STRUCTURAL EQUATION MODELING, *12*(3), 343–367

Sensitivity of Fit Indexes to Misspecified Structural or Measurement Model Components: Rationale of Two-Index Strategy Revisited

Xitao Fan
University of Virginia

Stephen A. Sivo
University of Central Florida

In previous research (Hu & Bentler, 1998, 1999), 2 conclusions were drawn: standardized root mean squared residual (SRMR) was the most sensitive to misspecified factor covariances, and a group of other fit indexes were most sensitive to misspecified factor loadings. Based on these findings, a 2-index strategy—that is, SRMR coupled with another index—was proposed in model fit assessment to detect potential misspecification in both the structural and measurement model parameters. Based on our reasoning and empirical work presented in this article, we conclude that SRMR is not necessarily most sensitive to misspecified factor covariances (structural model misspecification), the group of indexes (TLI, BL89, RNI, CFI, Gamma hat, Mc, or RMSEA) are not necessarily more sensitive to misspecified factor loadings (measurement model misspecification), and the rationale for the 2-index presentation strategy appears to have questionable validity.

The assessment of model fit in structural equation modeling (SEM) has long been a thorny issue in SEM application. As a result, the issues related to model fit assessment in SEM analysis have been at the forefront of theoretical and empirical research over the years. Research in this area has focused on different issues concerning the use and interpretation of model fit indexes. Studies typically examined the performance characteristics of different fit indexes under different data condi-

Requests for reprints should be sent to Xitao Fan, EDLF, Curry School of Education, University of Virginia, 405 Emmet Street South, Charlottesville, VA 22903–2495. E-mail: xfan@virginia.edu

tions; for example, sample size, estimation methods, and model misspecification (e.g., Fan, Thompson, & Wang, 1999; Fan & Wang, 1998; Gerbing & Anderson, 1993; Marsh, Balla, & Hau, 1996). More recently, research shifted to the search for empirically based cutoff criteria for model fit indexes, with the intention of providing more definitive guidelines for evaluating model fit (Enders & Finney, 2003; Hu & Bentler, 1998, 1999; Yu & Muthén, 2002).

TYPES OF MISSPECIFICATION AND TWO-INDEX STRATEGY

As of now, the most influential studies in this line of research are those by Hu and Bentler (1998, 1999). Their studies concluded that different fit indexes are differentially sensitive to either measurement model misspecification (e.g., a misspecified factor loading) or structural model misspecification (e.g., a misspecified covariance between two factors; Hu & Bentler, 1998). More specifically, two conclusions were drawn: (a) standardized root mean squared residual (SRMR) was the most sensitive to the misspecified factor covariances (misspecified structural model components), and (b) a group of other fit indexes (Tucker–Lewis Index [TLI], Bollen's delta [BL89], Relative Centrality Index [RNI], Comparative Fit Index [CFI], Gamma hat [Gamma], McDonald's Centrality Index [Mc], and Root Mean Squared Error of Approximation [RMSEA]) were most sensitive to misspecified factor loadings (misspecified measurement model components).

These conclusions naturally led to the proposal (Hu & Bentler, 1998) of a two-index strategy for model fit assessment: SRMR is always needed because of its sensitivity to misspecified structural model components, and another fit index (TLI, BL89, RNI, CFI, Gamma, Mc, or RMSEA) is also needed because of its sensitivity to misspecified measurement model components. Subsequently, some empirically based cutoff criteria for fit indexes were proposed (Hu & Bentler, 1999) in model fit assessment. The two-index strategy and the proposed cutoff criteria for fit indexes in model fit assessment have been gaining popularity in SEM applications (e.g., Corten et al., 2002; DiStefano, 2002; Glaser, 2001, 2002; Moulder & Algina, 2002; Pomplun & Omar, 2001).

These new developments on SEM model fit assessment have broken new ground. However, a close examination of these pioneering studies reveals that two important issues stand out: (a) The severity of model misspecification has not been quantified nor adequately controlled, making the internal validity of the conclusions from these studies questionable; and (b) there is an obvious lack of diversity in terms of the models and model parameters examined, raising the concern about the external validity (generalizability) of these conclusions.

To extend this line of research, Yu and Muthén (2002) incorporated categorical variables in their SEM models. Enders and Finney (2003) considered some important issues such as model complexity, choice of population parameter values, and the issue of differential power for misspecified SEM models. The findings from Enders and Finney suggest that the proposed cutoff criteria had limited generalizability, and did not fare well under different model and data conditions. More recently, Marsh, Hau, and Wen (2004) provided more detailed comments and analyses on the utility of using the proposed cutoff criteria in model fit assessment. Based on both theoretical and empirical grounds, Marsh, Hau, and Wen highlighted some important issues and problems in the practice of using the proposed cutoff criteria in model fit assessment within the framework of hypothesis testing.

SEVERITY OF MODEL MISSPECIFICATION

One critical issue that has not been adequately addressed in this area of research is severity (degree) of model misspecification. As discussed in Fan et al. (1999) and Fan and Wang (1998), model misspecification is a difficult issue, both because of the ambiguity and lack of efforts in quantifying the severity of misspecification, and because of the variety of forms in which model misspecification can occur. It is typically not very useful to just explain what has been misspecified in a model (e.g., a factor loading is fixed to zero). It is much more important and informative to specify the severity of misspecification. In other words, misspecification conditions should be quantified so that different misspecified models (e.g., a model with factor loadings misspecified, and a model with factor covariance misspecified) can be compared in terms of severity of misspecification. Unfortunately, this issue is not typically addressed, except in Enders and Finney's (2003) study in which the statistical power for rejecting the misspecified models was considered. This issue has important implications that are discussed later.

When we evaluate the sensitivity of fit indexes to different types of misspecification, it is logical that severity of misspecification should be considered. It appears that the most sensible approach is to use the noncentrality parameter (i.e., noncentral χ^2), and its associated statistical power for rejecting the misspecified model, to describe the severity of misspecification. The noncentrality parameter describes the amount of shift from central to noncentral χ^2 distributions due to model misspecification, regardless of the types of misspecification (e.g., misspecification in measurement vs. in structural components in a model).

The noncentrality parameter, together with its associated degrees of freedom, determines the statistical power for statistically rejecting the misspecified model. As a result, the statistical power for rejecting the misspecified model is blind to the types of model misspecification. It is thus reasonable to say that, if the power for rejecting two different misspecified models is comparable, the severity of misspecification for the two misspecified models should be considered comparable.

RATIONALE OF THE TWO-INDEX STRATEGY

Hu and Bentler (1998) considered two confirmatory factor analysis (CFA) models with different types of misspecifications, as shown in Figure 1. These two models were named simple (Figure 1a) and complex (Figure 1b) models, respectively.

In the simple model, misspecification occurred when the covariance(s) among the latent factors were misspecified to be zeros. There were two levels of misspecified simple model: In the first, one factor covariance was misspecified to be 0 ($s_{12} = 0$ in Figure 1a); in the second, two covariances were misspecified to be 0 (both $s_{12} = 0$ and $s_{13} = 0$ in Figure 1a). In the complex model, the factor loadings (measurement model components) were misspecified. Again, there were two lev-

