*1;
fw2 BY y5@1 y6-y8*1;
y1-y8*4;
fw1-fw2*2;
fw1 WITH fw2*0.65;
%BETWEEN%
fb1 BY y1@1 y2-y4*1;
fb2 BY y5@1 y6-y8*1;
y1-y8*.71;
fb1-fb2*.35;
fb1 WITH fb2*.31;
Output: tech9;

B. Bayesian Estimation Syntax

```
MONTECARLO:

names are y1-y8;
nobservations = 5000;
ncsizes = 1;
csizes = 200 (25);
seed = 50277;
nreps = 1000;
save = M2ICC2E2cond.dat;

ANALYSIS: TYPE = TWOLEVEL;
estimator = bayes;
MODEL POPULATION:
%WITHIN%
fw1 BY y1-y4@1;
fw2 BY y5-y8@1;
y1-y8*4;
fw1-fw2*2;
fw1 WITH fw2*0.65;
%BETWEEN%
fb1 BY y1-y4@1;
fb2 BY y5-y8@1;
y1-y8@.71;
fb1-fb2*.35;
fb1 WITH fb2*.31;
MODEL:
%WITHIN%
fw1 BY y1@1 y2-y4*1;
fw2 BY y5@1 y6-y8*1;
y1-y8*4;
fw1-fw2*2;
fw1 WITH fw2*0.65;
%BETWEEN%
fb1 BY y1@1 y2-y4*1;
fb2 BY y5@1 y6-y8*1;
y1-y8*.71;
fb1-fb2*.35;
fb1 WITH fb2*.31;
Output: tech9;
```

Appendix B

Within-Level Parameters Obtained From Multicollinearity Simulations

Table B.1. MLR Results of Within-Level Parameters From Within-Level Multicollinearity Simulations.

Multicollinearity		Intraclass correlation coefficient	Mean bias λ	Mean bias SE λ	Bias covariance Ψ_{12}	SE covariance Ψ_{12}
Low (.30)		Medium (.15)	0.20	1.24	−0.35	−1.59
		High (.25)	0.21	1.24	−0.35	−1.59
Medium (.50)		Medium (.15)	0.21	1.05	−0.32	−1.72
		High (.25)	0.21	0.96	−0.31	−1.71
High	.80	Medium (.15)	0.19	0.08	−0.29	−1.94
		High (.25)	0.19	0.12	−0.29	−2.08
	.90	Medium (.15)	0.18	−0.24	−0.27	−2.36
		High (.25)	0.18	−0.24	−0.27	−2.36
	.95	Medium (.15)	0.18	−0.23	−0.27	−2.46
		High (.25)	0.18	−0.23	−0.27	−2.46

Note. MLR = maximum likelihood with robust chi-squares and standard errors.

Table B.2. Bayesian Results of Within-Level Parameters From Within-Level Multicollinearity Simulations.

Multicollinearity		Intraclass correlation coefficient	Mean bias λ	Mean bias SE λ	Bias covariance Ψ_{12}	SE covariance Ψ_{12}
Low (.30)		Medium (.15)	0.17	−3.79	0.06	−2.55
		High (.25)	0.14	−5.35	0.05	−4.09
Medium (.50)		Medium (.15)	0.12	−3.75	0.07	−4.33
		High (.25)	0.17	−4.53	−0.11	−6.10
High	.80	Medium (.15)	0.12	−3.01	0.01	−4.48
		High (.25)	0.13	−3.44	−0.06	−4.64
	.90	Medium (.15)	0.10	−3.73	0.04	−3.51
		High (.25)	0.11	−5.06	−0.03	−4.13
	.95	Medium (.15)	0.13	−2.96	−0.15	−1.84
		High (.25)	0.09	−5.22	−0.02	−3.99

Table B.3. MLR Results of Within-Level Parameters From Between-Level Multicollinearity Simulations.

Multicollinearity		Intraclass correlation coefficient	Mean bias λ	Mean bias SEλ	Bias covariance Ψ_{12}	SE covariance Ψ_{12}
Low (.30)		Medium (.15)	0.20	1.24	−0.35	−1.59
		High (.25)	0.21	1.24	−0.35	−1.59
Medium (.50)		Medium (.15)	0.21	1.24	−0.35	−1.59
		High (.25)	0.21	1.29	−0.35	−1.59
High	.80	Medium (.15)	0.21	1.24	−0.35	−1.59
		High (.25)	0.21	1.29	−0.35	−1.59
	.90	Medium (.15)	0.21	1.24	−0.35	−1.59
		High (.25)	0.21	1.29	−0.35	−1.59
	.95	Medium (.15)	0.21	1.24	−0.35	−1.59
		High (.25)	0.21	1.29	−0.35	−1.59

Note. MLR = maximum likelihood with robust chi-squares and standard errors.

