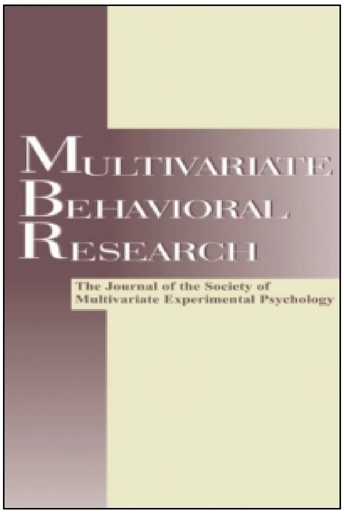

Multivariate Behavioral Research
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t775653673

Sensitivity of Fit Indices to Misspecification in Growth Curve Models

Wei Wu[a]; Stephen G. West[b]
[a] University of Kansas, [b] Arizona State University,

Online publication date: 07 June 2010

To cite this Article Wu, Wei and West, Stephen G.(2010) 'Sensitivity of Fit Indices to Misspecification in Growth Curve Models', Multivariate Behavioral Research, 45: 3, 420 — 452
To link to this Article: DOI: 10.1080/00273171.2010.483378
URL: http://dx.doi.org/10.1080/00273171.2010.483378

Multivariate Behavioral Research, 45:420–452, 2010

ISSN: 0027-3171 print/1532-7906 online
DOI: 10.1080/00273171.2010.483378

Psychology Press
Taylor & Francis Group

Sensitivity of Fit Indices to Misspecification in Growth Curve Models

Wei Wu
University of Kansas

Stephen G. West
Arizona State University

This study investigated the sensitivity of fit indices to model misspecification in within-individual covariance structure, between-individual covariance structure, and marginal mean structure in growth curve models. Five commonly used fit indices were examined, including the likelihood ratio test statistic, root mean square error of approximation, standardized root mean square residual, comparative fit index, and Tucker-Lewis Index. The fit indices were found to have differential sensitivity to different types of misspecification in either the mean or covariance structures with severity of misspecification controlled. No fit index was always more (or less) sensitive to misspecification in the marginal mean structure relative to those in the covariance structure. Specifying the covariance structure to be saturated can substantially improve the sensitivity of fit indices to misspecification in the marginal mean structure; this result might help identify the sources of specification error in a growth curve model. An empirical example of children's growth in math achievement (Wu, West, & Hughes, 2008) was used to illustrate the results.

Growth curve modeling (GCM) is developing into one of the more important analytic approaches in the behavioral sciences. It has been widely applied in many areas of psychology, including clinical, developmental, educational, personality, and learning and memory. Evaluating model fit is an important theoretical issue in GCM. A misspecified model could provide biased parameter estimates,

Correspondence concerning this article should be addressed to Wei Wu, University of Kansas, Department of Psychology, 1415 Jayhawk Boulevard, Fraser 411, Lawrence, KS 66045-7556. E-mail: wwei@ku.edu

leading to incorrect model inferences. One of the advanced methodologies used to estimate GCMs is structural equation modeling (SEM). One advantage of SEM is that it provides a variety of fit indices. Those fit indices, either based on the likelihood ratio test or residuals, tend to reflect the overall fit of a model by measuring the consistency between the model-implied and observed (a) mean responses and (b) covariance matrix (Singer & Willett, 2003; Wu, West, & Taylor, 2009). Evaluating model fit for GCMs is challenging. Misspecifications in GCMs can potentially come from the mean, covariance, or both structures (Wu et al., 2009). The great majority of past research on SEM-based fit indices has examined the performance of fit indices in detecting misspecification in the covariance structure, with only occasional studies examining their performance when misspecification in mean structure has occurred. There is little information about how fit indices will reflect the misspecification in only the mean structure or how joint misspecification in the mean and covariance structures may interact to affect model fit. The purpose of this article is to examine the sensitivity of five commonly used SEM-based fit indices that have shown good performance to several possible sources of misspecification in GCMs. We also investigated a possible way to differentiate misspecifications in mean and covariance structure. Finally, we used data from a longitudinal study of children's growth in math achievement to illustrate the approach. We hope that our study will provide a better understanding of the performance of these fit indices in growth curve models and help researchers locate potential misspecifications and further improve their model.

THE SEM APPROACH TO MODELING GROWTH: A BRIEF REVIEW

Model Specification

Meredith and Tisak (1990) showed that SEM can be used for GCMs if growth parameters (e.g., slope and intercept for linear GCMs) are treated as latent variables and repeated measures are treated as multiple indicators of the latent variables. A GCM is usually expressed in the following matrix form:

$$\mathbf{Y} = \boldsymbol{\Lambda} \times \boldsymbol{\eta} + \mathbf{e}$$

$$E(\boldsymbol{\eta}) = \boldsymbol{\alpha}, E(\mathbf{e}) = 0, COV(\boldsymbol{\eta}, \boldsymbol{\eta}') = \boldsymbol{\Phi}, COV(\mathbf{e}, \mathbf{e}') = \boldsymbol{\Psi}, \text{ and } COV(\boldsymbol{\eta}, \mathbf{e}) = 0. \quad (1)$$

Here $\mathbf{Y}$ represents a vector of repeated measures variables, $\boldsymbol{\eta}$ represents a vector of latent variables (growth parameters), and $\mathbf{e}$ represents a vector of residuals. $\boldsymbol{\Lambda}$ is a loading matrix. $\boldsymbol{\alpha}$ is a vector of the means of the latent variables. $\boldsymbol{\Phi}$ is the covariance matrix among the latent variables, which is also called the *between-individual covariance matrix* because it captures individual differences in growth parameters. $\mathbf{\Psi}$ is the covariance matrix of residuals that cannot be explained by the growth model, which is often called the *within-individual covariance matrix*. It is assumed that the latent variables and residuals are independent of each other. These matrices are then combined to calculate the model-implied marginal mean ($\hat{\boldsymbol{\mu}}$) and covariance matrices ($\hat{\mathbf{\Sigma}}$).

$$\hat{\boldsymbol{\mu}} = E(\mathbf{Y}) = \mathbf{\Lambda}\boldsymbol{\alpha} \tag{2}$$

$$\hat{\mathbf{\Sigma}} = COV(\mathbf{Y}, \mathbf{Y}') = \mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda} + \mathbf{\Psi} \tag{3}$$

Maximum Likelihood Estimator

We briefly introduce the maximum likelihood estimator because it is the basis for likelihood ratio tests of overall model fit and ultimately for the calculation of practical fit indices. Maximum likelihood aims to find the parameter estimates that maximize the likelihood of a given set of sample data given a specified distribution in the population. Under multivariate normality, the maximum likelihood function can be written as a discrepancy function (see Equation 4) in which the first term represents the discrepancy between the sample and model-implied vector of marginal means and the second term represents the discrepancy between the sample and model-implied covariance matrix. Note that for longitudinal data balanced on time with complete data, the mean vector and covariance matrix are assumed to be homogeneous across all of the individuals.

$$F_{ML} = (\overline{\mathbf{Y}} - \hat{\boldsymbol{\mu}})'\hat{\mathbf{\Sigma}}^{-1}(\overline{\mathbf{Y}} - \hat{\boldsymbol{\mu}}) + \ln|\hat{\mathbf{\Sigma}}| - \ln|\mathbf{S}| + tr\hat{\mathbf{\Sigma}}^{-1}\mathbf{S} - p, \tag{4}$$

where p is the number of repeated measures for each individual, $\hat{\mathbf{\Sigma}}$ is the model-implied covariance matrix, $\hat{\boldsymbol{\mu}}$ is the model-implied mean vector, $\overline{\mathbf{Y}}$ is the sample mean vector, and $\mathbf{S}$ is the sample covariance matrix.

Sources of Misfit

Wu et al. (2009) identified four different sources of potential misspecification(s) in GCMs: marginal mean structure, conditional mean structure, within-individual covariance structure, and between-individual covariance structure.

Misspecification in mean structures. If no time-varying or time-constant covariates are considered, the marginal mean structure refers to the specification of the mean growth trajectory and the conditional mean structure refers to the specification of individual growth trajectories. The mean and individual growth trajectories can take different forms in practice. For example, the individual growth trajectories might be quadratic; however, the average growth trajectory might be linear because the curvatures of individual trajectories get averaged out (Wu et al., 2009). In this case, even though the mean quadratic parameter is zero, there may be significant individual variability in the quadratic parameter.

Misspecification in covariance structures. The between-individual covariance structure refers to the specification of the variances and covariances among the growth parameters. In practice, the variance of some growth parameters might be wrongly constrained to be zero or the covariances between growth parameters might not be correctly specified. The within-individual covariance structure refers to the variances and covariances of residuals that cannot be explained by a specified GCM. The residuals are often assumed to have constant variance over time and no correlation with each other. However, in practice, the variances may differ across time. In addition, there might be relationships among the residuals that can not be accounted for by the time variable (Marsh & Hau, 1996; Sivo, Fan, & Witta, 2005; Wu et al., 2009). As shown earlier in Equation 3, the model-implied covariance matrix, which we also term the total covariance matrix in the following, is a combination of the between- and within-covariance matrices.

SEM-Based Fit Indices

Our study focuses on five commonly used SEM-based fit indices: the normal theory maximum likelihood-based likelihood ratio test statistic (T_{ML}), root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), comparative fit index (CFI), and Tucker-Lewis Index (TLI). We identified commonly used fit indices based on a computer and manual search of American Psychological Association journals (Taylor, 2008), subsequently supported by a published computer search of the PsychINFO database by Jackson, Gillaspy, and Purc-Stephenson (2009). From these, we selected fit indices that have shown good performance in simulation studies (e.g., Hu & Bentler, 1998; Marsh, Balla, & McDonald, 1988; Yu, 2002; see West, Taylor, & Wu, in press, for a review). The calculation formula and basic properties of each of the five fit indices are summarized in Table 1. RMSEA, CFI, and TLI are defined through T_{ML}, whereas SRMR is defined through residuals (Sun, 2005; Yuan, 2005).