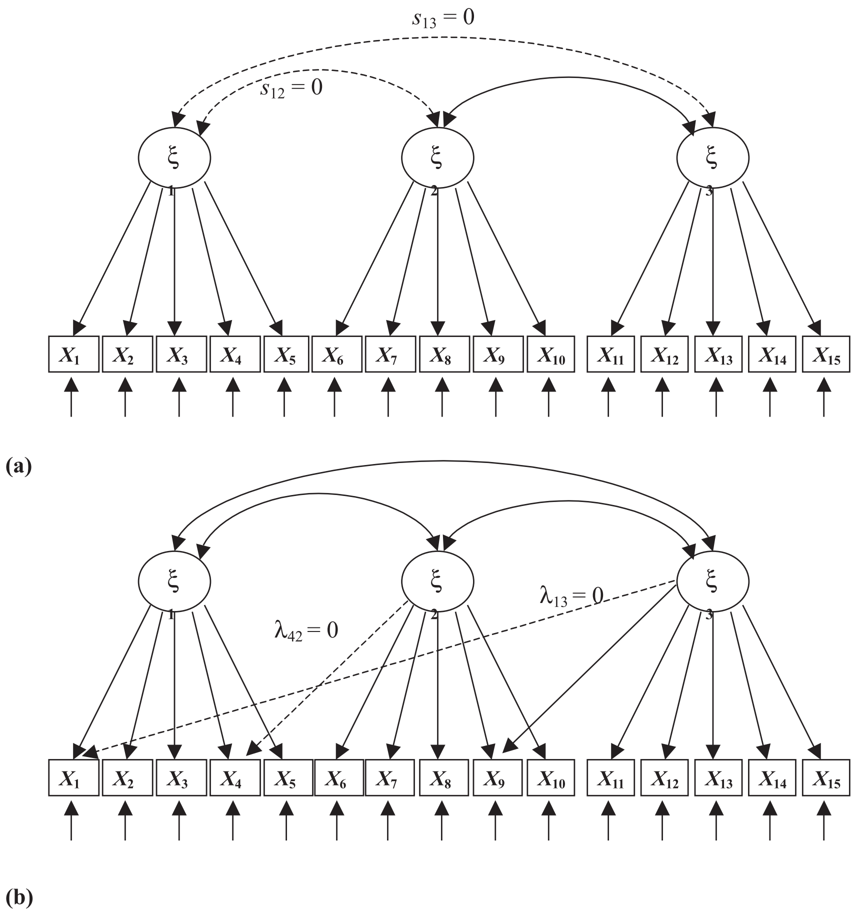

FIGURE 1 Simple and complex models.

els of misspecified complex model: In the first, one pattern coefficient was misspecified to be 0 ($\lambda_{13} = 0$ in Figure 1b); in the second, two coefficients were misspecified to be 0 (both $\lambda_{13} = 0$ and $\lambda_{42} = 0$ in Figure 1b).

Based on Monte Carlo simulation work that involved multiple design factors (e.g., model misspecification, sample size, estimation method, data distribution shape), two observations were made. First, the correlations among 15 different fit indexes (Hu & Bentler, 1998, Table 3) suggested two main clusters of fit indexes, with Normed Fit Index (NFI), Fit Index by Bollen (BL86; 1986), Goodness-of-Fit Index (GFI), Adjusted Goodness-of-Fit Index (AGFI), Rescaled Akaike's Information Criterion (CAK), and a Cross-validation Index (CK) being in one group, and TLI, BL89, RNI, CFI, Mc, and RMSEA forming another group. SRMR, however, was the least similar to either of these two clusters of fit indexes. This observation suggested a multifactor view for the fit indexes.

Second, the sensitivity of fit indexes to the two types of misspecification (misspecified factor covariances vs. misspecified factor loadings) was compared. The sensitivity of a fit index to model misspecification was quantified as the percentage of the total variation of a fit index attributable to the design factor of model misspecification. Quantitatively, percentage of variation attributable to a design factor is the η^2 ($\eta^2 = \textit{Sum of Squares}_{source} / \textit{Sum of Squares}_{total}$) derived from an analysis of variance (ANOVA) model. For a fit index, a large η^2 attributable to the design factor of model misspecification suggests high sensitivity of the index to model misspecification conditions. This approach is sensible, and it has been used by other researchers studying similar issues (e.g., Fan & Wang, 1998).

Hu and Bentler (1998) concluded that (a) SRMR was the most sensitive to the condition of misspecified factor covariances (simple model), and (b) a group of other fit indexes (TLI, BL89, RNI, CFI, Gamma, Mc, and RMSEA) were most sensitive to the condition of misspecified factor loadings (complex model). Table 1 reproduces a small portion of a table in Hu and Bentler (1998, Table 3, on p. 439) to illustrate these findings.

Table 1 shows that, for the simple model (misspecified factor covariances), the η^2 for SRMR (.914) is substantially larger than that (.653) for the complex model (misspecified factor loadings). On the other hand, for a group of other fit indexes (TLI, BL89, RNI, CFI, Gamma, Mc, and RMSEA), the η^2s for the complex model (misspecified factor loadings) are much larger than those for the simple model (misspecified factor covariance). These findings led to the conclusions that (a) SRMR is the most sensitive to misspecified factor covariances (structural model misspecification), (b) a group of other fit indexes are most sensitive to misspecified factor loadings (measurement model misspecification), and (c) to detect misspecification in both the structural and measurement model components, a two-index strategy should be used: SRMR coupled with another index (TLI, BL89, RNI, CFI, Gamma, Mc, or RMSEA). Subsequently, based on simulation work involving the same models and model misspecification conditions (see Figure 1), cutoff criteria for these indexes were proposed (Hu & Bentler, 1999).

TABLE 1
Partially Reproduced Hu and Bentler (1998, Table 3, p. 439) Data: η^2 Attributable to Model Misspecification for Simple and Complex Models

		Misspecification		
	(Other Factors)	*Simple*	*Complex*	*(Interactions)*
NFI	...	.151	.548	...
BL86 (Rho1)	...	.143	.534	...
TLI[a]	...	.315	.748	...
BL89 (Delta)[a]	...	.330	.763	...
RNI[a]	...	.326	.759	...
CFI[a]	...	.321	.759	...
GFI	...	.101	.471	...
AGFI	...	.094	.454	...
Gamma hat[a]	...	.309	.743	...
CAK	...	.061	.301	...
CK	...	.057	.286	...
Mc (Centrality)[a]	...	.339	.766	...
CN	...	.221	.256	...
SRMR[b]	...	.914	.653	...
RMSEA[a]	...	.466	.763	...

Note. NFI = Normed Fit Index; BL86 = Bollen's Fit Index; TLI = Tucker–Lewis Index; BL89 = Bollen's delta; RNI = Relative Noncentrality Index; CFI = Comparative Fit Index; GFI = Goodness-of-Fit Index; AGFI = Adjusted Goodness-of-Fit Index; CAK = Rescaled Akaike's Information Criterion; CK = Cross-validation Index; Mc = McDonald's Centrality Index; CN = Hoelter's Critical *N*; SRMR = Standardized Root Mean Squared Residual; RMSEA = Root Mean Squared Error of Approximation.

[a]Most sensitive to misspecified factor loadings (complex model). [b]Most sensitive to misspecified factor covariances ("simple" model).

UNRESOLVED ISSUES FOR THE TWO-INDEX STRATEGY

In the studies already reviewed, the severity of model misspecification was not controlled. As a result, the types of misspecification implemented for the simple and complex models (i.e., a misspecified factor loading vs. a misspecified factor covariance) might be confounded with the severity of model misspecification, and such confounding could have compromised the validity of the conclusions already discussed. To explore this issue, we examined the severity of misspecification for the simple and complex models as implemented in Hu and Bentler (1998), by fitting the misspecified models to the respective population covariance matrices. Population covariance matrices were derived from the model parameters ($\Sigma = \Lambda_X \Phi \Lambda_X' + \theta_\delta$). Figure 2 presents the original model parameters (Hu & Bentler, 1998), and the information about the severity of misspecification for the two types of misspecified models (the simple and complex models).

