Estimating Multilevel Linear Models
as Structural Equation Models

Daniel J. Bauer
North Carolina State University

Multilevel linear models (MLMs) provide a powerful framework for analyzing data collected at nested or non-nested levels, such as students within classrooms. The current article draws on recent analytical and software advances to demonstrate that a broad class of MLMs may be estimated as structural equation models (SEMs). Moreover, within the SEM approach it is possible to include measurement models for predictors or outcomes, and to estimate the mediational pathways among predictors explicitly, tasks which are currently difficult with the conventional approach to multilevel modeling. The equivalency of the SEM approach with conventional methods for estimating MLMs is illustrated using empirical examples, including an example involving both multiple indicator latent factors for the outcomes and a causal chain for the predictors. The limitations of this approach for estimating MLMs are discussed and alternative approaches are considered.

Keywords: *factor analysis, hierarchical linear models, mediation, multilevel models, structural equation models*

Multilevel linear models and structural equation models constitute two of the dominant analytic frameworks in contemporary social and behavioral science. Each approach offers distinct advantages. Multilevel linear models (MLMs) generalize simple regression models to circumstances in which the assumption of independent observations is compromised, as occurs when there is a "nested" or "clustered" data structure. A classic example is the observation of students within schools. Students within a school are likely to be more similar to one another than a simple random sample. Prior to the development of MLMs, this situation was handled either by estimating an ordinary least squares regression despite nonindependence, inflating Type I error rates, or by aggregating the data at the highest level (i.e., schools), resulting in a substantial loss of information and power (see Bryk & Raudenbush, 1992; Goldstein, 1995). MLMs offer the opportunity of analyzing both levels of the data simultaneously (i.e., student and school), thereby accounting for the nonindependence of the observations, and can even be used when clusters are overlapping (i.e., when exam-

ining individuals within both schools and neighborhoods). Further, MLMs capitalize on this data structure, allowing the effects of predictors to vary over clusters.

Despite these advantages, MLMs have several traditional limitations. First, it is difficult to incorporate a measurement model for the outcomes; that is, where the outcomes are multiple indicator latent factors with estimated factor loadings. Traditionally, a measurement model is added as a third level of the model (the item level), but in this approach the factor loadings of the items cannot be estimated and must be specified *a priori* instead (Goldstein, 1995; Raudenbush, Rowan, & Kang, 1991). A second traditional limitation of MLMs, shared with other regression models, is that complex causal processes, including mediational pathways, cannot be modeled directly. Current approaches for estimating mediational pathways involve either post-estimation matrix manipulations of estimates from a noncausal model (Raudenbush & Sampson, 1999) or the combination of information from multiple models (Krull & MacKinnon, 1999, 2001). Finally, inferential tests of the overall fit of MLMs are not routinely available (Raudenbush, 2001; but see du Toit & du Toit, in press).

The traditional limitations of the multilevel linear model are the primary strengths of structural equation modeling (SEM). SEMs are composed of two parts, an estimated measurement model relating the outcomes to the latent factors and a latent variable model that specifies the causal relations among the latent factors. One common structure for these relationships is a mediational pathway. In addition to obtaining coefficient estimates for the model, a key focus of SEM analyses is to assess the overall fit of the model, for which an asymptotic chi-square test (and numerous other fit indices) are available. However, use of SEMs has traditionally been limited to simple random samples.