T_{ML}, RMSEA, and SRMR are absolute fit indices because they evaluate the fit of the hypothesized model without comparison to a baseline model. In contrast, CFI and TLI are relative fit indices because they assess the specific improvement in fit of the hypothesized model relative to a baseline model (Bollen & Curran, 2006). Specification of an acceptable baseline model for relative fit indices is important for their interpretation. The standard baseline model (independence model) used by nearly all SEM packages (e.g., LISREL, Mplus) is *not* appropriate for commonly used GCMs (Widaman & Thompson, 2003).

TABLE 1
The Properties of the Five Fit Indices for Covariance Structure Models

Fit Indices	*Cutoff Criteria (Hu & Bentler, 1999)*	*Sensitive to Model Misspecification*	*Overreject Trivially Misspecified Models?*	*Sensitive to N*
$T_{ML} = (N-1)F_{ML}$ Bollen (1989)	p-value $\leq .05$	Yes	Yes	Yes
$RMSEA^a = \sqrt{\max(d_h/df_h, 0)}$ Steiger & Lind (1980)	$RMSEA \leq .06$	Yes	No	Yes to small N
$SRMR^b = \sqrt{\sum_j \sum_k r_{jk}^2 / p*}$ Jöreskog and Sörbom (1981)	$SRMR \leq .08$	Yes	No	No
$CFI = 1 - \max[d_h, 0] / \max[d_h, d_b, 0]$ Bentler (1990)	$CFI \geq .95$	Yes	No	No
$TLI = [(\chi_b^2/df_b) - (\chi_h^2/df_h)]/[(\chi_b^2/df_b) - 1$ Tucker & Lewis (1973)	$TLI \geq .95$	Yes	No	Yes to small N

Note. T_{ML} = maximum likelihood ratio test statistic; *RMSEA* = root mean square error of approximation; *SRMR* = standardized root mean square residual; *CFI* = comparative fit index; *TLI* = Tucker-Lewis Index; h = hypothesized model; b = baseline model.

[a] $d = (\chi^2 - df)/(N-1)$, where $\chi^2 - df$ is the sample estimate of the noncentrality parameter.

[b] r_{jk} is a standardized residual from the jth row and kth column of the covariance matrix.

p^* is the number of nonduplicated elements in the covariance matrix.

In our study, we used the correct baseline model for GCMs to calculate the CFI and TLI (described in more detail later).

Wu et al. (2009) argued that SEM-based fit indices directly detect misspecification(s) in the marginal mean and covariance structures but *might not* detect misspecification(s) in the conditional mean structures. T_{ML} is built on the minimized fitting function, which takes into account both the marginal mean and covariance structures (see Equation 4). Similarly, the model fit indices based on T_{ML} (such as RMSEA, CFI, and TLI) should also reflect the fit of the model to these structures. A residual-based fit index (such as SRMR) should reflect fit for the marginal mean structure as long as its calculation takes into account the residuals in the marginal means (deviation of the sample means from the model-implied means). SEM-based fit indices cannot be used to detect the misspecification in the conditional mean structure; they operate at the aggregated matrix level not at the level of individual observed and predicted scores.

LITERATURE REVIEW

Nearly all Monte Carlo studies have been conducted to examine the performance of SEM-based fit indices in detecting misfit in covariance structures in the context of confirmatory factor analysis (CFA) models. Few studies have addressed fit issues in GCMs. We briefly review studies addressing sensitivity of SEM-based fit indices to model misspecification.

Covariance Structure Only Model

Hu and Bentler (1998) investigated the sensitivity of a large set of model fit indices (e.g., RMSEA, SRMR, CFI, and TLI) to underparameterized misspecification in CFA models. They used the proportion of variation (η^2) in a fit index accounted for by a specific misspecification to indicate the sensitivity of the fit index to that misspecification. Two types of misspecification were examined: misspecification in factor covariances and misspecification in factor loadings. They found the RMSEA, CFI, and TLI were more sensitive to misspecification of the factor loadings (η^2 ranged from .70 to .77) than to misspecification of factor covariances (η^2 ranged from .30 to .49). In contrast, the SRMR was more sensitive to misspecification in factor covariances ($\eta^2 = .91$) than to misspecification in factor loadings ($\eta^2 = .65$). Given that SRMR showed sensitivity to a different type of misspecification than the other fit indices, Hu and Bentler (1998, 1999) recommended a two-index strategy of using SRMR in combination with either the CFI or RMSEA to evaluate fit. One limitation of Hu and Bentler's (1998) study is that they did not quantify the severity of misspecification. Fan and Sivo (2005) defined severity of misspecification using the estimated noncentrality parameter and its associated power (Saris & Satorra, 1993; Satorra & Saris, 1985; see also Enders & Finney, 2003). They showed that, after controlling for the severity of misspecification, Hu and Bentler's (1998, 1999) conclusion that the fit indices were differently sensitive to different types of misspecification no longer held. Fan and Sivo (2007) further examined the effect of misspecification on fit indices across five types of covariance structure models: three CFA models and two structural models with exogenous and endogenous latent variables. They controlled the severity of misspecification at two levels (low and high) in each model type using the Saris and Satorra approach. They found that model type influenced a variety of fit indices including RMSEA, TLI, CFI, and SRMR with severity of misspecification and sample size controlled. Model type and severity of misspecification also interacted to affect SRMR. Note that the (a) type of misspecification varied across the (b) different model types, thus the effect of these two factors cannot be clearly distinguished in their simulation study.

Growth Curve Model

Leite and Stapleton (2006) investigated the power of CFI, TLI, RMSEA, and SRMR to reject misspecification in the functional form of growth at different sample sizes, levels of severity, and number of repeated measures. They created misspecification by fitting a linear GCM to data generated from nonlinear (mix of both linear and nonlinear trajectories) or piecewise linear GCMs. They quantified the severity of misspecification using the same power-based approach used in

Fan and Sivo (2005; see Saris & Satorra, 1993; Satorra & Saris, 1985). The severity of misspecification was manipulated at three levels based on $n = 100$: slight (power = .40), moderate (power = .70), and severe (power = .90). Leite and Stapleton correctly noted that the standard baseline model for relative fit indices (CFI and TLI) is not applicable to GCMs; however, their attempted solution did not fully address this issue. On the other hand, they were able to examine RMSEA and SRMR for appropriate models involving both a mean and covariance structure.

They found that RMSEA had >.80 power to detect misspecification in functional form using a criterion of RMSEA $\leq$ .06 (average power = .89, .97, and 1.00 for slight, moderate, and severe misspecification across the other conditions). In contrast, the SRMR had unacceptably low power to reject models with misspecified mean and covariance structures with a cutoff criterion of SRMR $\leq$ .08 (average power = .12, .41, and .36 for slight, moderate, and severe misspecification, respectively, across the other conditions) with these nonmonotonic results likely reflecting the differential sensitivity of the SRMR to misspecification in the mean and covariance structures, which were not separately manipulated in this study. As a comparison, they tested the same linear models when the marginal mean structure was specified as saturated (freely estimating all elements in the vector of marginal means). Under this condition SRMR had higher power to reject the misspecification in covariance structure (average power = .31, .71, and .83 for slight, moderate, and severe misspecification, respectively, across the other conditions). The average power for RMSEA under the different levels of severity of misspecification did not change substantially when the marginal mean structure was specified as saturated. Another interesting finding is that SRMR performed inconsistently between the nonlinear and piecewise true growth trajectories. SRMR had more power to detect the misspecification when the true growth trajectory was piecewise rather than nonlinear.

Yu (2002) extended Hu and Bentler's (1998, 1999) studies to GCMs. Using a multivariate normal distribution in the population, he generated data from quadratic GCMs with either five or eight timepoints. He created misspecified models by dropping the quadratic effect from the models. The results showed that the cutoff criteria proposed in Hu and Bentler (1999) were generally suitable for growth models when $N \geq 250$. Based on the suggested cutoff criteria, the fit indices had >.80 power to reject misspecified GCMs when $N \geq 250$ except for SRMR. However, with a cutoff criterion of .08, SRMR had unacceptably low power to reject the misspecified GCM with five timepoints, which was particularly problematic with large N (power = .31 at $N = 1{,}000$). However, SRMR did have adequate power to reject the misspecified GCM with eight timepoints. This latter result may have occurred because the severity of misspecification of the GCM with eight timepoints was so large that all fit indices achieved 1.00 power. None of the fit indices overrejected the true model at greater than the nominal alpha level of .05 except at the smallest sample size, $N = 100$.

QUANTIFYING SEVERITY OF MODEL MISSPECIFICATION

As described earlier, Fan and Sivo (2005, 2007) and Leite and Stapleton (2006) used the power to detect misspecification as a measure of severity of model misspecification. They utilized Saris and Satorra's (1993) two-step approach to estimate the power, which was originally proposed for models with only a covariance structure. Muthén and Curran (1997) extended the Saris and Satorra approach to models including both mean and covariance structures. The first step is to obtain the estimated mean and covariance matrices for the correctly specified model. In this step, one needs to choose the values of parameters in the correctly specified model. The second step is to estimate the misspecified model (with one or more parameters in the correctly specified model constrained) using the estimated mean and covariance matrices from Step 1. The T_{ML} obtained in the second step is used to obtain an approximation of the noncentrality parameter. Then the power to reject the misspecification(s) at a desired level of alpha can be obtained by comparing the noncentral chi-square distribution defined by the noncentrality parameter and the *df* of the misspecified model to a central chi-square distribution with the same degrees of freedom.

Addressing Statistical Limitations

An important statistical limitation of the power-based approach occurs because all parameters other than the misspecified parameter are freely estimated. It is possible that a misspecification in one parameter might be manifested as biased estimate(s) of other parameter(s) without substantially changing the overall estimated mean and covariance matrices (Gerbing & Anderson, 1993; Tomarken & Waller, 2003, 2005). Given that the likelihood ratio test statistic can only detect substantial change in the overall marginal mean and covariance matrices, this approach cannot reflect misspecification in one parameter that is canceled out by the other parameters.