"Simple" Model Parameters:

$$\Phi = \begin{bmatrix} 1.00 & & \\ .50 & 1.00 & \\ .40 & .30 & 1.00 \end{bmatrix}$$

$$\Lambda_X\,(Transposed) = \begin{bmatrix} .70 & .70 & .75 & .80 & .80 & .00 & .00 & .00 & .00 & .00 & .00 & .00 & .00 & .00 & .00 \\ .00 & .00 & .00 & .00 & .00 & .70 & .70 & .75 & .80 & .80 & .00 & .00 & .00 & .00 & .00 \\ .00 & .00 & .00 & .00 & .00 & .00 & .00 & .00 & .00 & .00 & .70 & .70 & .75 & .80 & .80 \end{bmatrix}$$

$$\theta_\delta\,(Diagonal) = \begin{bmatrix} .51 & .51 & .4375 & .36 & .36 & .51 & .51 & .4375 & .36 & .36 & .51 & .51 & .4375 & .36 & .36 \end{bmatrix}$$

"Complex" Model Parameters:

$$\Phi = [same\ \ as\ \ above]$$

$$\Lambda_X\,(Transposed) = \begin{bmatrix} .70 & .70 & .75 & .80 & .80 & .00 & .00 & .00 & .00 & .00 & .00 & .00 & .00 & .00 & .00 \\ .00 & .00 & .00 & \underline{.70} & .00 & .70 & .70 & .75 & .80 & .80 & .00 & .00 & .00 & .00 & .00 \\ \underline{.70} & .00 & .00 & .00 & .00 & .00 & .00 & .00 & \underline{.70} & .00 & .70 & .70 & .75 & .80 & .80 \end{bmatrix}$$

$$\theta_\delta = [same\ \ as\ \ above]$$

Non-comparable Severity of Misspecification for Simple and Complex Models (N = 100)

	χ^2	*df*	*Power*[a]
Simple model (factor covariances misspecified)			
One misspecified factor covariance	20.83	88	0.43
Two misspecified factor covariances	28.33	89	0.61
Complex model (factor pattern coefficients misspecified)			
One misspecified factor loading	40.68	85	0.84
Two misspecified factor loadings	75.11	86	0.99

[a]Power estimation is based on Satorra-Saris method (Satorra & Saris, 1993; Saris & Satorra, 1983).

FIGURE 2 Original model parameters in Hu and Bentler (1998) and severity of misspecification.

The χ^2s in the table embedded in Figure 2 are actually noncentrality parameters for the misspecified models. For a misspecified model fitted to the population covariance matrix, the nonzero χ^2 value represents the shift from central χ^2 to noncentral χ^2; that is, the noncentrality parameter. Figure 2 shows that the severity of model misspecification is different for the simple and complex models. Based on the estimated statistical power (Saris & Satorra, 1993; Satorra & Saris, 1983) for rejecting the misspecified models, misspecification is less severe for the misspecified simple models (for $\alpha = .05$, power = .43 and .61, respectively, for rejecting the two misspecified simple models) than for the misspecified complex models (for $\alpha = .05$, power = .84 and .99, respectively, for rejecting the two misspecified complex models). This indicates that severity of misspecification is confounded with types of misspecification. Here and elsewhere in the article, we use the estimated noncentrality and its associated statistical power to operationally define severity of misspecification. This approach was used primarily because we are not aware of a better approach for this purpose. We are, however, open to other suggestions or alternative approaches.

The information in Figure 2 indicates that it may not be the types of misspecification (misspecified factor covariances vs. factor loadings) that contributed to the results observed by Hu and Bentler (1998; see Table 1), but rather, the severity of misspecification. The confounding between types of misspecification and severity of misspecification may have undermined the validity of the previous conclusion that some indexes (TLI, BL89, RNI, CFI, Gamma, Mc, and RMSEA) were more sensitive to misspecified factor loadings, whereas SRMR was more sensitive to misspecified factor covariances.

PURPOSE AND SCOPE OF THE STUDY

This study intended to evaluate the validity of the two-index strategy by partially replicating the study by Hu and Bentler (1998). The critical issue in this replication design was to control the severity of model misspecification. The findings from the previous studies (Hu & Bentler, 1998, 1999) were not based on appropriate study design; consequently, the confounding of severity of misspecification and types of misspecification might have led to incorrect conclusions. Before the proposed two-index strategy and the related cutoff criterion of fit indexes can be embraced by SEM researchers in general, the issues raised in this article should be addressed.

It is important to point out that this study was limited to the evaluation of the rationale of the two-index strategy, but there was no attempt to study the implementation of the two-index strategy in the form of using cutoff values of fit indexes in model fit assessment. In this regard, Marsh et al. (2004) recently provided detailed comments and analyses on the utility of using the proposed cutoff criteria in SEM model fit assessment, and they highlighted some issues and problems in such practice. Although the article by Marsh et al. may appear to be similar to this article, the issues discussed here are actually different from those in Marsh et al. Because the scope of this study is much narrower compared with those by Hu and Bentler (1998, 1999), and many issues addressed by Hu and Bentler were not studied in this article, this study should not be considered a full replication of the Hu and Bentler studies.

METHODS

Models and Model Misspecification

A Monte Carlo simulation experiment was conducted to study the issues previously discussed. The same two CFA models (simple and complex models; see Figure 1) as those in Hu and Bentler (1998, 1999) were used in this study, and the misspecified components in the simple and the complex models were identical to those in Hu and Bentler (1998, 1999). The purpose of this study requires that the severity of misspecification for the simple and complex models be comparable. To accomplish this, population model parameters were adjusted in such a way that the misspecified simple (misspecified factor covariances) and the complex (misspecified factor loadings) models had comparable severity of misspecification. For the purpose of providing a broader context for model misspecification, we created two conditions of misspecifications: slight misspecification (Condition I) and moderate misspecification (Condition II). Under each condition, there were two levels of misspecification for the simple and the complex models, in the same manner as shown in Figure 1, and the two levels of misspecification for the simple and the complex models represented comparable degrees of misspecification.

The slight misspecification condition (Condition I) mirrors the degree of misspecification for the original simple model in Hu and Bentler (1998), as shown in Figure 2 . This condition was created by adjusting the complex model parameter values in the $\mathbf{\Lambda}_X$ matrix (factor pattern coefficients) such that the severity of misspecification for the two misspecified complex models would be comparable to the original misspecified simple models in Hu and Bentler (1998). The adjusted complex model $\mathbf{\Lambda}_X$ matrix is shown here (two underlined parameters, adjusted down from the original parameter of 0.70; see original complex model in Figure 2 for comparison):

$$\mathbf{\Lambda}_X\,(Transposed) = \begin{bmatrix} .70 & .70 & .75 & .80 & .80 & .00 & .00 & .00 & .00 & .00 & .00 & .00 & .00 & .00 & .00 \\ .00 & .00 & .00 & \underline{\underline{.28}} & .00 & .70 & .70 & .75 & .80 & .80 & .00 & .00 & .00 & .00 & .00 \\ \underline{\underline{.455}} & .00 & .00 & .00 & .00 & .00 & .00 & .00 & .70 & .00 & .70 & .70 & .75 & .80 & .80 \end{bmatrix}$$

Except for the parameter adjustments just shown, all other parameters are exactly the same as those in Figure 2. The upper panel of Table 2 shows that, after these adjustments in the $\mathbf{\Lambda}_X$ matrix, the statistical power for rejecting the misspecified simple and complex models is comparable.

Under the moderate misspecification condition (Condition II), the severity of misspecification mirrors that of the original complex models in Hu and Bentler (1998) in Figure 2. This condition was created by adjusting the simple model parameters in the $\mathbf{\Phi}$ matrix (factor covariance matrix) such that the severity of misspecification for the two misspecified simple models would be comparable to the original misspecified complex models in Hu and Bentler (1998). The adjusted simple model parameters in the $\mathbf{\Phi}$ matrix are shown here (three underlined parameters, adjusted up from the original parameters of .50, .40, and .30, respectively; see the original simple model parameters in Figure 2 for comparison):

$$\mathbf{\Phi} = \begin{bmatrix} 1.00 & & \\ \underline{\underline{.67}} & 1.00 & \\ \underline{\underline{.747}} & \underline{\underline{.50}} & 1.00 \end{bmatrix}$$

Except for these parameter adjustments in the $\mathbf{\Phi}$ matrix, all other parameters are exactly the same as those in Figure 2. The lower panel of Table 2 shows that, af-

TABLE 2
Two Misspecification Conditions With Comparable Severity of Misspecification for Simple and Complex Models

	χ^2	*df*	*Power*[a]
Misspecification condition I (slight misspecification)			
Simple model (factor covariances misspecified)			
One misspecified factor covariance	20.83	88	0.43
Two misspecified factor covariances	28.33	89	0.61
Complex model (pattern coefficients misspecified)			
One misspecified factor loading	20.47	85	0.43
Two misspecified factor loadings	27.82	86	0.61
Misspecification condition II (moderate misspecification)			
Simple model (factor covariances misspecified)			
One misspecified factor covariance	41.24	88	0.84
Two misspecified factor covariances	75.75	89	0.99
Complex model (pattern coefficients misspecified)			
One misspecified factor loading	40.68	85	0.84
Two misspecified factor loadings	75.11	86	0.99

[a]Power estimation is based on Satorra–Saris method (Satorra & Saris, 1993; Saris & Satorra, 1983) for $N = 100$.

ter these adjustments made in the $\mathbf{\Phi}$ matrix, the statistical power for rejecting the misspecified simple and complex models is comparable.