Given the complementary strengths and limitations of MLMs and SEMs, a number of investigators have worked to synthesize the two models (e.g., Goldstein & MacDonald, 1988; Muthén, 1994; Muthén & Sattora, 1995; Rabe-Hesketh, Skrondal & Pickles, 2002). Most of these approaches require an expansion of the analytic model and highly specialized software. However, Rovine and Molenaar (1998, 2000, 2001) recently demonstrated that linear mixed-effects models (including MLMs) could be estimated as conventional SEMs. Bauer & Curran (2002) extended this approach, showing that even MLMs for unbalanced data could be estimated as SEMs. This article provides an expanded demonstration of this approach to multilevel modeling. The first aim is to illustrate, both analytically and with empirical examples, how MLMs for unbalanced data can be estimated as SEMs. Two approaches for handling unbalanced data are discussed and contrasted. The second aim of the article is to demonstrate how this approach may be used to integrate the strengths of the two modeling frameworks. Specifically, a new approach for estimating multilevel structural equation models will be illustrated through an empirical example that includes both a measurement model for the outcomes and a causal chain among the upper-level predictors.

The article begins by introducing the structural equation modeling framework. The multilevel linear model is then presented as a linear mixed-effects model. Initially, several important simplifying assumptions will be made in order to facilitate

translation of the multilevel model to an SEM. Specifically, no upper-level covariates will be included in the model, and the data will be assumed to be balanced. However, these assumptions will be progressively relaxed, first to accommodate unbalanced data, then to accommodate upper-level covariates. I will then demonstrate how MLMs can be extended within the SEM framework to include multiple indicator latent factors and complex causal structures. Finally, this approach to multilevel SEM will be contrasted with two alternative methods. Empirical examples will be used throughout to illustrate the major points.

Structural Equation Modeling

Structural equation models consist of a linear system of equations of two basic types. The first set of equations describes the relations between the latent variables, with a distinction made between *endogenous* latent variables $\boldsymbol{\eta}$, which are predicted by other variables in the model, and *exogenous* latent variables $\boldsymbol{\xi}$, which are external predictors whose own causes are unmodeled. Using Jöreskog and Sörbom's (1993) LISREL notation, the latent variable model is

$$\boldsymbol{\eta} = \boldsymbol{\alpha} + \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta} \tag{1}$$

where $\boldsymbol{\alpha}$ is a vector of intercept terms for the equations, $\mathbf{B}$ is the matrix of co-efficients giving the impact of the endogenous latent variables on each other, $\boldsymbol{\Gamma}$ is the coefficient matrix giving the effects of the exogenous latent variables on the endogenous latent variables, and $\boldsymbol{\zeta}$ is the vector of disturbances of $\boldsymbol{\eta}$ with covariance matrix $\boldsymbol{\Psi}$. The mean vector and covariance matrix of the exogenous latent variables are conventionally designated by $\boldsymbol{\kappa}$ and $\boldsymbol{\Phi}$, respectively. It is assumed that the expected value of $\boldsymbol{\zeta}$ is zero, that $\boldsymbol{\xi}$ and $\boldsymbol{\zeta}$ are uncorrelated, and that $(\mathbf{I} - \mathbf{B})$ is nonsingular.

The second set of equations defines the measurement model for the latent factors. It is here that the observed variables of the model are related to the latent variables they are thought to represent. The set of observed variables representing the exogenous latent variables is designated as $\mathbf{x}$ and is distinguished from the set of observed variables representing the endogenous latent variables, which are designated as $\mathbf{y}$. The measurement model then consists of two equations of the same basic form, one regressing $\mathbf{x}$ on $\boldsymbol{\xi}$ and another regressing $\mathbf{y}$ on $\boldsymbol{\eta}$:

$$\mathbf{x} = \boldsymbol{\upsilon}_x + \boldsymbol{\Lambda}_x\boldsymbol{\xi} + \boldsymbol{\delta} \tag{2}$$

$$\mathbf{y} = \boldsymbol{\upsilon}_y + \boldsymbol{\Lambda}_y\boldsymbol{\eta} + \boldsymbol{\epsilon}, \tag{3}$$