An alternative approach is to use the fixed root mean squared residual (FRMR) to measure the severity of misspecification when the true model is known (i.e., in a simulation study). Olsson, Troye, and Howell (1999) defined FRMR as

$$FRMR = [2\Sigma\Sigma(\sigma - \sigma_F)^2/(p^*)(p^* + 1)]^{0.5},$$

where σ is an element in the population covariance matrix. σ_F is an element in the covariance matrix for a misspecified model with one or more parameters of the true model fixed at zero and other parameters fixed at their population values. Because each parameter other than the misspecified parameter is fixed at its population value, misspecifications in one or more parameters will not spread to the other parameters. However, the FRMR has its own statistical limitations.

First, it reflects only misspecification in the covariance structure. Second, it is not standardized. Thus the severity of a given model misspecification may change as the measurement scale changes.

Following Olsson et al.'s (1999) idea, we proposed using the true model fixed likelihood ratio test statistic (TMFLR) to measure the severity of model misspecification when the true model is known. To calculate TMFLR, the misspecified model is fit to the population mean and covariance matrix with the parameters other than the misspecified parameter fixed at their population value. The likelihood ratio test (T_{ML}) obtained in this way is the TMFLR. Similar to FRMR, TMFLR can prevent misspecifications in one or more parameters from spreading to the other parameters. But in contrast to FRMR, TMFLR is able to take into account misspecification(s) in the marginal mean structure as well as those in the covariance structure. In addition, as a likelihood ratio test statistic, it is comparable across different measurement scales. We used the TMFLR in this study to provide an index of the severity of misspecification (an often neglected feature in previous research) that avoids the statistical limitations associated with power-based approaches and the FRMR.

A second important limitation of previous research is that only *joint misspecification* in both mean and covariance structures was considered, thus one is not able to reach a conclusion about how fit indices reflect misspecifications in only the mean structure and or only the covariance structure. In this study, we considered the separate effects of misspecification in the mean structure and covariance structure as well as their combined effects. Several different general forms of misspecification can occur in the covariance structure; we created one exemplar of each of these general forms to probe their effects.

A final statistical limitation of most of the studies is that they did not use the correct baseline model in the calculation of relative fit indices for GCMs. Thus the performance of the relative fit indices in detecting model misspecification based on the correct baseline model is unknown.

Philosophical Limitation

Two philosophical traditions have been implicitly used in the evaluation of the fit of statistical models. The first tradition follows a straightforward tack: Act as if the proposed model is true and evaluate the effect on the fit of the model as known misspecifications are introduced. The second tradition follows George Box's (1979) dictum "All models are wrong, some are useful" (p. 202). Browne and Cudeck (1993) have introduced an important variant of this idea for SEM in which there is assumed to be a (hopefully small) error of approximation between the unknown true model and the known tested model in the population. This position is undoubtedly more realistic than the first tradition; unfortunately, there is not a current consensus on how to implement this position in studies of the performance of fit indices as specific misspecifications are introduced. Consequently, in the design of our study we followed the first tradition, which has been widely used in previous studies, deferring fuller consideration of the issue of error of approximation until the discussion.

RESEARCH QUESTIONS

This study focused on five commonly used SEM-based fit indices—T_{ML}, RMSEA, SRMR, CFI, and TLI—that have shown good performance and have been recommended in previous simulation studies (e.g., Hu & Bentler, 1998, 1999; Yu, 2002). Three research questions, the first of which has three subparts, were examined in this study:

1) How sensitive are the five fit indices in detecting different types of misspecification in the structures identified earlier?
   a) How sensitive are the five fit indices in detecting misspecification in only the marginal mean structure?
   b) How sensitive are the five fit indices in detecting different general types of misspecification in the covariance structure? Here both misspecification in the within- and between-individual covariance structures were considered.
   c) Do the fit indices have different differential sensitivity to the misspecification in the marginal mean structure as opposed to misspecification in the covariance structure?
2) If there is misspecification in both the mean and covariance structures, do they interact with each other to affect the SEM-based fit indices?
3) Can saturating one structure improve the sensitivity of fit indices in detecting misspecification in the other structure? In other words, can saturating the covariance structure help detect misspecification in the marginal mean structure? Can saturating the marginal mean structure help detect misspecification in the covariance structure?

METHOD

Design Overview

In this study, three general factors were manipulated in a factorial design. First was the type of misspecification: Tested models were misspecified in (a) the marginal mean structure, or (b) one of four different ways in the covariance structure (see the following), or (c) a combination of both. Second was the degree of misspecification with three levels: slight, moderate, and severe, which is defined by the power associated with TMFLR for $N = 250$. Third was the sample size with four levels: 125, 250, 500, and 1,000. In the following, we describe the population models used to generate data and how we created different types of misspecification.

Population Models

All of the data generation models are quadratic GCMs with five timepoints. The general quadratic model is presented as a path diagram in Figure 1. In the general quadratic GCM, we centered time at the third timepoint so that the intercept represents the status of the outcome variable at Time 3. The linear slope

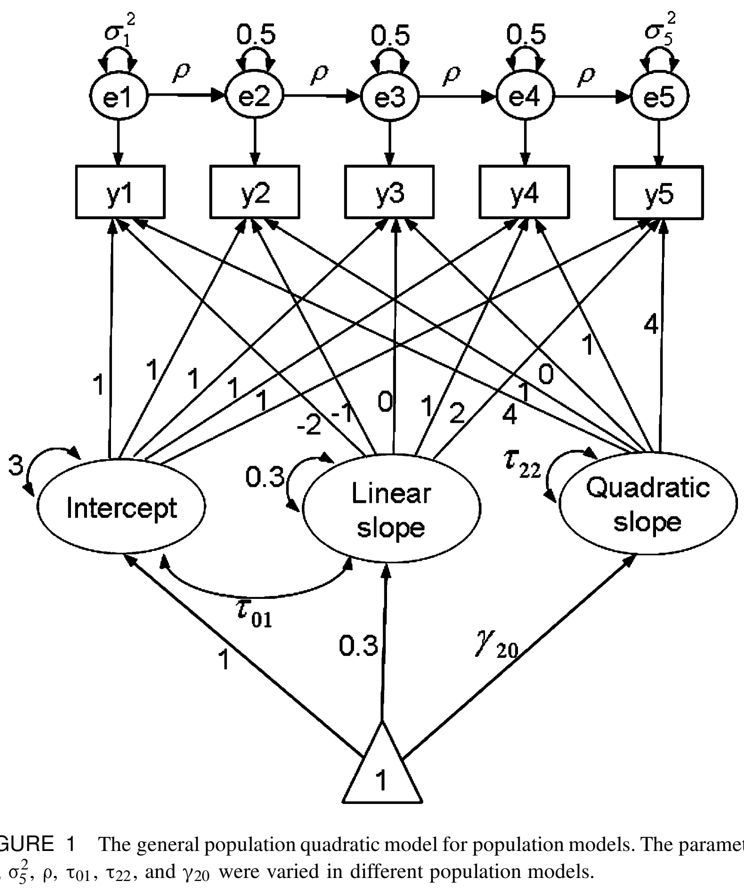

FIGURE 1 The general population quadratic model for population models. The parameters σ_1^2, σ_5^2, ρ, τ_{01}, τ_{22}, and γ_{20} were varied in different population models.

represents the instantaneous slope at Time 3, which is also approximately equal to the average slope over time. The following matrices show the parameters in each population model used to generate the data:

$$\boldsymbol{\Lambda} = \begin{bmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix}, \quad \boldsymbol{\alpha} = \begin{bmatrix} 1 \\ 0.3 \\ \gamma_{20} \end{bmatrix}, \boldsymbol{\Phi} = \begin{bmatrix} 3 & \tau_{01} & 0 \\ \tau_{10} & 0.3 & 0 \\ 0 & 0 & \tau_{22} \end{bmatrix},$$

$$\boldsymbol{\Psi} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1^2 & \rho^2\sigma_1^2 & \rho^3\sigma_1^2 & \rho^4\sigma_1^2 \\ \rho\sigma_1^2 & \rho^2\sigma_1^2 + 0.5 & \rho(\rho^2\sigma_1^2 + 0.5) & \rho^2(\rho^2\sigma_1^2 + 0.5) & \rho^3(\rho^2\sigma_1^2 + 0.5) \\ \rho^2\sigma_1^2 & \rho(\rho^2\sigma_1^2 + 0.5) & \rho^4\sigma_1^2 + \rho^2 0.5 + 0.5 & \rho(\rho^4\sigma_1^2 + \rho^2 0.5 + 0.5) & \rho^2(\rho^4\sigma_1^2 + \rho^2 0.5 + 0.5) \\ \rho^3\sigma_1^2 & \rho^2(\rho^2\sigma_1^2 + 0.5) & \rho(\rho^4\sigma_1^2 + \rho^2 0.5 + 0.5) & \rho^6\sigma_1^2 + \rho^4 0.5 + \rho^2 0.5 + 0.5 & \rho(\rho^6\sigma_1^2 + \rho^4 0.5 + \rho^2 0.5 + 0.5) \\ \rho^4\sigma_1^2 & \rho^3(\rho^2\sigma_1^2 + 0.5) & \rho^2(\rho^4\sigma_1^2 + \rho^2 0.5 + 0.5) & \rho(\rho^6\sigma_1^2 + \rho^4 0.5 + \rho^2 0.5 + 0.5) & \rho^8\sigma_1^2 + \rho^6 0.5 + \rho^4 0.5 + \rho^2 0.5 + \sigma_5^2 \end{bmatrix}$$

In the marginal mean structure ($E(\mathbf{Y}) = \boldsymbol{\Lambda\alpha}$), the parameters are the mean intercept, linear slope, and quadratic parameter. In the between-individual covariance matrix ($\boldsymbol{\Phi}$), the parameters are the variances for the intercept, linear slope, and quadratic parameter and the covariance between the intercept and linear slope. The within-individual covariance matrix ($\boldsymbol{\Psi}$) represents a first-order autoregressive structure with nonconstant residual variances across time (Hedeker & Gibbons, 2006, pp. 116–117). In the matrix, the parameters are residual variances and the autoregressive coefficient between residuals at adjacent timepoints (ρ). With this specification, we allowed the residuals to be correlated with each other and residual variances to vary across time instead of following standard assumptions.