Monte Carlo Simulation Design and Data Conditions

This study considered multivariate normal data only. Although other studies in this area typically involved nonnormal data conditions, not enough is known about the issues under study even when the data conditions are ideal. The decision was made to evaluate these issues under ideal data conditions without the potential complication of nonnormal data conditions.

Both maximum likelihood (ML) and generalized least squares (GLS) estimation methods were used, and these two estimation methods were implemented separately. Under each of the two misspecification conditions (slight vs. moderate), there were two models (the simple model and the complex model), with each model having three levels of misspecification (true model with no misspecification, a model with one parameter misspecified, and a model with two parameters misspecified). Ten sample size conditions were implemented, ranging from 150 to 1,500 at an interval of 150. Under each sample size condition, 500 random samples (i.e., replications) were generated based on the population covariance matrices for the simple and complex models, respectively. This Monte Carlo simulation design called for the generation of 40,000 samples ($2 \times 2 \times 2 \times 10 \times 500$). Because each sample dataset was fitted to each of three levels of misspecified model (simple and complex models, respectively)—the true model, the model with one misspecified parameter, and the model with two misspecified parameters (see Figure 1)—the total number of model fittings was 120,000 ($40,000 \times 3$). Relevant model fit indexes from each model fitting were saved for later analyses.

Data Source and Analyses

Based on the matrix decomposition procedures (Fan, Felsovalyi, Sivo, & Keenan, 2002; Kaiser & Dickman, 1962), data were simulated using a combination of SAS macro, SAS BASE, and SAS PROC IML (Interactive Matrix Language). Model fitting and estimation were implemented through SAS/PROC CALIS. For each random sample fitted to the three levels of misspecified models, the appropriate fit indexes were saved and accumulated for later analyses.

ANOVA was conducted for each fit index, with the fit index value as the dependent variable, and sample size (10 levels) and model misspecification (three levels: true model, one misspecified parameter, two misspecified parameters), and the interaction between the two, as the independent variables. ANOVA was conducted separately for the simple and complex models, separately for ML and GLS estimation conditions, and separately for Condition I and Condition II of misspecification. The same 15 fit indexes used in Hu and Bentler (1998) were evaluated. For each index, the Type III sum of squares attributable to each factor

(SS^{III}_{source}) and the total sum of squares (SS_{total}) were used to compute $\eta^2 : \eta^2_{source} = \dfrac{SS^{III}_{source}}{SS_{total}}$. The η^2 represents the percentage of variation in a fit index attributable to a factor in the ANOVA model (e.g., sample size, model misspecification). As discussed previously, because a fit index is designed to detect model misspecification, ideally a large proportion of variation in a fit index would be attributable to misspecification conditions. In other words, the η^2 value for a fit index represents the sensitivity of the fit index to misspecified model conditions.

RESULTS AND DISCUSSION

The upper panel of Table 3 presents the correlations among the fit indexes for the design conditions previously discussed. Hu and Bentler (1998) discussed that the correlation pattern of the fit indexes suggested multiple clusters of fit indexes: NFI, BL86, GFI, AGFI, CAK, and CK appeared to form a cluster, whereas TLI, BL89, RNI, CFI, Mc, and RMSEA grouped into another cluster. SRMR was the least similar to either of the two groups.

Examination of the upper panel of Table 3 shows that, after controlling for the severity of model misspecification in this study, correlation patterns similar to those in Hu and Bentler (1998) were also observed. SRMR showed lower correlations with other fit indexes in general, although not nearly as low as those shown in Hu and Bentler (1998, Table 3). This suggests the possible multifactor view for the fit indexes. To evaluate this proposition, an exploratory factor analysis was conducted for the fit indexes. Because CAK, CK, and CN are qualitatively different from other fit indexes, and typically, they are not used as stand-alone indexes for evaluating model fit, these three indexes were excluded in the exploratory factor analysis.

The results of the exploratory factor analysis indicated that a single factor adequately explains the correlation pattern among the fit indexes (the upper panel of Table 3). For the 12 fit indexes included in the analysis, the first three eigenvalues (principal component extraction) were 10.64, 0.79, and 0.46, respectively, indicating that a single dominant factor was sufficient in accounting for the correlation pattern among the fit indexes (89% of variance accounted for by this single factor). Another extraction method (e.g., principal factor extraction) led to the same conclusion. So for the design conditions implemented in this study, it appears that there is insufficient evidence to conclude that a multifactor view of the fit indexes should be adopted. The transposed vector of the pattern coefficients (based on principal component extraction) on this single factor is shown here:

$$P' = \begin{bmatrix} \text{NFI} & \text{BL86} & \text{TLI} & \text{BL89} & \text{RNI} & \text{CFI} & \text{GFI} & \text{AGFI} & \text{GAMMA} & \text{Mc} & \text{SRMR} & \text{RMSEA} \\ .95 & .95 & .98 & .98 & .98 & .98 & .91 & .90 & .98 & .98 & -.72 & -.95 \end{bmatrix}$$

TABLE 3
Correlations Among Fit Indexes (Maximum Likelihood Estimation)

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*
Complex and Original Simple Models (Comparable Misspecification Severity)															
1. NFI															
2. BL86 (Rho1)	0.999														
3. TLI	0.881	0.879													
4. BL89 (Delta)	0.880	0.877	0.999												
5. RNI	0.882	0.879	0.999	0.999											
6. CFI	0.886	0.883	0.998	0.999	0.999										
7. GFI	0.960	0.968	0.822	0.814	0.817	0.820									
8. AGFI	0.952	0.960	0.813	0.804	0.807	0.810	0.999								
9. Gamma hat	0.871	0.871	0.995	0.993	0.993	0.992	0.833	0.826							
10. CAK	–0.942	–0.947	–0.705	–0.700	–0.703	–0.709	–0.962	–0.960	–0.708						
11. CK	–0.934	–0.939	–0.685	–0.680	–0.684	–0.690	–0.956	–0.954	–0.689	0.999					
12. Mc Centrality	0.870	0.869	0.994	0.992	0.992	0.990	0.833	0.826	0.999	–0.705	–0.685				
13. CN	0.735	0.740	0.651	0.648	0.648	0.650	0.754	0.753	0.653	–0.688	–0.680	0.665			
14. SRMR	–0.682	–0.662	–0.708	–0.724	–0.723	–0.724	–0.504	–0.478	–0.667	0.496	0.484	–0.672	–0.540		
15. RMSEA	–0.850	–0.851	–0.958	–0.955	–0.955	–0.951	–0.829	–0.823	–0.963	0.691	0.673	–0.971	–0.761	0.677	

(continued)