where $\boldsymbol{\Lambda}_x$ is the factor loading matrix relating $\mathbf{x}$ to $\boldsymbol{\xi}$ and $\boldsymbol{\Lambda}_y$ is the factor loading matrix relating $\mathbf{y}$ to $\boldsymbol{\eta}$. The intercepts of $\mathbf{x}$ and $\mathbf{y}$ are contained in $\boldsymbol{\upsilon}_x$ and $\boldsymbol{\upsilon}_y$, respectively. Finally, the vectors $\boldsymbol{\delta}$ and $\boldsymbol{\epsilon}$ represent the residuals of $\mathbf{x}$ and $\mathbf{y}$, with covariance matrices $\boldsymbol{\Theta}_\delta$ and $\boldsymbol{\Theta}_\epsilon$, respectively. It is assumed that the expected values of $\boldsymbol{\delta}$ and $\boldsymbol{\epsilon}$ are zero and that $\boldsymbol{\delta}, \boldsymbol{\epsilon}, \boldsymbol{\xi}$ and $\boldsymbol{\zeta}$ are uncorrelated with one another.

Several additional features of the SEM are noteworthy. First, the general model reduces to many other familiar modeling approaches. For example, if only observed variables are present in the model, then **x** and **y** can be substituted for **ξ** and **η** in Equation 1 and the model reduces to a simultaneous equation model (or traditional path analysis). Alternatively, if there is a measurement model but no specific causal structure for the latent variables (only correlations) then the model reduces to confirmatory factor analysis (CFA). Second, SEM is often referred to as covariance structure analysis because the model equations imply a specific form for the population covariance matrix **Σ** and mean vector **μ.** The CFA model will be important for subsequent developments, so it provides a useful example. If the CFA is modeled on the "y-side" then the model-implied covariance matrix and mean vector are

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \boldsymbol{\Lambda}_y \boldsymbol{\Psi} \boldsymbol{\Lambda}_y' + \boldsymbol{\Theta}_\epsilon$$

$$\boldsymbol{\mu}(\boldsymbol{\theta}) = \boldsymbol{\nu}_y + \boldsymbol{\Lambda}_y \boldsymbol{\alpha},$$

(4)

where **θ** is the vector of model parameters from $\boldsymbol{\Lambda}_y$, **Ψ, Θ$_\epsilon$,** $\boldsymbol{\nu}_y$, and **α.**

A key objective of structural equation modeling is to assess the fit of the hypothesized moment structure. The null hypothesis is that

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}$$

$$\boldsymbol{\mu}(\boldsymbol{\theta}) = \boldsymbol{\mu},$$

(5)

that is, that the model perfectly reproduces the population variances, covariances and means of the observed variables. This hypothesis is tested by of a likelihood-ratio χ^2 test, involving the estimation and comparison of two nested models. The first is the structural equation model of interest. The second is the *saturated model* where no restrictions are placed on the estimated covariance matrix or mean vector. The saturated model provides a benchmark of perfect fit for assessing the degree of misfit of the structural equation model. Formally, using LL_0 to represent the log-likelihood of the structural model and LL_1 to represent the log-likelihood of the saturated model, $-2(LL_0 - LL_1)$ is asymptotically distributed as a chi-square with degrees of freedom equal to the difference in the number of parameters for the two models. Small and nonsignificant chi-squares suggest that the structural equation model fits well (i.e., cannot be rejected).[1]

I now turn to the multilevel linear model and how it may be translated into an SEM.

A Simple Multilevel Linear Model

Consider the special case of a two-level linear model that includes only Level 1 covariates. The Level 1 model may be written as

$$y_{ij} = \pi_{0j} + \sum_{p=1}^{P} \pi_{pj} x_{pij} + r_{ij}.$$

(6)

For ease of presentation assume that i indexes individual and j indexes group so that y_{ij} may be interpreted as the response of individual i nested within group j. This response is modeled as a function of a random intercept term (π_{0j}), and P Level 1 covariates with random slopes ($\pi_{1j}, \ldots, \pi_{Pj}$). The residuals of y_{ij}, denoted r_{ij}, are assumed to be multivariate normally distributed as

$$\mathbf{r}_j \ \sim \ MVN\,(\mathbf{0}, \mathbf{\Sigma}_{r_j}). \tag{7}$$

Often $\mathbf{\Sigma}_{r_j}$ is assumed to be diagonal (reflecting conditional independence), and the residual variances are frequently constrained to be equal. These constraints on $\mathbf{\Sigma}_{r_j}$ are sometimes necessary to estimate the model while at other times they are not.