To facilitate the generation of different degrees of misspecification, we varied population values for (a) the variance of the quadratic parameter (τ_{22}), (b) the covariance between intercept and slope (τ_{01}), (c) the residual variances at Times 1 and 5 (σ_1^2 and σ_5^2), (d) the autoregressive coefficient (ρ), and (e) the mean quadratic slope (γ_{20}). As an illustration, suppose a misspecification is created by fixing τ_{22} to be zero. Then as we increase the population value for τ_{22}

in the data generation model, fixing it to be zero would yield more severe misspecification. Given space limitations, we are not able to present all of the population parameter values used to generate the data. The complete list of population parameter values can be found in the supplemental materials (see Acknowledgment). In the aforementioned matrices, the parameter values that are constant for the quadratic models are indicated by a specific numerical value; the parameter values that were varied are indicated by the corresponding symbol for the parameter.

Misspecified Models

Consistent with previous studies, we only considered underspecification in the study. That is, misspecification(s) were created by constraining parameter(s) in the population model used to generate the data (true model).

Misspecifications in the covariance structure. Exemplars representing four different general types of misspecification of the covariance structure in CGM were created:

1. The variance of the quadratic slope was fixed to be 0 ($\tau_{22} = 0$);
2. The covariance between the intercept and linear slope was fixed to be 0 ($\tau_{01} = 0$);
3. The Level 1 residual variances (at Time 1 and Time 5) differ in the population model but they were constrained to be equal in the tested model ($\sigma_1^2 = \sigma_5^2$); and
4. The autoregressive coefficient among the residuals was fixed to be 0 ($\rho = 0$).

For ease of presentation, we term the four general types of misspecification tao22, tao01, sigma, and rho. Our goal was to represent the full range of general types of misspecification. We recognize that in most substantive research contexts tao01 will be freely estimated and many other exemplars of each general type of misspecification could have been used. For each type of misspecification in the covariance structure, we varied its level of severity based on the TMFLR test statistic and the power associated with TMFLR for $N = 250$. There are three levels of TMFLR: 4.90 (slight, power = 0.60), 7.85 (moderate, power = 0.80), 38.00 (severe, power $\approx$ 1.0) for $N = 250$. Note that these power values correspond to the focused one degree of freedom LR test, whereas the fit indices reflect global misspecification in the entire model.

To compare how the sensitivity of fit indices to each of these misspecifications in the covariance structure varied when the marginal mean structure was true as opposed to saturated, we examined the misspecifications under two conditions: (a) misspecified covariance structure and true marginal mean structure (TMMC) and (b) misspecified covariance structure and saturated marginal mean structure (SMMC). To saturate the marginal mean structure,[1] we freely estimated the intercepts for all waves of repeated measures and fixed the means of intercept, linear slope, and quadratic acceleration to be zero. Example Mplus syntax for fitting TMMC and SMMC models can be found in the supplemental materials.

Misspecifications in the marginal mean structure. We included only one variant of misspecification in the marginal mean structure, fixing the mean quadratic parameter to be 0 ($\gamma_{20} = 0$). We term this type of misspecification gamma20. To equate the severity of the misspecification in the marginal mean structure to the severity of those in the covariance structure, we set the severity of the misspecification in the marginal mean structure at the same three levels as those in covariance structure according to the TMFLR and its associated power.

We compared two ways of examining the sensitivity of fit indices to misspecification in the marginal mean structure. First, we examined theoretical models not realizable in practice, which had a misspecified marginal mean structure and true covariance structure (MMTC). Second, we examined more realistic models, which had misspecified marginal mean structure and saturated covariance structure (MMSC). To saturate the covariance structure, we freely estimated all of the elements in the within-individual covariance structure and fixed all of the elements in the between-individual covariance structure to zero. Example Mplus syntax for fitting MMTC and MMSC models can be found in the supplemental materials.

Misspecifications in both the marginal mean and covariance structures. We combined the different levels of misspecification in marginal mean structure and the different types and levels of misspecification in the covariance structure to create models with misspecifications in both the marginal mean and covariance structure (MMMC). There was one type of misspecification in the marginal mean structure with three levels ($1 \times 3 = 3$) combined with each of the four types of misspecification in the covariance structure with three levels ($4 \times 3 = 12$). Thus, there were in total $3 \times 12 = 36$ combinations of misspecifications in both the mean and covariance structures.

[1]There are two ways to specify the marginal mean structure. One is to estimate the intercepts of the repeated measures. The other is to estimate the means of the growth parameters. One cannot estimate the intercepts and the means of the growth parameters simultaneously. Usually, researchers freely estimate the mean growth parameters with intercepts fixed at zero, in which case marginal means are functions of mean growth parameters. To saturate the marginal mean structure, the intercepts need to be freely estimated, but the mean growth parameters must be fixed to be zero, in which case no structure is imposed on the marginal means, and the estimated marginal means will be equal to the sample means.

In summary, there were 6 models with misspecification in only the marginal mean structure: 3 with saturated covariance structure (MMSC) and 3 with true covariance structure (MMTC). There were 24 models with misspecification in only the covariance structure: 12 with saturated marginal mean structure (SMMC) and 12 with true marginal mean structure (TMMC). There were 36 models with misspecifications in both marginal mean and covariance structures (MMMC). In total, 66 misspecified models plus one true model were examined in the study. All of the models were examined under four sample sizes. Thus, in all there were $(1 + 66) \times 4 = 268$ conditions with 1,100 replications of each condition. All replications successfully converged.[2]

Baseline Model for CFI and TLI

We calculated CFI and TLI using an acceptable baseline model. Figure 2 depicts the baseline model for the true marginal mean, true covariance structure model (TMTC), TMMC, MMTC, and MMMC. The baseline model is an intercept-only model in which only the mean of the intercept and the constant unique variance were freely estimated (Widaman & Thompson, 2003). Corresponding to the saturated covariance structure in MMSC and the true marginal mean, saturated covariance structure model (TMSC), we specified the covariance structure in the baseline model to be saturated so that the relative fit indices only reflect the improvement in model fit due to the improvement in the marginal mean structure. Corresponding to the saturated marginal mean structure in SMMC and the saturated marginal mean, true covariance structure model (SMTC), we specified the marginal mean structure in the baseline model to be saturated so that the relative fit indices only reflect the improvement in model fit due to the improvement in the covariance structure. We followed three steps to produce correct relative fit indices using appropriate baseline models: (a) Estimate the model of interest to obtain $T_{ML(h)}$ and df_h; (b) estimate the correct baseline model to obtain $T_{ML(b)}$ and df_b; (c) use the values of $T_{ML(h)}$, df_h, $T_{ML(b)}$, and df_b from the previous steps to calculate CFI and TLI using the equations presented in Table 1, where h stands for hypothesized or tested model and b stands for baseline model.

[2]In some conditions, Heywood cases occur. These represent cases in which the estimated variance for the quadratic rate or the estimated residual variance at Time 1 was negative, common when the variance component is zero or small in the population. Programs simply fix these estimates to zero (as do researchers typically). We calculated the mean and standard deviation of those parameter estimates for which Heywood cases occurred. The mean of the parameter estimates was very close to zero. For example, for the model with severe misspecification in the mean structure and $N = 125$, negative quadratic variances occurred in 389 cases among the 1,100 replications. However, the mean quadratic rate variance across the 1,100 replications was .005 with $SD = .01$.

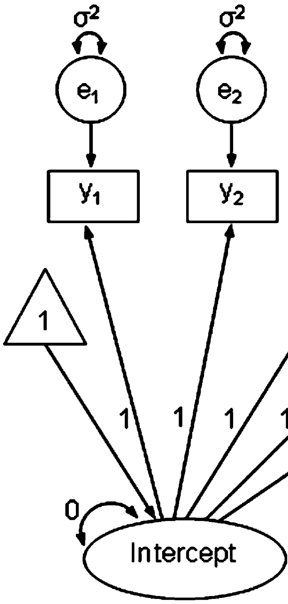

FIGURE 2 The general baseline model used to calculate the Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) in this study. In the baseline model, only the mean intercept and the residual variances that are constant over time are estimated.

Data Analysis

Mplus 5.1 (L. K. Muthén & Muthén, 1998–2006) was used to estimate each of the models. The values for T_{ML}, RMSEA, CFI, TLI, and SRMR were saved for each replication of each of the GCMs. A series of analyses of variance (ANOVAs) were conducted to examine potential factors that might influence the fit indices under the different types of models. For the models with misspecification in only the marginal mean structure (MMSC or MMTC), two design factors were considered: the level of severity of misspecification (low, medium, high) and N. For the models with misspecification in only the covariance structure (SMMC or TMMC), three factors were considered: type of misspecification (tao22, tao01, sigma, and rho), severity of misspecification (low, medium, high), and N. Finally, for the models with joint misspecification in the marginal mean and covariance structures (MMMC), three factors were considered: severity of misspecification in the marginal mean structure, severity of misspecification in the covariance structure, and N. For all of the ANOVAs, we reported the proportion of variation in the fit indices (η^2) that was accounted for by each of the design factors and their interactions. Following Cohen (1988), we used

$\eta^2 = .01$, .06, and .14 to define small, moderate, and large effects, respectively. Any η^2 that was smaller than .01 was deemed ignorable.