TABLE 3 *(Continued)*

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*
Complex and Original Simple Models (Comparable Misspecification Severity)															
1. NFI															
2. BL86 (Rho1)	0.999														
3. TLI	0.920	0.918													
4. BL89 (Delta)	0.919	0.916	0.999												
5. RNI	0.921	0.918	0.999	0.999											
6. CFI	0.923	0.920	0.998	0.999	0.999										
7. GFI	0.962	0.952	0.898	0.906	0.907	0.910									
8. AGFI	0.970	0.961	0.904	0.911	0.912	0.914	0.999								
9. Gamma hat	0.913	0.912	0.996	0.995	0.995	0.994	0.899	0.905							
10. CAK	–0.947	–0.951	–0.763	–0.759	–0.762	–0.766	–0.892	–0.904	–0.763						
11. CK	–0.938	–0.943	–0.745	–0.741	–0.744	–0.749	–0.883	–0.895	–0.746	0.999					
12. Mc Centrality	0.910	0.909	0.995	0.994	0.994	0.992	0.898	0.904	0.999	–0.760	–0.742				
13. CN	0.777	0.782	0.719	0.715	0.715	0.718	0.766	0.778	0.727	–0.739	–0.731	0.736			
14. SRMR	–0.844	–0.863	–0.846	–0.828	–0.829	–0.830	–0.725	–0.749	–0.849	0.793	0.783	–0.853	–0.786		
15. RMSEA	–0.896	–0.897	–0.971	–0.968	–0.968	–0.966	–0.884	–0.892	–0.978	0.756	0.739	–0.983	–0.796	0.877	

Note. NFI = Normed Fit Index; BL86 = Bollen's Fit Index; TLI = Tucker–Lewis Index; BL89 = Bollen's delta; RNI = Relative Noncentrality Index; CFI = Comparative Fit Index; GFI = Goodness-of-Fit Index; AGFI = Adjusted Goodness-of-Fit Index; CAK = Rescaled Akaike's Information Criterion; CK = Cross-Validation Index; Mc = McDonald's Centrality Index; CN = Hoelter's Critical *N*; SRMR = Standardized Root Mean Squared Residual; RMSEA = Root Mean Squared Error of Approximation.

Table 4 presents the η^2 values for the 15 fit indexes under two conditions of model misspecification (light and moderate misspecifications) obtained with ML estimation for model fitting. These are the same 15 fit indexes evaluated in Hu and Bentler (1998). As explained previously, Condition I (slight misspecification) and Condition II (moderate misspecification) represent different levels of severity of misspecification (see Methods section for details).

The findings in Table 4 should be interpreted and discussed in relation to Table 1. As shown in Table 1, for model misspecification, a group of fit indexes (TLI, BL89, RNI, CFI, Gamma, Mc, RMSEA) have much higher η^2 values (~.75) under the complex model than those under the simple model (~.35), suggesting that these indexes are more sensitive to misspecified factor loadings (complex model) than to misspecified factor covariances (simple model). On the other hand, SRMR has a much higher η^2 value (.91) under the simple model than under the complex model (.65), suggesting that SRMR is more sensitive to misspecified factor covariances.

As argued previously, the validity of Hu and Bentler's (1998) results and conclusions might have been compromised by the confounding between types of model misspecification and severity of model misspecification. This argument is largely supported by the results in Table 4, obtained while holding the severity of model misspecification comparable between the simple and complex models. It is obvious that the group of fit indexes (TLI, BL89, RNI, CFI, Gamma, Mc, RMSEA) had very comparable η^2 values for the simple and complex models. For Condition I (slight misspecifications), the η^2 values (in the range of .85–.90) were all comparable across the two types of models. For Condition II (moderate misspecifications), the findings under Condition I were replicated, and the η^2 values were again quite comparable across the two types of misspecified models. Because Condition II represented more severely misspecified models, it makes sense that the η^2 values were higher (generally above .95) than those under Condition I, because a fit index should be more sensitive to more severely misspecified models. These findings suggest that, when severity of model misspecification is controlled, these indexes are not differentially sensitive to different types of model misspecification (misspecified factor covariance vs. misspecified factor loadings) as concluded in Hu and Bentler (1998). Very similar results were obtained with the GLS estimation method for model fitting, as shown in Table 5.

Considerations for SRMR

The fit index SRMR was advocated (Hu & Bentler, 1998) as the most sensitive to structural model misspecification (i.e., misspecified factor covariances). It still appears to show more sensitivity to misspecified factor covariances (Table 4 for Condition I, $\eta^2 = .965$ vs. .771 for simple and complex models, respectively; for Condition II, $\eta^2 = .988$ vs. .902 for simple and complex models, respectively), even when

TABLE 4
Percentage of Variance (η^2) Attributable to Sample Size and Model Misspecification for Simple and Complex Models—Maximum Likelihood Estimation

	Sample Size		*Misspecification*		*Interaction*	
	Simple	*Complex*[a]	*Simple*	*Complex*	*Simple*	*Complex*
Condition I[b] (slight misspecifications)						
NFI	.554	.551	.393	.382	.000	.000
BL86 (Rho1)	.563	.560	.382	.372	.000	.000
TLI[c]	.002	.002	.861	.843	.000	.000
BL89 (Delta)[c]	.002	.002	.864	.846	.000	.000
RNI[c]	.002	.003	.862	.844	.000	.000
CFI[c]	.005	.005	.867	.846	.001	.001
GFI	.624	.567	.327	.377	.000	.001
AGFI	.634	.577	.317	.366	.001	.001
Gamma hat[c]	.003	.003	.853	.847	.000	.000
CAK	.818	.829	.154	.144	.000	.000
CK	.833	.844	.141	.131	.000	.000
Mc Centrality[c]	.002	.002	.859	.853	.000	.000
CN	.216	.214	.522	.523	.214	.214
SRMR[d]	.006	.147	.965	.771	.003	.017
RMSEA[c]	.004	.004	.892	.889	.005	.006
Condition II (moderate misspecifications)						
NFI	.169	.159	.810	.807	.000	.000
BL86 (Rho1)	.173	.164	.805	.801	.001	.001
TLI[c]	.001	.000	.973	.959	.000	.000
BL89 (Delta)[c]	.000	.000	.974	.960	.000	.000
RNI[c]	.001	.000	.974	.959	.000	.000
CFI[c]	.001	.001	.975	.960	.000	.000
GFI	.297	.182	.679	.790	.002	.002
AGFI	.305	.187	.671	.784	.002	.002
Gamma hat[c]	.001	.000	.968	.961	.000	.000
CAK	.432	.442	.548	.534	.000	.000
CK	.458	.470	.523	.507	.000	.000
Mc Centrality[c]	.000	.000	.970	.964	.000	.000
CN	.137	.136	.611	.608	.212	.214
SRMR[d]	.001	.050	.988	.902	.001	.011
RMSEA[c]	.002	.001	.965	.963	.002	.002

Note. NFI = Normed Fit Index; BL86 = Bollen's Fit Index; TLI = Tucker–Lewis Index; BL89 = Bollen's delta; RNI = Relative Noncentrality Index; CFI = Comparative Fit Index; GFI = Goodness-of-Fit Index; AGFI = Adjusted Goodness-of-Fit Index; CAK = Rescaled Akaike's Information Criterion; CK = Cross-validation Index; Mc = McDonald's Centrality Index; CN = Hoelter's Critical *N*; SRMR = Standardized Root Mean Squared Residual; RMSEA = Root Mean Squared Error of Approximation.

[a]Simple model: factor covariances misspecified (structural model misspecification); complex model: factor pattern coefficients misspecified (measurement model misspecification).The severity of misspecification is comparable across the two models. [b]Under Condition I, misspecification is slight (see Table 2 and related discussion for details).Under Condition II, misspecification is moderate (see Table 2 and related discussion for details). [c]In Hu and Bentler (1998), these were characterized as the most sensitive to misspecified factor loadings (complex model). [d]In Hu and Bentler (1998), these were characterized as the most sensitive to misspecified factor covariances (simple model).