Table B.4. Bayesian Results of Within-Level Parameters From Between-Level Multicollinearity Simulations.

Multicollinearity		Intraclass correlation coefficient	Mean bias λ	Mean bias SEλ	Bias covariance Ψ_{12}	SE covariance Ψ_{12}
Low (.30)		Medium (.15)	0.17	−3.79	0.06	−2.55
		High (.25)	0.13	−5.24	0.09	−4.47
Medium (.50)		Medium (.15)	0.13	−3.95	0.17	−3.89
		High (.25)	0.11	−6.08	0.11	−4.65
High	.80	Medium (.15)	0.15	−4.62	0.23	−3.89
		High (.25)	0.15	−5.30	0.03	−4.64
	.90	Medium (.15)	0.15	−3.44	0.31	−4.09
		High (.25)	0.16	−4.29	0.05	−3.70
	.95	Medium (.15)	0.17	−3.63	0.31	−4.45
		High (.25)	0.15	−4.05	0.17	−3.52

Declaration of Conflicting Interests

Funding

Note

1. In Mplus, the default prior distributions for means and intercepts of observed and latent continuous variables, thresholds of observed categorical dependent variables, factor loadings, and regression coefficients are normal distributions with a prior mean of zero and an infinitive large prior variance. For the prior distributions of variances and residual variances of observed and latent parameters, gamma distributions are used, but an inverse Wishart distribution is used if more than one latent variable is estimated, and a Dirichlet distribution is used for categorical (latent) variables.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

Asparouhov, T., & Muthén, B. (2010a). *Bayesian analysis of latent variable models using Mplus*. Manuscript submitted for publication.

Asparouhov, T., & Muthén, B. (2010b). *Bayesian analysis using Mplus: Technical implementation*. Manuscript submitted for publication.

Belsley, D. A. (1991). *Conditioning diagnostics: Collinearity and weak data in regression*. New York, NY: Wiley.

Bollen, K. A. (1989). *Structural equations and latent variables*. New York, NY: Wiley.

Boomsma, A. (1983). *On the robustness of LISREL (maximum likelihood estimation) against small sample size and nonnormality*. Amsterdam, Netherlands: Sociometric Research Foundation.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/ correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Fabrycy, M. Z. (1975). Multicollinearity caused by specification errors. *Applied Statistics*, *24*, 250-254.

Farrar, D. E., & Glauber, R. R. (1967). Multicollinearity in regression analysis: The problem revisited. *Review of Economics and Statistics*, *49*, 92-107.

Field, A. (2013). *Discovering statistics using IBM SPSS: And sex and drugs and rock 'n' roll* (4th ed.). London, England: Sage.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis* (2nd ed.). London, England: Chapman & Hall/CRC Press.

Gordon, R. A. (1968). Issues in multiple regression. *American Journal of Sociology*, *73*, 592-616.

Grewal, R., Cote, J. A., & Baumgartner, H. (2004). Multicollinearity and measurement error in structural equation models: Implications for theory testing. *Marketing Science*, *23*, 519-529.

Haitovsky, Y. (1969). Multicollinearity in regression analysis: A comment. *Review of Economics and Statistics*, *51*, 486-489.

Harvey, A. C. (1977). Some comments on multicollinearity in regression. *Applied Statistics*, *26*, 188-191.

Hays, W. L. (1981). *Statistics* (3rd ed.). New York, NY: Holt, Rinehart & Winston.

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

Howard, G. S., Maxwell, S. E., & Fleming, K. J. (2000). The proof of the pudding: An illustration of the relative strengths of null hypothesis, meta-analysis, and Bayesian analysis. *Psychological Methods*, *5*, 315-332.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York, NY: Routledge.

Hox, J. J., & Maas, C. J. M. (2001). The accuracy of multilevel structural equation modeling with pseudobalanced groups and small samples. *Structural Equation Modeling*, *8*, 157-174.

Jagpal, H. S. (1982). Multicollinearity in structural equation models with unobservable variables. *Journal of Marketing Research*, *19*, 431-439.