RESULTS

True Model

The results for the true model replicated those of past research. The means and standard deviation of each fit index are shown in Table 2. The mean of T_{ML} (9.170) approximated 9 for all Ns as was expected for the chi-square distribution with $df = 9$. The mean values of CFI and TLI were close to 1 (.998 and 1.000, respectively, across the four sample size conditions) and were unaffected by sample size. Mean values of RMSEA and SRMR were close to 0 (.016 and .017, respectively, across the four sample size conditions); mean values decreased as N increased.

Models With Misspecification in Covariance Structure Only

The means of T_{ML}, RMSEA, SRMR, and CFI for different Ns, types, and levels of severity of misspecification in covariance structures are shown in a series

TABLE 2
Means and Standardized Deviations of the Five Fit Indices: True Model

		N			
Fit Index		*125*	*250*	*500*	*1,000*
$T_{ML}{}^a$	*M*	9.132	9.170	9.188	9.190
	SD	4.190	4.246	4.298	4.245
RMSEA	*M*	0.024	0.017	0.012	0.009
	SD	0.031	0.022	0.016	0.011
SRMR	*M*	0.027	0.019	0.013	0.009
	SD	0.009	0.006	0.004	0.003
CFI	*M*	0.994	0.997	0.999	0.999
	SD	0.009	0.005	0.002	0.001
TLI	*M*	1.000	1.000	1.000	1.000
	SD	0.023	0.012	0.006	0.003

[a] $df = 9$.

Note. T_{ML} = maximum likelihood ratio test statistic; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; CFI = comparative fit index; TLI = Tucker-Lewis Index.

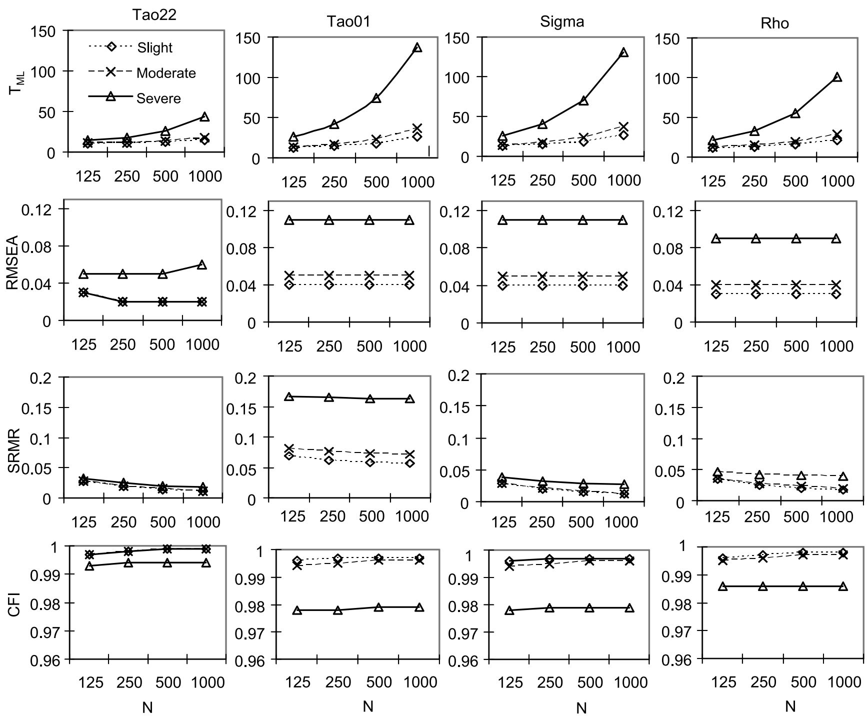

FIGURE 3 The means of fit indices for models with slight, moderate, or severe misspecification in each of the four types of misspecification in covariance structure over sample sizes. *Note.* Tao22 = drop the quadratic rate variance; Tao01 = drop the covariance between intercept and linear rate; Sigma = equate the residual variances at Time 1 and Time 5; Rho = drop the autoregressive coefficient; ◇ = slight misspecification; x = moderate misspecification; △ = severe misspecification.

of panels in Figure 3. The means of TLI across Ns showed a pattern very similar to the means of CFI; they are not plotted in Figure 3. ANOVAs were conducted to examine the effect of type and severity of misspecification in the covariance structure, N, and the two-way and three-way interactions on the fit indices (Table 3).

Main effect of severity of misspecification. Paralleling previous simulation studies of covariance structure models, severity of misspecification accounted for a large amount of variation in T_{ML}, RMSEA, CFI, and TLI ($\eta^2 s$ ranged from .325 to .470) and a moderate amount of variation in SRMR ($\eta^2 =$ .127).

TABLE 3
η^2 Derived From a 4 × 3 × 4 Analysis of Variance (Type of Misspecification × Severity of Misspecification × Sample Size) Preformed on the Models With Misspecification in Covariance Structure With True Marginal Mean Structure

Fit Index	*Type[a] (A)*	*Severity (B)*	*N (C)*	*A × B*	*A × C*	*B × C*	*A × B × C*
T_{ML}	0.069	0.325	0.216	0.053	0.034	0.167	0.027
RMSEA	0.109	0.455	0.001	0.036	0.001	0.001	0.001
SRMR	0.578	0.127	0.013	0.162	0.001	0.001	<0.001
CFI	0.100	0.470	0.004	0.088	<0.001	0.001	<0.001
TLI	0.100	0.459	<0.001	0.078	<0.001	<0.001	<0.001

Note. There are four types of misspecification in covariance structure: tao22 = 0, tao01 = 0, sigma1 = sigma2, and rho = 0. The analyses are based on 1,100 converged replications. Values ≥ .01 are underlined.

T_{ML} = maximum likelihood ratio test statistic; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; CFI = comparative fit index; TLI = Tucker-Lewis Index.

[a]There are four types of misspecification in covariance structure: Tao22 = drop the quadratic rate variance; Tao01 = drop the covariance between intercept and linear rate; Sigma = equate the residual variances at Time 1 and Time 5; Rho = drop the autoregressive coefficient. The analyses are based on 1,100 converged replications. Values .01 are underlined.

Main effect for sample size. T_{ML} was affected by N ($\eta^2 = .216$) across type of misspecification. T_{ML} increased as N increased. For most forms of misspecification, SRMR decreased slightly as N increased ($\eta^2 = .013$). RMSEA, CFI, and TLI were generally not sensitive to N, consistent with previous findings that population based fit indices and relative fit indices are less sensitive to N.

Main effect of type of misspecification. Unique to this study, four types of misspecification (tao22, tao01, sigma, rho) in the covariance structure were investigated. Type of misspecification[3] accounted for a large amount of variation in SRMR ($\eta^2 = .578$) and a moderate amount of variation in T_{ML}, RMSEA, CFI, and TLI ($\eta^2 s$ ranged from .069 to .109). Of interest, all of the fit indices,

[3]Because tao01 is normally freely estimated, we examined how the ANOVA results would be changed by dropping the misspecification of tao01. We conducted a second ANOVA in which the factor of type of misspecification in the covariance matrix had only three levels (tao22, sigma, and rho). The results showed that with tao01 excluded, type of misspecification in the covariance matrix explained slightly more variation in T_{ML}, RMSEA, CFI, and TLI but much less variation in SRMR (η^2 dropped from .578 to .146) compared with the results of the original analysis presented in Table 3. However, dropping tao01 did *not* change the conclusions regarding the main and interaction effects of type, severity of misspecification, and sample size.

but particularly SRMR, showed differential sensitivity to the different types of misspecification in the covariance structure.

Interaction effects. Type and severity of misspecification interacted to affect all of the fit indices ($\eta^2 s$ ranged from .036 to .162). As depicted in Figure 3, the sensitivity of each of the fit indices to the four types of misspecification followed the order presented here:

T_{ML}: tao01, sigma, and rho > tao22;
RMSEA, CFI, and TLI: tao01 and sigma > rho > tao22; and
SRMR: tao01 > sigma and rho > tao22.

As an illustration, T_{ML} was almost equivalently sensitive to misspecification in tao01, sigma, and rho, but it was less sensitive to tao22 than to the other three misspecifications. It appears that SRMR was more sensitive to the misspecification in tao01 and less sensitive to the other three types of misspecification than the other SEM fit indices. All of the fit indices were least sensitive to the misspecification in tao22.

Type of misspecification and N interacted to affect only T_{ML}. T_{ML} was less vulnerable to sample size for tao22 than for the other three types of misspecification. In addition, there was an interaction between severity of misspecification and sample size such that T_{ML} increased more with increasing N for more severely misspecified models for all four types of misspecification.

Models With Misspecification in Only Marginal Mean Structure

The means of each fit index for different Ns and levels of severity of misspecification in marginal mean structure are shown in a series of panels in Figure 4. ANOVAs were conducted to examine the effect of severity of misspecification, N, and their two-way interactions on the fit indices (see Table 4).

Main effect of severity of misspecification. Severity of misspecification accounted for a large amount of variation in all of the fit indices. RMSEA, CFI, and TLI ($\eta^2 s$ ranged from .511 to .528) were more sensitive to misspecification in marginal mean structure than T_{ML} and SRMR ($\eta^2 s = .389$ and .332).

Main effect of sample size. Only T_{ML} and SRMR were affected by N ($\eta^2 = .262$ and .247). T_{ML} increased as N increased; SRMR decreased as N increased; and RMSEA, CFI, and TLI were not affected by N.

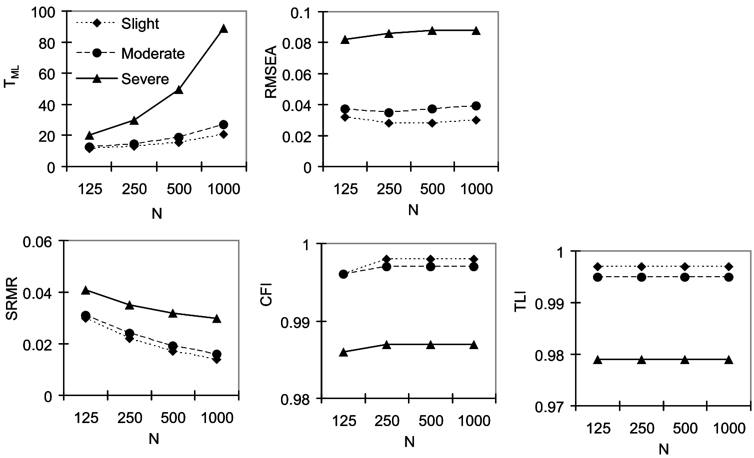

FIGURE 4 The means of fit indices for models with slight, moderate, or severe misspecification in marginal mean structure over sample sizes. T_{ML} = maximum likelihood ratio test statistic; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; CFI = comparative fit index; TLI = Tucker-Lewis Index.