TABLE 5
Percentage of Variance (η^2) Attributable to Sample Size and Model Misspecification for Simple and Complex Models—Generalized Least Squares Estimation

	Sample Size		*Misspecification*		*Interaction*	
	Simple	*Complex*[a]	*Simple*	*Complex*	*Simple*	*Complex*
Condition I[b] (slight misspecifications)						
NFI	.545	.532	.410	.422	.007	.006
BL86 (Rho1)	.556	.544	.397	.408	.009	.008
TLI[c]	.012	.007	.816	.807	.000	.000
BL89 (Delta)[c]	.019	.014	.826	.821	.002	.001
RNI[c]	.012	.007	.818	.810	.000	.000
CFI[c]	.003	.001	.847	.842	.008	.006
GFI	.721	.705	.235	.244	.001	.000
AGFI	.730	.715	.225	.233	.001	.001
Gamma hat[c]	.021	.014	.819	.810	.004	.002
CAK	.898	.900	.084	.081	.001	.000
CK	.908	.911	.076	.073	.001	.000
Mc centrality[c]	.022	.015	.819	.811	.004	.002
CN	.270	.272	.470	.465	.207	.205
SRMR[d]	.013	.178	.949	.734	.003	.006
RMSEA[c]	.007	.004	.856	.850	.014	.012
Condition II (moderate misspecifications)						
NFI	.216	.197	.748	.775	.016	.012
BL86 (Rho1)	.223	.204	.739	.767	.018	.013
TLI[c]	.005	.002	.938	.949	.000	.000
BL89 (Delta)[c]	.011	.007	.938	.950	.002	.001
RNI[c]	.005	.002	.939	.951	.000	.000
CFI[c]	.001	.000	.954	.964	.003	.002
GFI	.482	.398	.488	.571	.001	.001
AGFI	.491	.407	.477	.560	.002	.001
Gamma hat[c]	.014	.009	.926	.938	.003	.002
CAK	.761	.716	.223	.268	.001	.001
CK	.782	.740	.204	.245	.001	.001
Mc centrality[c]	.014	.010	.926	.938	.003	.002
CN	.186	.184	.550	.553	.220	.216
SRMR[d]	.002	.070	.984	.875	.001	.004
RMSEA[c]	.003	.002	.936	.940	.008	.006

Note. NFI = Normed Fit Index; BL86 = Bollen's Fit Index; TLI = Tucker–Lewis Index; BL89 = Bollen's delta; RNI = Relative Noncentrality Index; CFI = Comparative Fit Index; GFI = Goodness-of-Fit Index; AGFI = Adjusted Goodness-of-Fit Index; CAK = Rescaled Akaike's Information Criterion; CK = Cross-validation Index; Mc = McDonald's Centrality Index; CN = Hoelter's Critical *N*; SRMR = Standardized Root Mean Squared Residual; RMSEA = Root Mean Squared Error of Approximation.

[a]Simple model: factor covariances misspecified (structural model misspecification); complex model: factor pattern coefficients misspecified (measurement model misspecification). The severity of misspecification is comparable across the two models. [b]Under Condition I, misspecification is slight (see Table 2 and related discussion for details). Under Condition II, misspecification is moderate (see Table 2 and related discussion for details). [c]In Hu and Bentler (1998), these were characterized as the most sensitive to misspecified factor loadings (complex model). [d]In Hu and Bentler (1998), these were characterized as the most sensitive to misspecified factor covariances (simple model).

the severity of model misspecification was held comparable. This finding contradicted our expectation, and made us wonder why this should be the case.

Closer examination of the two levels of misspecified simple model (see Figure 1a) led to the realization that the misspecified simple models represented a somewhat unusual situation: A single misspecified parameter ($s_{12} = 0$ in Figure 1a) resulted in a large number of covariances being zeros in the model-implied covariance matrix ($\hat{\Sigma}_{model}$). The additional misspecified parameter ($s_{13} = 0$ in Figure 1a) doubled the number of covariances being zeros in the model-implied covariance matrix. More specifically, when the first factor covariance was misspecified to be zero ($s_{12} = 0$), it resulted in $\left[\left(\frac{25}{(15 \times 14)/2}\right) \times 100 \approx 24\%\right]$ of the covariances being zeros in the model-implied covariance matrix. When the second factor covariance was misspecified to be zero ($s_{12} = 0$, and $s_{13} = 0$), this resulted in $\left[\left(\frac{50}{(15 \times 14)/2}\right) \times 100 \approx 48\%\right]$ of the covariances being zeros in the model-implied covariance matrix.

Substantively, the misspecified simple models represent the situation in which correlated factors are misspecified as orthogonal factors. The observation that one or two misspecified parameters result in such a large number of zero covariances in the model-implied covariance matrix was alarming. Furthermore, SRMR most directly reflects the condition of having numerous zero elements in the model-implied covariance matrix, because SRMR is directly influenced by the discrepancy between the corresponding elements in the sample and the model-implied covariance matrices. SRMR is obtained from the following (Bentler, 1995; Hu & Bentler, 1998; Jöreskog & Sörbom, 1981):

$$\text{SRMR} = \sqrt{\left\{2\sum_{i=1}^{p}\sum_{j=1}^{i}[(s_{ij} - \hat{\sigma}_{ij}) / s_{ii}s_{jj}]^2\right\} / p(p+1)}$$

where s_{ij} is a sample covariance between variables i and j, $\hat{\sigma}_{ij}$ is the model-implied covariance between the two, s_{ii} and s_{jj} are sample standard deviations for variables i and j, and p is the number of variables in the model analysis.

Regardless of how a model is misspecified, s_{ij}, s_{ii}, s_{jj}, and p will not be affected, and only $\hat{\sigma}_{ij}$ is affected by model misspecification. If model misspecification results in many zero $\hat{\sigma}_{ij}$ (model-implied covariances being zeros) whereas their counterparts (s_{ij}) are not, the impact will be directly reflected as large values of ($s_{ij} - \hat{\sigma}_{ij}$), because now $(s_{ij} - \hat{\sigma}_{ij}) = s_{ij}$.

Although many other fit indexes are also based on the discrepancy between sample covariance matrix and model-based covariance matrix, the computation of root mean squared residual (RMSR) and its standardized version (SRMR) lead to the expectation that the large number of zero covariances caused by the misspecified model parameters will make this index especially sensitive to this kind of model misspecification. As a result, the finding that SRMR is more sensitive to structural model misspecification (e.g., misspecified factor covariances) may not be generalizable beyond this condition.

To evaluate this hypothesis, a different type of misspecified factor covariance was considered in the simple model. Instead of factor covariances being misspecified to be zeros ($s_{12} = 0$, $s_{13} = 0$ in Figure 1a), the correlation between two factors was misspecified to be 1.00 (i.e., in Figure 1a, correlation between ξ_1 and ξ_2, and that between ξ_1 and ξ_3 were misspecified to be $r = 1.00$). Substantively, this misspecification represents the situation where two (highly) correlated factors are misspecified to be the same factor. For this purpose, we created a new simple model with parameters in the $\mathbf{\Phi}$ matrix (factor covariance matrix) being such that, when the factor correlations were misspecified to be 1.00, the severity of misspecification for the misspecified new simple models would be comparable to those of the original misspecified complex models in Hu and Bentler (1998). The new simple model's $\mathbf{\Phi}$ matrix parameters are shown here (three underlined parameters; see the original simple model parameters in Figure 2 for comparison):

$$\mathbf{\Phi} = \begin{bmatrix} 1.00 & & \\ \underline{\underline{.79}} & 1.00 & \\ \underline{\underline{.80}} & \underline{\underline{.80}} & 1.00 \end{bmatrix}$$

Except for these parameter adjustments in the $\mathbf{\Phi}$ matrix, all other parameters were exactly the same as those in Figure 2. Table 6 shows that the estimated statistical power was comparable for rejecting the misspecified new simple and complex models.

The first misspecified new simple model ($r_{12} = 1.0$) represents the situation where two correlated factors (ξ_1 and ξ_2) were misspecified to be the same factor, and a three-factor model was misspecified to be a two-factor model. The second misspecified new simple model ($r_{12} = 1.0$ and $r_{13} = 1.0$) represented the situation where a three-factor model was misspecified to be a one-factor model. Although the misspecification was still related to factor covariances, unlike the original simple model (Figure 1a), this type of misspecification for factor covariances did not force the model-implied covariances to be zeros. In addition, the severity of model misspecification between the new simple model and the original complex model was comparable, as shown in Table 6 by the estimated power for rejecting the misspecified new simple and complex models.