Given the absence of Level 2 covariates, the Level 2 model is simply

$$\begin{aligned}
\pi_{0j} &= \beta_0 + u_{0j} \\
\pi_{pj} &= \beta_p + u_{pj},
\end{aligned} \tag{8}$$

where β_0 represents the average intercept and β_p represents the average effect of x_p, while group-level deviations from these averages are represented by the disturbances u_{0j} and u_{pj}. These disturbances are conventionally assumed to be multivariate normally distributed as

$$\mathbf{u}_j \ \sim \ MVN\,(\mathbf{0}, \mathbf{T}), \tag{9}$$

where $\mathbf{T}$ is typically an unrestricted covariance matrix.

Equation 8 may then be substituted into Equation 7 to obtain the reduced-form equation for the model

$$y_{ij} = (\beta_0 + u_{0j}) + \sum_{p=1}^{P} (\beta_p + u_{pj})\, x_{pij} + r_{ij}. \tag{10}$$

By expanding and rearranging the terms, this equation may be rewritten as

$$y_{ij} = \left[\beta_0 + \sum_{p=1}^{P} \beta_p x_{pij} \right] + \left[u_{0j} + \sum_{p=1}^{P} u_{pj} x_{pij} \right] + r_{ij}. \tag{11}$$

Rewriting the model in this way segregates the fixed effects of the model (the first bracketed term) from the random effects (the second bracketed term). Note that while the fixed effects are constant over schools, the random effects are subscripted by j to indicate that they vary over schools. This illustrates that the model is a special case of the general linear mixed-effects model given by Laird & Ware (1982) as

$$\mathbf{y}_j = \mathbf{X}_j \mathbf{\beta} + \mathbf{Z}_j \mathbf{u}_j + \mathbf{r}_j, \tag{12}$$

where $\mathbf{y}_j$ is the response vector for group j, $\mathbf{X}_j$ is the design matrix for the fixed effects $\boldsymbol{\beta}$, $\mathbf{Z}_j$ is the design matrix of the random effects $\mathbf{u}_j$, and $\mathbf{r}_j$ is a vector of residuals. Assuming the random effects and residuals are multivariate normally distributed, the implied marginal distribution for $\mathbf{y}_j$ is then

$$\mathbf{y}_j \sim MVN\,(\mathbf{X}_j\boldsymbol{\beta},\, \mathbf{Z}_j\mathbf{T}\mathbf{Z}_j' + \boldsymbol{\Sigma}_{r_j}). \tag{13}$$

As Verbeke & Molenberghs (2000) noted, it is often this marginal model for $\mathbf{y}_j$ that is actually fit to the data.

A second reflection on Equation 11 is that the fixed and random parts of the model are strictly parallel. This parallelism follows from the absence of Level 2 covariates. Importantly, because of this simplification, the design matrices for the fixed and random effects are equivalent and $\mathbf{X}_j$ may be substituted for $\mathbf{Z}_j$ to yield

$$\mathbf{y}_j \sim MVN\,(\mathbf{X}_j\boldsymbol{\beta}, \mathbf{X}_j\mathbf{T}\mathbf{X}_j' + \boldsymbol{\Sigma}_{r_j}). \tag{14}$$

If the design is balanced (i.e., $\mathbf{X}_j = \mathbf{X}$ for all j) then the marginal model may be further simplified to

$$\mathbf{y}_j \sim MVN\,(\mathbf{X}\boldsymbol{\beta}, \mathbf{X}\mathbf{T}\mathbf{X}' + \boldsymbol{\Sigma}_r). \tag{15}$$