TABLE 4
η^2 Derived From a 3 × 4 Analysis of Variance (Severity of Misspecification × Sample Size) Preformed on the Models With Misspecification in Marginal Mean Structure With True Covariance Structure

Fit Index	*Severity*	*N*	*N × Severity*
T_{ML}	0.389	0.262	0.199
RMSEA	0.517	0.001	0.002
SRMR	0.332	0.247	0.008
CFI	0.528	0.006	0.001
TLI	0.511	<0.001	<0.001

Note. N = 1,100 converged replications. Values ≥ .01 are underlined.

T_{ML} = maximum likelihood ratio test statistic; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; CFI = comparative fit index; TLI = Tucker-Lewis Index.

Interaction effect. N interacted with severity of misspecification to affect T_{ML} ($\eta^2 = .199$). T_{ML} was more sensitive to severity of misspecification when N is large, consistent with previous studies.

Do the Fit Indices Differ in Sensitivity to Misspecification in Marginal Mean Structure Versus Misspecification in Covariance Structure?

We compared the η^2 due to severity of misspecification in marginal mean structure with the η^2 due to severity of misspecification for each of the four types of misspecification in covariance structure (see Figure 5). As shown in Figure 5, none of the fit indices was always more sensitive to the misspecification in the mean structure or to those in the covariance structure. It seems like the misspecification in the mean structure and the misspecification of the autoregressive coefficient had a similar amount of influence on the fit indices. RMSEA, CFI, and TLI were less sensitive to the misspecification in the mean structure than to the misspecification of tao01 and sigma.

Do the Misspecifications in Marginal Mean Structure and Covariance Structure Interact to Affect the Fit Indices?

To answer the question, we examined the models with two simultaneous misspecifications: one in the marginal mean structure (gamma20) and the other in the covariance structure (one of the four types of misspecification: tao22,

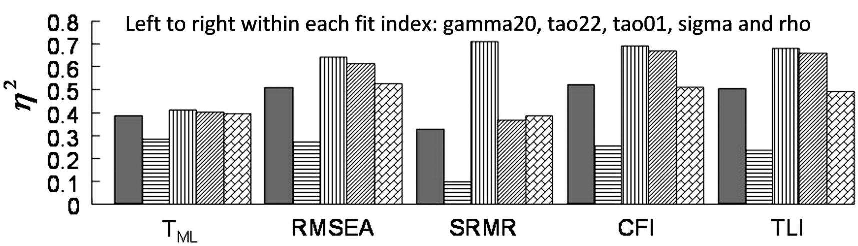

FIGURE 5 The proportion of variation (η^2) in each of the fit indices accounted for by the misspecification in marginal mean structure versus the proportion of variation in each of the fit indices accounted for by the four types of misspecification in the covariance structure. *Note.* From left to right within each fit index: Gamma20 = fix the mean quadratic rate at zero, tao22 = drop the quadratic rate variance, tao01 = drop the covariance between intercept and linear rate, sigma = equate the residual variances at Time 1 and Time 5, rho = drop the autoregressive coefficient; T_{ML} = maximum likelihood ratio test statistic; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; CFI = comparative fit index; TLI = Tucker-Lewis Index.

TABLE 5
η^2 Derived From a 3 × 3 × 4 Analysis of Variance (Severity of Misspecification in Mean Structure × Severity of Misspecification in Covariance Structure × Sample Size) Performed on the Models With Misspecification in Both Mean Structure and Covariance Structures

Fit Index	*Misspecification in Marginal Mean*	*Misspecification in Covariance*	*Severity (Mean) (A)*	*Severity (Covariance) (B)*	*N (C)*	*A × B*	*A × C*	*B × C*	*A × B × C*
T_{ML}	Gamma20	Tao22	0.256	0.034	0.357	0.001	0.137	0.019	<.001
		Tao01	0.072	0.235	0.429	0.006	0.037	0.120	0.003
		Sigma	0.068	0.226	0.435	0.009	0.035	0.116	0.005
		Rho	0.142	0.184	0.460	0.001	0.073	0.095	0.001
		Average	0.135	0.170	0.420	0.004	0.070	0.087	0.002
RMSEA	Gamma20	Tao22	0.356	0.061	0.003	0.008	0.001	0.001	<.001
		Tao01	0.157	0.412	0.002	0.031	<.001	<.001	<.001
		Sigma	0.146	0.384	0.002	0.035	<.001	<.001	<.001
		Rho	0.340	0.420	0.002	0.005	<.001	<.001	<.001
		Average	0.250	0.319	0.002	0.020	<.001	<.001	<.001
SRMR	Gamma20	Tao22	0.307	0.088	0.151	0.006	0.005	0.001	<.001
		Tao01	0.009	0.732	0.006	0.005	<.001	0.001	<.001
		Sigma	0.246	0.144	0.187	0.002	0.004	0.001	<.001
		Rho	0.220	0.408	0.068	<.001	0.001	0.003	<.001
		Average	0.196	0.343	0.103	0.003	0.003	0.002	<.001
CFI	Gamma20	Tao22	0.384	0.058	0.003	0.001	0.001	0.001	<.001
		Tao01	0.136	0.458	0.001	0.012	<.001	<.001	<.001
		Sigma	0.110	0.461	0.002	0.014	<.001	<.001	<.001
		Rho	0.393	0.374	<.001	0.001	<.001	<.001	<.001
		Average	0.256	0.338	0.002	0.007	<.001	<.001	<.001
TLI	Gamma20	Tao22	0.371	0.058	<.001	0.001	<.001	<.001	<.001
		Tao01	0.136	0.452	<.001	0.013	<.001	<.001	<.001
		Sigma	0.111	0.454	<.001	0.015	<.001	<.001	<.001
		Rho	0.392	0.373	<.001	<.001	<.001	<.001	<.001
		Average	0.253	0.334	<.001	0.007	<.001	<.001	<.001

Note. The analyses are based on 1,100 converged replications. Values ≥ .01 are underlined.

T_{ML} = maximum likelihood ratio test statistic; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; CFI = comparative fit index; TLI = Tucker-Lewis Index. Gamma20 = fix the mean quadratic rate at 0; Tao22 = drop the quadratic rate variance; Tao01 = drop the covariance between intercept and linear rate; Sigma = equate the residual variances at Time 1 and Time 5; and Rho = drop the autoregressive coefficient. The analyses are based on 1,100 converged replications. Values .01 are underlined.

tao01, sigma, and rho). The severity (degree) of misspecification in the marginal mean and covariance structures were comparable but manipulated independently. ANOVAs were conducted separately for each combination to examine how the fit indices were affected by the severity of misspecification in marginal mean structure, the severity of misspecification in the covariance structure, and the interaction between the severity of misspecification in the two structures. The ANOVA results are summarized in Table 5.

The main effects of severity of misspecification in marginal mean structure and covariance structure. In general, severity of misspecification in the covariance structure accounted for a larger proportion of variation in the fit indices (averages η^2 ranged from .170 to .383) than severity of misspecification in the marginal mean structure (averages $\eta^2 s$ ranged from .135 to .256). The exceptions include all of the fit indices when tao22 was misspecified, SRMR when sigma was misspecified, and CFI and TLI when rho was misspecified. For those cases, the severity of misspecification in marginal mean structure accounted for more variation in those fit indices. When the source of the misspecification in the covariance structure was tao01, SRMR tended to be affected *only* by severity of misspecification in the covariance structure ($\eta^2 =$ .732).

Note that when both gamma20 and tao22 were fixed at zero, the misspecified model is equivalent to a linear GCM. Otherwise stated, the misspecification is equivalent to fitting a linear functional form to the data generated from a quadratic GCM. In this case, all of the fit indices were more determined by the misspecification in the marginal mean structure.

Interaction effect. The interaction between severity of misspecification in the two structures accounted for a small amount of variation in RMSEA ($\eta^2 s =$ .031 and .035) and CFI ($\eta^2 s = .012$ and .014) when the misspecification in covariance structure was in tao01 or sigma. As the severity of misspecification in tao01 or sigma increased, the sensitivity of RMSEA and CFI to the misspecification in the marginal mean structure decreased, and vice versa. In other words, the more severe the misspecification was in one structure, the less sensitive the two fit indices were to the misspecification in the other structure.

Consistency Among SEM-Based Fit Indices

We examined the correlations among the five fit indices for different types of models. RMSEA, CFI, and TLI showed high consistency with each other for all types of models examined in the study ($|r|$ ranged from .87 to .97). Generally, T_{ML} was more consistent with RMSEA, CFI, and TLI than with SRMR. SRMR showed high consistency with RMSEA, CFI, or TLI only when tao01 was misspecified ($|r|$ ranged from .92 to .94 in this condition).

Can Saturating Marginal Mean Structure Improve the Sensitivity of the Fit Indices to Misspecification in the Covariance Structure?

To answer this question, we saturated the marginal mean structure for those models with misspecification in only the covariance structure. Paralleling the previous analyses, we conducted ANOVAs with three factors: type, severity of misspecification, and N. We compared the η^2 due to severity of misspecification in the covariance structure between models with correctly specified marginal mean structure versus saturated marginal mean structure. This comparison indicated that specifying marginal mean structure to be saturated did not change the sensitivity of the fit indices to misspecification in the covariance structure except for SRMR. Saturating the marginal mean structure *slightly* increased the average sensitivity of SRMR to the misspecifications in the covariance structure.