Jedidi, K., & Ansari, A. (2001). Bayesian structural equation models for multilevel data. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 129-157). Mahwah, NJ: Lawrence Erlbaum.

Johnston, J. (1963). *Econometric methods*. New York, NY: McGraw-Hill.

Julian, M. W. (2001). The consequences of ignoring multilevel data structures in nonhierarchial covariance modeling. *Structural Equation Modeling*, *8*, 325-352.

Kaplan, D. (1994). Estimator conditioning diagnostics for covariance structure models. *Sociological Methods*, *23*, 220-229.

Kleinbaum, D., Kupper, L., & Muller, K. (1988). *Applied regression analysis and other multivariable methods*. Boston, MA: PWS-Kent.

Kreft, I. G. G., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.

Kubitschek, W. N., & Hallinan, M. T. (1999). Collinearity, bias and effect size: Modeling the effect of track on achievement. *Social Science Research*, *28*, 380-402.

Lee, S.-Y. (2007). *Structural equation modeling: A Bayesian approach*. Chichester, England: Wiley.

Lee, M. D., & Wagenmakers, E.-J. (2005). Bayesian statistical inference in psychology: Comment on Trafimow (2003). *Psychological Review*, *112*, 662-668.

Lynch, S. (2007). *Introduction to applied bayesian statistics and estimation for social scientists*. New York, NY: Springer.

Malinvaud, E. (1966). *Statistical methods for econometrics*. Amsterdam, Netherlands: North Holland.

Marsh, H. W., Dowson, W., Pietsch, J., & Walker, R. (2004). Why multicollinearity matters: A reexamination of relations between self-efficacy, self-concept, and achievement. *Journal of Educational Psychology*, *96*, 518-522.

Mason, C., & Perrault, W. (1991). Collinearity, power and interpretations of multiple regression analysis. *Journal of Marketing Research*, *28*, 268-280.

Mehta P. R., & Neale M. C. (2005). People are variables too: multilevel structural equations modeling. *Psychological Methods*, *10*, 259-284.

Morrow-Howell, N. (1994). The M word: Multicollinearity in multiple regression. *Social Work Research*, *18*, 247-251.

Muthén, B. O. (1990, June). *Mean and covariance structure analysis of hierarchical data*. *Paper presented at the Psychometric Society meeting*, Princeton, NJ. UCLA Statistics Series 62.

Muthén, L. K., & Muthén, B. O. (2013). Mplus (Version 7.11) [Computer software]. Los Angeles, CA: Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (1998-2014). *Mplus users' guide*. Los Angeles, CA: Muthén & Muthén.

Neter, J., Wasserman, W., & Whitmore, G. (1978). *Applied statistics*. Boston, MA: Allyn & Bacon.

Paxton, P., Curran, P. J., Bollen, K. A., Kirby, J., & Fen, C. (2001). Monte Carlo experiments: Design and implementation. *Structural Equation Modeling*, *8*, 287-312.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research* (2nd ed.). New York, NY: Holt, Rinehart & Winston.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Fort Worth, TX: Harcourt Brace Jovanovich.

Schmidt, P., & Muller, E. N. (1978). The problem of multicollinearity in a multistage causal alienation model: A comparison of ordinary least squares, maximum-likelihood, and ridge estimators. *Quality and Quantity*, *12*, 267-297.

Shieh, Y., & Fouladi, R. T. (2003). The effect of multicollinearity on multilevel structural equation modeling parameter estimates and standard errors. *Educational and Psychological Methods*, *63*, 951-985.

Skrondal, A. (2000). Design and analysis of Monte Carlo experiments: Attacking the conventional wisdom. *Multivariate Behavioral Research*, *35*, 137-167.

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences* (4th ed.). Mahwah, NJ: Lawrence Erlbaum.

Tabachnick, B. G., & Fidell, L. S. (2006). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon.

van de Schoot, R., Hoijtink, H., Mulder, J., Van Aken, M. A. G., Orobio de Castro, B., Meeus, W., & Romeijn, J.-W. (2011). Evaluating expectations about negative emotional states of aggressive boys using Bayesian model selection. *Developmental Psychology*, *47*, 203-212.

Walker, L. J., Gustafson, P., & Hennig, K. H. (2001). The consolidation/transition model in moral reasoning development. *Developmental Psychology*, *37*, 187-197.