The lower panel of Table 3 presents the correlations among the fit indexes under these new conditions. The correlation pattern indicated that, in general, SRMR had slightly lower correlations with other fit indexes, although to the same degree as

TABLE 6
Comparable Severity of Misspecification for New Simple and Complex Models

	χ^2	*df*	*Power*[a]
New simple model (factor covariances misspecified)			
One misspecified factor covariance	42.26	89	0.85
Two misspecified factor covariances	75.67	90	1.00
Complex model (pattern coefficients misspecified)			
One misspecified factor loading	40.68	85	0.84
Two misspecified factor loadings	75.11	86	0.99

[a]Power estimation is based on Satorra–Saris method (Satorra & Saris, 1993; Saris & Satorra, 1983) for $N = 100$.

shown previously (Hu & Bentler, 1998, Table 3). An exploratory factor analysis was conducted for the fit indexes (with CAK, CK, and CN excluded). The results of the exploratory factor analysis suggested that a single factor explained the correlations among the fit indexes well. For the 12 fit indexes included in the analysis, the first three eigenvalues (principal component extraction) were 11.21, 0.41, and 0.30, respectively, indicating again that a single dominant factor was sufficient in explaining the correlations among the fit indexes (93% of variance accounted for by this single factor). The findings here suggest that, for the design conditions implemented in this study, there is insufficient evidence to support a multifactor view for the fit indexes. The transposed vector of the pattern coefficients (based on principal component extraction) on this single factor was:

$$P' = \begin{bmatrix} \text{NFI} & \text{BL86} & \text{TLI} & \text{BL89} & \text{RNI} & \text{CFI} & \text{GFI} & \text{AGFI} & \text{GAMMA} & \text{Mc} & \text{SRMR} & \text{RMSEA} \\ .96 & .96 & .98 & .98 & .98 & .98 & .94 & .95 & .98 & .98 & -.87 & -.97 \end{bmatrix}$$

Table 7 presents the percentage of variation (η^2) in sample fit indexes attributable to the factors of sample size and model misspecification for the new simple and the original complex models. The findings in Table 7 confirmed suspicions. Under ML estimation, the η^2 for SRMR is .834 and .902 for the new simple and complex models, respectively. Under GLS estimation, the η^2 for SRMR is .730 and .875 for the new simple and complex models. So contrary to the previous observation (Hu & Bentler, 1998), SRMR did not appear to be more sensitive to the misspecified factor covariances (simple model) than to the misspecified measurement factor loadings (complex model).

The findings in Table 7 provide support for our hypothesis: The previous conclusion concerning SRMR (i.e., SRMR was most sensitive to misspecified factor covariances) was very likely the result of a special type of misspecification that caused a large number of covariances to become zeros in the model-implied

TABLE 7
Percentage of Variance (η^2) Attributable to Sample Size and Model Misspecification for the New Simple Model[a] and the Original Complex Model

	Sample Size		*Misspecification*		*Interaction*	
	Simple	*Complex*[a]	*Simple*	*Complex*	*Simple*	*Complex*
Maximum likelihood estimation						
NFI	.169	.159	.779	.807	.000	.000
BL86 (Rho1)	.175	.164	.772	.801	.000	.001
TLI[b]	.001	.000	.940	.959	.000	.000
BL89 (Delta)[b]	.001	.000	.942	.960	.000	.000
RNI[b]	.001	.000	.941	.959	.000	.000
CFI[b]	.001	.001	.941	.960	.000	.000
GFI	.092	.182	.861	.790	.001	.002
AGFI	.096	.187	.856	.784	.002	.002
Gamma hat[b]	.001	.000	.950	.961	.000	.000
CAK	.426	.442	.542	.534	.000	.000
CK	.452	.470	.518	.507	.000	.000
Mc centrality[b]	.000	.000	.955	.964	.000	.000
CN	.138	.136	.611	.608	.213	.214
SRMR[c]	.106	.050	.834	.902	.016	.011
RMSEA[b]	.002	.001	.958	.963	.002	.002
Generalized least squares estimation						
NFI	.249	.197	.721	.775	.016	.012
BL86 (Rho1)	.261	.204	.706	.767	.018	.013
TLI[b]	.003	.002	.920	.949	.000	.000
BL89 (Delta)[b]	.008	.007	.928	.950	.001	.001
RNI[b]	.003	.002	.923	.951	.000	.000
CFI[b]	.000	.000	.949	.964	.002	.002
GFI	.542	.398	.419	.571	.000	.001
AGFI	.557	.407	.402	.560	.001	.001
Gamma hat[b]	.008	.009	.904	.938	.002	.002
CAK	.797	.716	.184	.268	.001	.001
CK	.815	.740	.167	.245	.001	.001
Mc centrality[b]	.008	.010	.904	.938	.002	.002
CN	.222	.184	.519	.553	.211	.216
SRMR[c]	.207	.070	.730	.875	.014	.004
RMSEA[b]	.002	.002	.911	.940	.008	.006

Note. NFI = Normed Fit Index; BL86 = Bollen's Fit Index; TLI = Tucker–Lewis Index; BL89 = Bollen's delta; RNI = Relative Noncentrality Index; CFI = Comparative Fit Index; GFI = Goodness-of-Fit Index; AGFI = Adjusted Goodness-of-Fit Index; CAK = Rescaled Akaike's Information Criterion; CK = Cross-validation Index; Mc = McDonald's Centrality Index; CN = Hoelter's Critical *N*; SRMR = Standardized Root Mean Squared Residual; RMSEA = Root Mean Squared Error of Approximation.

[a]In this new simple model, factor correlations were misspecified to be 1.0 (instead of fixed to be zeros as before; structural model misspecification). The severity of misspecification is comparable for the new simple and complex models, as shown in Table 6. [b]In Hu and Bentler (1998), these were characterized as the most sensitive to misspecified factor loadings (complex model). [c]In Hu and Bentler (1998), these were characterized as the most sensitive to misspecified factor covariances (simple model).

covariance matrix. As a result, that conclusion does not appear to be generalizable beyond that specific condition. Consequently, SRMR should not be considered the most sensitive to misspecified structural parameters in general.

CONCLUSIONS

Previously, Hu and Bentler (1998, 1999) concluded that (a) the correlation pattern of the fit indexes suggests a multifactor view for the fit indexes; (b) SRMR was the most sensitive to misspecified factor covariances (i.e., misspecified structural model parameters); (c) a group of other fit indexes (e.g., TLI, BL89, RNI, CFI, Gamma, Mc, or RMSEA) were the most sensitive to misspecified factor loadings (i.e., misspecified measurement model parameters); and (d) a two-index strategy is recommended for model fit assessment: SRMR is needed (for detecting misspecified structural model components), and it should be supplemented by another index (e.g., TLI, BL89, RNI, CFI, Gamma, Mc, or RMSEA) that is sensitive to misspecified measurement model components.

Based on our reasoning and the empirical findings presented here, we conclude: (a) there is insufficient evidence to support the multifactor view for the fit indexes; (b) SRMR is not generally most sensitive to misspecified factor covariances (structural model misspecification), and (b) the group of indexes (TLI, BL89, RNI, CFI, Gamma, Mc, or RMSEA) are not more sensitive to misspecified factor loadings. Consequently, the validity of the rationale for the proposed two-index strategy is in question.

We argued in this article that the design problems in the previous studies (Hu & Bentler, 1998, 1999) compromised the validity of the rationale that led to the two-index strategy proposal. First, types of model misspecification (misspecified factor covariance vs. misspecified factor loadings) and the severity of model misspecification were confounded, and this confounding led to the incorrect conclusion that some indexes were differentially sensitive to different types of model misspecification.

Second, the misspecified factor covariances in the simple model represented a somewhat unusual type of misspecification: A large number of covariances in the model-based covariance matrix were forced to be zeros by one or two misspecified factor covariances. SRMR appears to be sensitive to this condition, thus leading to the conclusion that it was the most sensitive to misspecified factor covariances in general. This type of misspecification might not be representative of models with misspecified structural parameter(s), and as a result, the conclusion would not generalize to other kinds of structural parameter misspecification.