What is especially interesting about Equation 15 is that it has the same basic structure as the CFA model. Specifically, the LISREL matrices in Equation 4 can be mapped onto the multilevel notation as follows

$$\boldsymbol{\Lambda}_y = \mathbf{X}$$
$$\boldsymbol{\alpha} = \boldsymbol{\beta}$$
$$\boldsymbol{\upsilon}_y = \mathbf{0}$$
$$\boldsymbol{\Psi} = \mathbf{T}$$
$$\boldsymbol{\Theta}_\varepsilon = \boldsymbol{\Sigma}_r$$

The key aspect of this translation is that the design matrix of the Level 1 covariates $\mathbf{X}$ is represented by the factor loading matrix $\boldsymbol{\Lambda}_y$. That is, the loadings for the latent factors are fixed to the values of the covariates in $\mathbf{X}$.[2] In turn, the random regression coefficients $\boldsymbol{\pi}$ of the MLM are represented as latent factors $\boldsymbol{\eta}$ in the SEM, with factor means equal to the fixed effects, and with factor variances and covariances equal to the variances and covariances of the random effects.

An apparent limitation of this translation is that $\mathbf{X}$ can only be represented as $\boldsymbol{\Lambda}_y$ by assuming a balanced design. Use of this approach for estimating MLMs as SEMs has thus been restricted to certain special cases, such as growth curve models

(Chou, Bentler, & Pentz 1998; Meredith & Tisak, 1984, 1990; MacCallum et al., 1997; Willett & Sayer, 1994) and models for dyadic data (Newsom, 2002), where balanced designs are common. As a starting point, this approach to estimating MLMs with balanced data as SEMs is illustrated below. Subsequently, I discuss how unbalanced designs can be accommodated within the SEM.

Fitting MLMs with Balanced Data as SEMs

Consider a study in which the language proficiency of 3 male and 3 female students was assessed in each of N schools. The primary interest is whether there is a sex-difference in language ability, and whether this sex difference varies significantly over schools. Sex has been dummy coded so that male = 0 and female = 1. The MLM then has two levels, a student level and a school level. The Level 1 (student-level) model is written as

$$lang_{ij} = \pi_{0j} + \pi_{1j}sex_{ij} + r_{ij}, \tag{16}$$

with constant residual variance σ. The Level 2 (school-level) model is:

$$\pi_{0j} = \beta_0 + u_{0j}$$
$$\pi_{1j} = \beta_1 + u_{1j}, \tag{17}$$

where

$$\mathbf{T} = \begin{pmatrix} \tau_{00} & \\ \tau_{10} & \tau_{11} \end{pmatrix}. \tag{18}$$

The same model can be parameterized as an SEM. Assume that the data vector for each school is ordered so that the three male students' scores always precede the three female students' scores. Then the measurement model may be written as

$$\begin{pmatrix} lang_{m1j} \\ lang_{m2j} \\ lang_{m3j} \\ lang_{f1j} \\ lang_{f2j} \\ lang_{f3j} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \pi_{0j} \\ \pi_{1j} \end{pmatrix} + \begin{pmatrix} r_{m1j} \\ r_{m2j} \\ r_{m3j} \\ r_{f1j} \\ r_{f2j} \\ r_{f3j} \end{pmatrix}, \tag{19}$$

where for convenience the subscripts on the variables (and their disturbances) designate the sex of the student (*m* for male and *f* for female) as well as the order of the observations within sex (1, 2, or 3). The design matrix for the fixed and random

141

effects $\mathbf{X}$ is represented by Λ_y, which is composed of a column vector of ones for the intercept parameter and a column vector containing the dummy coded values of the sex covariate for the slope parameter.