Can Saturating the Covariance Structure Improve the Sensitivity of the Fit Indices to Misspecification in the Marginal Mean Structure?

To answer this question, we saturated the covariance structure for those models with misspecification in only the marginal mean structure. Paralleling the previous analyses, we conducted ANOVAs with two factors: severity of misspecification and N. Figure 6 compared the η^2 due to severity of misspecification in marginal mean structure between the models with correctly specified covariance structure versus saturated covariance structure. The results show that the misspecification in marginal mean structure accounted for more variation in RMSEA, SRMR, CFI, and TLI (η^2 ranged from .541 to .699) when the covariance structure was saturated rather than correctly specified (η^2 ranged from .332 to .528). The only exception was T_{ML}. Misspecification in marginal mean structure accounted for less variation in T_{ML} ($\eta^2 = .249$) when the covariance structure was saturated rather than when it was correctly specified ($\eta^2 = .389$). In addition, with a saturated covariance structure, T_{ML} was more sensitive to

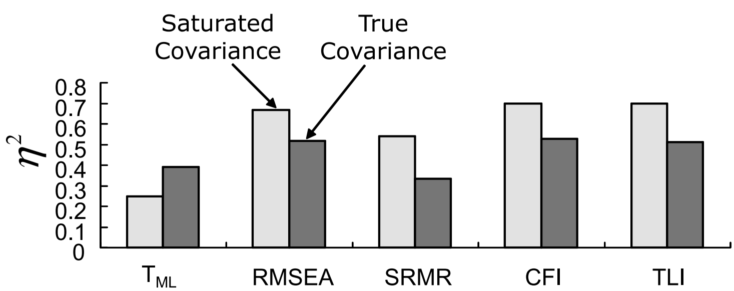

FIGURE 6 Comparison of proportion of variation in each of the fit indices explained by misspecification in marginal mean structure when the covariance structure was correctly specified versus when the covariance structure was saturated. T_{ML} = maximum likelihood ratio test statistic; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; CFI = comparative fit index; TLI = Tucker-Lewis Index.

N ($\eta^2 = .568$). It is also worth noting that the effect of N on SRMR was substantially reduced when the covariance structure was saturated (η^2 due to N was reduced from .247 to .011).

AN EMPIRICAL EXAMPLE

An empirical example was used to illustrate how saturating the covariance structure can help detection of incorrect specification of the functional form for average growth trajectory (marginal mean structure). This example analyzed one condition from the longitudinal data set used in Wu, West, and Hughes (2008). In this study, elementary school students were measured on their math achievement once per year for 4 years from Grade 1 to Grade 4. At the end of the Grade 1, some of the students ($N = 165$) were retained in Grade 1. Based on both theoretical and empirical considerations, we hypothesize that a quadratic growth curve model would fit these data. One complication in our illustration arises because some observations were missing in the original data set. Using SAS PROC MI (SAS Institute, Inc., 1999), we created a single complete data set with the missing data imputed to prevent confounding the effect of missing data on model fit (see Wu et al., 2009). We then fit a quadratic model to the complete data set for the 165 retained students. The fit indices for the quadratic model, $T_{ML} = 2.97$ with $p = .56$, CFI = 1.00, TLI = 1.00, RMSEA = 0, and SRMR = 0.04, suggested adequate model fit of the quadratic model. The mean quadratic rate is also significant with Wald $z = 2.61$, $p = .009$. In comparison, we then fit a linear trajectory on the same data. For the linear model, $T_{ML} = 9.73$ with $p = .08$, CFI = .99, TLI = .99, RMSEA = .08, and SRMR = 0.09. Although SRMR suggested possible misfit based on Hu and Bentler's (1999) rule-of-thumb cutoff criterion, T_{ML}, RMSEA, CFI, and TLI all indicated acceptable fit of the linear model. Note that in both models we allowed the residual variance to vary across time. The CFI and TLI were corrected using an intercept-only model as the baseline model. However, when we saturated the covariance structure, $T_{ML} = 9.64$ with $p = .01$, CFI = .98, TLI = .97, RMSEA = .15, SRMR = .04. Notice that now both RMSEA and T_{ML} suggested misfit of the linear model. This example also reaffirms the inconsistency of SRMR with the other fit indices observed in the simulation study. In contrast to the other fit indices, SRMR suggested better fit of the linear model when the covariance structure was saturated.

Finally, our example also serves as an illustration of the danger of freeing the parameters in the error covariance matrix before settling the optimal mean structure. When we added an autoregressive coefficient to the linear model, all of the fit indices indicated acceptable fit: $T_{ML} = .02$ with $p = .99$, CFI = 1.00, TLI = 1.00, RMSEA = .00, and SRMR = 0.05. Such results could lead researchers to accept a linear trajectory for the mean structure, contrary to the hypothesized quadratic relationship. A better procedure would be first to use the saturated covariance structure to detect the misspecification in the mean structure and then to free the elements in the error covariance structure.

DISCUSSION

This study addressed the question of how sensitive SEM fit indices are to different sources of model misspecification in GCMs. The five fit indices were differentially sensitive to the different types of misspecification in the growth curve model even with severity of misspecification carefully controlled. Generally, the fit indices were most sensitive to the misspecification when the covariance between intercept and linear slope was fixed to zero or when the residual variances at Time1 and Time 5 were equated and least sensitive to the misspecification when the variance of quadratic acceleration was fixed to zero. SRMR was extremely sensitive to the misspecification when the covariance between intercept and linear rate was fixed at zero, but it was less sensitive to the other types of misspecifications. No fit index was always more (or less) sensitive to misspecification in the marginal mean structure relative to those in the covariance structure.

Generally, RMSEA, CFI, and TLI were more sensitive to the examined misspecifications in GCMs than T_{ML} and SRMR (except for tao01), potentially making them better fit indices. In addition, RMSEA, CFI, and TLI were not sensitive to N, whereas T_{ML} was highly affected by N and SRMR was affected by N under some conditions.

Of importance, saturating the covariance structure substantially improved the sensitivity of the practical fit indices to misspecification in the marginal mean structure. Saturating the marginal mean structure did not change the sensitivity of the SEM-based fit indices except for SRMR to misspecification in the covariance structure. These findings suggest that specifying one structure to be saturated improves the detection of misspecification in the marginal mean structure without decreasing detection of misspecification in covariance structures.

When there are misspecifications in both the marginal mean and covariance structures, whether a fit index is more sensitive to the misspecification in the marginal means or in the covariance structure depends on the type of misspecification. When a linear GCM was fit to the data generated from a quadratic model, all of the fit indices were strongly determined by the misspecification in marginal mean structure. The misspecifications in marginal mean structure and covariance structures interacted with each other to affect only the RMSEA and CFI.

Differentiating Different Sources of Misspecification in GCMs

An interesting question is whether one can differentiate different sources of misspecification in GCMs based on the available fit indices. If that is possible, then in addition to knowing that a model provides an inadequate fit to data, one can have a basis for identifying the source of the misfit and how to further improve the model. Following, we consider the possibility of identifying the source of misspecification(s) based on the findings in this study.

Misspecification in the marginal mean structure versus in the covariance structure. SEM-based fit indices reflect the misspecification(s) in the marginal mean and in both of the covariance structures. To differentiate the misspecification(s) in the marginal mean structure from those in the covariance structure, we proposed specifying one structure to be saturated to detect misspecification in the other structure. As described earlier, this approach worked especially well with the marginal mean structure. RMSEA, SRMR, CFI, and TLI were more sensitive to the misspecification in the marginal mean structure when the covariance structure was specified as saturated. Inspection of residuals in the marginal mean structure suggested that specifying the covariance structure as saturated increased the magnitude of the residuals in the marginal mean structure. Consequently it dramatically increased the part of the minimized discrepancy function (F_{ML}) that is related to the fit of the marginal mean structure.[4] The amount of increase exceeded the amount of the decrease in the part of the minimized discrepancy function that is related to the fit of the covariance structure.[5] Thus F_{ML} and T_{ML}, which equals $(N-1)F_{ML}$, were increased.

Specifying the marginal mean structure to be saturated did not substantially change the sensitivity of the SEM fit indices to misspecification in the covariance structure. Degrees of freedom (*df*) might play a role here. Usually, there would be more power to detect a specific misspecification with fewer *df* in a model. In our case, specifying the covariance structure to be saturated dramatically reduced the *df*. For the models with misspecification in only the marginal mean structure, saturating the covariance structure reduced the *df* from 10 to 3. However, saturating the marginal mean structure only reduced the *df* from 10 to 8. The relative benefits of specifying saturated models will depend on the relative reduction in the *df* for the tested models, which, in turn, will depend in large part on the number of measurement waves.

---

[4]See $(\overline{\mathbf{Y}} - \hat{\boldsymbol{\mu}})'\hat{\boldsymbol{\Sigma}}^{-1}(\overline{\mathbf{Y}} - \hat{\boldsymbol{\mu}})$ in Equation 4.

[5]See $\ln|\hat{\boldsymbol{\Sigma}}| - \ln|\mathbf{S}| + tr\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{S} - p$ in Equation 4.

Differentiating misspecification in the within-individual covariance structure versus in the between-individual covariance structure. Fit indices can be used to detect misspecification in both the within-individual and between-individual covariance structures. An approach is to follow the advice of Fitzmaurice, Laird, and Ware (2004) to free as many parameters in the between-individual covariance matrix as possible before specifying the parameters in the within-individual covariance matrix.[6] In this way, any lack of fit in the covariance structure must come from the misspecification in the within-individual matrix. Note that one cannot specify the within-individual matrix to be saturated to detect the misspecification in the between-individual matrix because the model will be underidentified (Kwok, West, & Green, 2007).