We partially replicated the study by Hu and Bentler (1998) to reevaluate the validity of the rationale of the proposed two-index strategy. In the study design, there were two important changes: (a) Two types of model misspecifications (misspecified factor covariances vs. misspecified factor loadings) had comparable severity of misspecification, and (b) misspecified factor covariances did not result in a large number of zeros in the model-based covariance matrix. Empirical findings showed that the previous conclusions concerning the proposed two-index strategy were not supported. Consequently, SEM researchers should reconsider the applicability of the two-index strategy in model fit assessment in SEM applications.

It should be emphasized that this study examined the rationale for the two-index strategy as proposed in Hu and Bentler (1998). Hu and Bentler (1999) later extended the line of research by specifying cutoff values of different indexes, and examined the optimal use of the two-index strategy in making model rejection decisions. This study did not evaluate the actual application of the two-index strategy with regard to using specific cutoff values in making model rejection decisions (Hu & Bentler, 1999).

It is also important to point out that, although the two-index strategy and the later proposed fit index cutoff values have become extremely popular (see more detailed discussion on this issue in Marsh et al., 2004), Hu and Bentler (1998) were cautious about potential overgeneralization of their findings, as they discussed that "the performance of fit indices is complex and that additional research with a wider class of models and conditions is needed, to provide final answers on the relative merits of many of these indices" (p. 446). The findings of this study indicate that this caution is warranted. As suggested by Marsh et al. (2004), many SEM practitioners have disregarded Hu and Bentler's (1998) caution, and overgeneralized Hu and Bentler's findings and tentative conclusions.

The reasoning and findings in this article questioned the rationale of the two-index strategy as advocated in Hu and Bentler (1998, 1999), but some similarities were observed with regard to the performance of some indexes across the two studies. Hu and Bentler (1998) showed that a group of fit indexes (e.g., NFI, BL86, GFI, AGFI) were more sensitive to sample size condition (i.e., larger η^2 values attributable to sample size variation), an undesirable feature of a model fit index. This finding was largely replicated in this study, as shown by the relatively larger η^2 values associated with these indexes in Tables 4, 5, and 7.

Although tempting to conclude that these fit indexes were less useful because of this undesirable characteristic, we refrain from drawing such a definite conclusion based on these findings. As discussed in Marsh et al. (2004), the majority of the misspecified models considered in Hu and Bentler (1998, 1999) were actually acceptable models with very minor degrees of misspecification (acceptable misspecified models). The η^2 value is a relative term. It is possible that an index is not sensitive to very minor model misspecifications, and as a result, it may appear that factors other than model misspecification contribute to its sampling variation. However, when model misspecification becomes more unacceptable, the situation may change. Before we discount this group of indexes as being less useful, the performance of these indexes should be further evaluated by involving more severe model misspecification conditions (unacceptable misspecified models).

Limitations

This study has some obvious limitations. First, although the confounding between severity of model misspecification and types of model misspecification was tackled, research was limited to replicating previously used models; as a result, we did not study a wider class of models. This means that the generalizability of the findings in this study can be potentially limited. Future research may extend to different and more complex types of models (e.g., from CFA models to full structural equation models with latent exogenous and endogenous variables).

Second, we concluded that SRMR may be sensitive to a large number of zero covariances in the model-based covariance matrix. This finding, however, is not analytically based, but empirically based. It is possible that this finding and its related conclusion may not be able to stand the test of further scrutiny as future research involves a wider class of models and model misspecification conditions. Future research in this area may examine this issue more closely.

Third, in this study, a narrow focus on the issue of differential sensitivity of fit indexes to different types of model misspecification was chosen. As such, this study did not address many other issues studied by Hu and Bentler (1998, 1999), such as those related to data distribution (e.g., nonnormality), alternative estimation methods (e.g., asymptotic distribution-free method [ADF]), different combination rules involving fit indexes in model fit assessment, and so on.

This conclusion about the questionable validity of the rationale of the two-index strategy does not mean that we advocate the use of a single fit index in model fit assessment, instead of looking for convergence of multiple indexes. It simply means that the validity of the specific two-index strategy as proposed by Hu and Bentler (1998) is questionable. Because SEM researchers do not fully understand the strengths and weaknesses of the individual fit indexes, the reliance on a single index is more likely to lead to incorrect conclusions about model fit than relying on the triangulation of several fit indexes. For this reason, the use of multiple fit indexes in model fit assessment makes good sense. However, to use multiple fit indexes successfully in model fit assessment, we need to demonstrate that the information provided by different indexes is complementary rather than redundant. The work by Hu and Bentler (1998, Table 3) has laid some groundwork for research in this direction.

Model fit assessment in SEM is a complicated issue that does not appear to have clear-cut solutions at this time. The validity and generalizability of the findings from any particular study, including the one presented in this article, should be evaluated in broader contexts, such as using different models with varying model complexity, with different parameter values, and under different data conditions.

REFERENCES

Bentler, P. M. (1995). *EQS structural equations program manual.* Encino, CA: Multivariate Software.

Bollen, K. A. (1986). Sample size and Bentler and Bonett's nonnormal fit index. *Psychometrika, 51*, 375–377.

Corten, I. W., Saris, W. E., Coenders, G., van der Veld, W., Aalberts, C. E., & Kornelis, C. (2002). Fit of different models for multitrait-multimethod experiments. *Structural Equation Modeling, 9,* 213–232.

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

Enders, C., & Finney, S. (2003, April). *SEM fit index criteria re-examined: An investigation of ML and robust fit indices in complex models.* Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Fan, X., Felsovalyi, A., Sivo, S. A., & Keenan, S. (2002) *SAS for Monte Carlo studies: A guide for quantitative researchers.* Cary, NC: SAS Institute.

Fan, X., Thompson, B., & Wang, L. (1999). The effects of sample size, estimation methods, and model specification on SEM fit indices. *Structural Equation Modeling, 6,* 56–83.

Fan, X., & Wang, L. (1998). Effects of potential confounding factors on fit indices and parameter estimates for true and misspecified SEM models. *Educational and Psychological Measurement, 58,* 699–733.

Gerbing, D. W., & Anderson, J. C. (1993). Monte Carlo evaluations of goodness-of-fit indices for structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 40–59). Newbury Park, CA: Sage.

Glaser, D. N. (2001). [Review of *A first course in structural equation modeling* by Tenko Raykov & George A. Marcoulides]. *Structural Equation Modeling, 8,* 316–323.

Glaser, D. N. (2002). [Review of *Structural equation modeling: Foundation and extensions* by David Kaplan]. *Structural Equation Modeling, 9,* 141–150.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3,* 424–453.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55.

Jöreskog, K. G., & Sörbom, D. (1981). *LISREL V: Analysis of linear structural relationships by the method of maximum likelihood.* Chicago: National Educational Resources.

Kaiser, H. F., & Dickman, K. (1962). Sample and population score matrices and sample correlation matrices from an arbitrary population correlation matrix. *Psychometrika, 27,* 179–182.

Marsh, H. W., Balla, J. R., & Hau, K. T. (1996). An evaluation of incremental fit indices: A clarification of mathematical and empirical properties. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 315–353). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11,* 320–341.

Moulder, B. C., & Algina, J. (2002). Comparison of methods for estimating and testing latent variable interactions. *Structural Equation Modeling, 9,* 1–19.

Pomplun, M., & Omar, M. H. (2001). Score comparability of a state reading assessment across selected groups of students with disabilities. *Structural Equation Modeling, 8,* 257–274.

Saris, W. E., & Satorra, A. (1993). Power evaluations in structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 181–204). Newbury Park, CA: Sage.

Satorra, A., & Saris, W. E. (1983). The accuracy of a procedure for calculating the power of the likelihood ratio test as used within the LISREL framework. In C. O. Middendorp (Ed.), *Sociometric research 1982* (pp. 129–190). Amsterdam: Sociometric Research Foundations.

Yu, C., & Muthén, B. (2002, April). *Evaluation of the model fit indices for latent variable models with categorical and continuous outcomes*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.