Limitations and Strengths of This Study and Directions for Future Research

As noted in the introduction, we followed the approach of acting as if the model were correct and investigating the effect of specific misspecifications on fit indices. In this case, this straightforward approach permitted us to examine carefully the separate and combined effects of exemplars of the full set of different general types of misspecification in the mean and covariance structures. It also permitted us to place our research findings in the context of the large body of previous research investigating the performance-of-fit indices (e.g., Hu & Bentler, 1998; Marsh et al., 1988; Yu, 2002). In contrast, following Box (1979) and Browne and Cudeck (1993), a second tradition assumes that models are at best useful approximations of an unknown true model. Within this tradition, no clear consensus has developed as to how the effects of more focused forms of misspecification could be studied. Cudeck and Browne (1992) proposed and Chun and Shapiro (2010) have extended a method of adding a perturbation matrix to a covariance matrix that satisfies a specified model and then using the covariance matrix with perturbation as the population for simulation study. However, we were unable to locate any simulation studies that have utilized this method in the context of evaluation of specific types of model misspecification. With the possible exception of the RMSEA, which invokes the convention that .05 is a close fitting model, it is not clear how the results of a simulation study could be interpreted because T_{ML} and practical fit indices

[6]There is a trade-off between freeing parameters in the between-individual and within-individual covariance matrices (Kwok et al., 2007). We propose to free as many parameters as possible in the between-individual covariance matrix fixing the remainder to zero. As an increasing number of parameters are freed in the between-individual covariance matrix, simpler error structures with fewer parameters will need to be entertained in the within-subjects covariance matrix or the model will not be identified.

involve comparisons with a perfectly fitting saturated model. The second and probably more realistic philosophical position is appealing but does not yet yield a clear path to answering the present research questions. A parallel observation can be made about the Satorra and Saris (1985) and MacCallum, Browne, and Sugawara (1996) methods of computing statistical power in SEM. The Satorra-Saris method estimates the power to detect a specific misspecification of fixed magnitude relative to a true model; the MacCallum et al. (1996) method estimates power given an unspecified global misspecification of fixed magnitude relative to a model with a specified error of approximation. Both approaches can be useful; they answer different questions (MacCallum, Browne, & Cai, 2006).

Our study examined the effects of only one misspecification in marginal mean structure, or one misspecification in the covariance structure, or both. Previous studies have failed to separate these two general sources of misspecification. In practice, more than two misspecifications can be present simultaneously in a model. Within the scope of this study, we could not examine a greater variety of forms of misspecification. Regardless of the type of misspecification in the covariance structure, our results suggested that saturating the covariance structure can help detect misspecifications in the mean structure.

We studied exemplars of five general types of misspecification in the study. These five general types represent the full set of misspecifications that could theoretically occur in growth curve models. These types of misspecification likely occur with different frequencies in practice with the failure to specify a covariance between an intercept and slope occurring only rarely in specialized models. In addition, within each of these general types, different forms of misspecification can occur. For example, in the error covariance structure, the relationship among residuals may be represented by structures other than the commonly assumed first-order autoregressive structure that we investigated, for example, moving average, Toeplitz, or autoregressive moving average structures. Failure to specify those relationships in the error covariance structure can lead to biased estimates of the variances and covariances of the growth factors, incorrect standard errors, and substantially reduced model fit (Kwok et al., 2007; Sivo et al., 2005). The current simulation studies considered only outcome variables with a multivariate normal distribution. In practice, the multivariate normality assumption is commonly violated and the current findings might not generalize fully to nonnormal data (Yuan, Bentler, & Zhang, 2005).

The current simulation studies considered only models in which time is the only variable that explains the change in an outcome variable. In many applications, there might be other time-varying covariates that also explain the change in the outcome variable, or time-invariant covariates that might explain individual differences in the growth trajectory, or both (Bollen & Curran, 2006; Grimm, 2007; Hedeker & Gibbons, 2006; Singer & Willett, 2003; Wu et al., 2008). If additional time-varying or time-invariant covariates exist, including or ignoring these covariates might make an appreciable difference in model fit. Including significant predictors of change in the outcome variable can typically improve the fit of the marginal mean structure. In addition, introducing additional covariates can change the magnitude of the total covariance matrix, which might have effects on the fit of the covariance structure (Kreft & de Leeuw, 1998; Singer & Willett, 2003).

We close by noting that this study has a number of key strengths compared with previous work: (a) It used the correct baseline model to calculate relative fit indices for GCMs, (b) it separately examined misspecification in the marginal mean and covariance structures, (c) it explored the possibility of constraining one structure to be saturated as a means of detecting the misspecification in the other structure, and (d) it carefully controlled for the severity of misspecification in the two structures using the TMFLR. Future research is needed that builds on these foundations and expands the range of models and distributions that are considered.

ACKNOWLEDGMENT

Supplementary materials for this article are available at http://people.ku.edu/∼wwei or by contacting Wei Wu. The collection of the data in the empirical example was supported by a grant from the National Institute of Child Health and Human Development to Jan N. Hughes (5 RO1 HD 39367).

REFERENCES

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

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

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. New York: Wiley.

Box, G. E. P. (1979). Robustness in the strategy of scientific model building. In R. L. Launer and G. N. Wilkinson (Eds.), *Robustness in statistics* (pp. 201–236). New York: Academic.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Chun, S. Y., & Shapiro, A. (2010). Construction of covariance matrices with a specified discrepancy function minimizer, with application to factor analysis. *SIAM Journal on Matrix Analysis and Applications, 31,* 1570–1583.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cudeck, R., & Browne, M. W. (1992). Constructing a covariance matrix that yields a specified minimizer and a specified minimum discrepancy value. *Psychometrika, 57,* 357–369.

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, IL.

Fan, X., & Sivo, S. A. (2005). Sensitivity of fit indices to misspecified structural or measurement model components: Rationale of two-index strategy revisited. *Structural Equation Modeling, 12,* 343–367.

Fan, X., & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research, 42,* 509–529.

Fitzmaurice, G., Laird, N., & Ware, J. (2004). *Applied longitudinal data analysis.* Hoboken, NJ: Wiley.

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. 40–65). Newbury Park, CA: Sage.

Grimm, K. J. (2007). Multivariate longitudinal methods for studying developmental relationships between depression and academic achievement. *International Journal of Behavioral Development, 34,* 328–339.

Hedeker, D., & Gibbons, R. D. (2006). *Longitudinal data analysis.* Hoboken, NJ: Wiley.

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

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

Jackson, D. L., Gillaspy, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods, 14,* 6–23.

Jöreskog, K. G., & Sörbom, D. (1981). *LISREL VI: Analysis of linear structural relationship by maximum likelihood and least squares methods.* Chicago: National Educational Resources.

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

Kwok, O., West, S. G., & Green, S. B. (2007). The impact of misspecifying the within-subject covariance structure in multiwave longitudinal multilevel models: A Monte Carlo study. *Multivariate Behavioral Research, 42,* 557–592.

Leite, W. L., & Stapleton, L. M. (2006). *Sensitivity of fit indices to detect misspecifications of growth shape in latent growth modeling.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

MacCallum, R. C., Browne, M. W., & Cai, L. (2006). Testing differences between nested covariance structure models: Power analysis and null hypotheses. *Psychological Methods, 11,* 19–35.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1,* 130–149.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103,* 391–410.

Marsh, H. W., & Hau, K.-T. (1996). Assessing goodness of fit: Is parsimony always desirable? *Journal of Experimental Education, 64,* 364–390.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55,* 107–122.

Muthén, B. O., & Curran, P. J. (1997). General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods, 2,* 371–402.

Muthén, L. K., & Muthén, B. O. (1998–2006). *Mplus user's guide.* Los Angeles: Author.

Olsson, U., Troye, S. V., & Howell, R. D. (1999). Theoretic fit and empirical fit: The performance of maximum likelihood versus generalized least squares estimation in structural equation models. *Multivariate Behavioral Research, 34,* 31–58.

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.

SAS Institute, Inc. (1999). *SAS procedures guide, version 8.* Cary, NC: Author.

Satorra, A., & Saris, W. E. (1985). Power of the likelihood ratio test in covariance structure analysis. *Psychometrika, 50,* 83–90.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

Sivo, S. A., Fan, X., & Witta, E. L. (2005). The biasing effects of unmodeled ARMA time series processes on latent growth curve model estimates. *Structural Equation Modeling, 12,* 215–231.

Steiger, J. H., & Lind, J. C. (1980, May). *Statistically based tests for the number of common factors.* Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.

Sun, J. (2005). Assessing goodness of fit in confirmatory factor analysis. *Measurement and Evaluation in Counseling and Development, 37,* 240–256.

Taylor, A. B. (2008). *Two new methods of studying the performance of SEM fit indexes*. Unpublished doctoral dissertation, Arizona State University, Tempe, AZ.

Tomarken, A. J., & Waller, N. G. (2003). Potential problems with well fitting models. *Journal of Abnormal Psychology, 112,* 578–98.

Tomarken, A. J., & Waller, N. G. (2005). Structural equation modeling: Strength, limitations, and misconceptions. *Annual Review of Clinical Psychology, 1,* 31–65.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38,* 1–10.

West, S. G., Taylor, A. B., & Wu, W. (in press). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling*. New York: Guilford.

Widaman, K. F., & Thompson, J. S. (2003). On specifying the null model for incremental fit indices in structural equation modeling. *Psychological Methods, 8,* 16–37.

Wu, W., West, S. G., & Hughes, J. N. (2008). Effect of retention in first grade on children's achievement trajectories over four years: A piecewise growth analysis using propensity score matching. *Journal of Educational Psychology, 100,* 727–740.

Wu, W., West, S. G., & Taylor, A. B. (2009). Evaluating model fit for growth curve models: Integration of fit indices from SEM and MLM frameworks. *Psychological Methods, 14,* 183–201.

Yu, C. Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes.* Unpublished doctoral dissertation, University of California, Los Angeles.

Yuan, K.-H. (2005). Fit indices versus test statistics. *Multivariate Behavioral Research, 40,* 115–148.

Yuan, K.-H., Bentler, P. M., & Zhang, W. (2005). The effect of skewness and kurtosis on mean and covariance structure analysis. *Sociological Methods & Research, 34,* 240–258.