The latent variable model for this MLM can then be written

$$\begin{pmatrix} \pi_{0j} \\ \pi_{1j} \end{pmatrix} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix}, \tag{20}$$

where β_0 and β_1 are captured as latent variable means ($\boldsymbol{\alpha}$), and the random coefficients u_{0j} and u_{1j} are represented as disturbances from those means ($\boldsymbol{\zeta}$). The covariance matrix of the latent factors is equivalent to the covariance matrix of the random effects from the multilevel model, or

$$\boldsymbol{\Psi} = \mathbf{T} = \begin{pmatrix} \tau_{00} \\ \tau_{10} & \tau_{11} \end{pmatrix}. \tag{21}$$

A path diagram showing the SEM representation of this model is presented in Figure 1.

It is important to recognize that the order of the observations within sex is arbitrary. That is, although the scores of students $m1$, $m2$, and $m3$ are represented as dif-
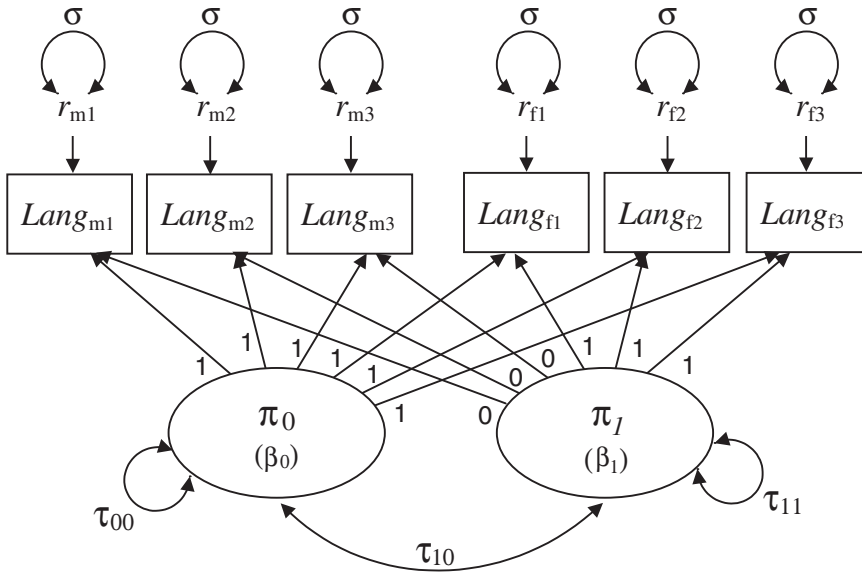

FIGURE 1. *A path diagram of a multilevel linear structural equation model for balanced data with a categorical Level 2 covariate.*

ferent observed variables in the SEM, there is nothing in the model that distinguishes one male student from another within a given school; they are viewed as interchangeable. The arbitrary ordering of the observations must be accounted for by employing equality constraints on the elements of the residual covariance matrix (this is true of the multilevel linear model as well). In this case, constant variance is assumed so that

$$\boldsymbol{\Theta}_{\epsilon} = \boldsymbol{\Sigma}_{r} = \sigma\mathbf{I}. \tag{22}$$

Such constraints ensure that the model-implied variances, covariances, and means are identical over students of a given sex, so that the ordering of the observations has no impact on the parameter estimates, standard errors or fit of the model. This can be seen in the model-implied covariance matrix and mean vector for this example, given in Table 1.

The arbitrary ordering of the observations also requires some modification of the usual χ^2 test of overall model fit typically used with SEMs. Specifically, the appropriate saturated model for judging the fit of the hypothesized model must also be estimated with equality constraints, again following the notion that students of a given sex are essentially interchangeable. For this example, the appropriate saturated model is given in the lower half of Table 1, where elements designated by the same letter are constrained to be equal. The letter *a* is the variance in language scores among males, and *b* is the covariance among male students within schools (accounting for the nonindependence of the observations). Similar interpretations may be given to *d* and *e* for female students. The letter *c* then captures the covariance between male and female students within schools and *f* and *g* are the means of the male and female students, respectively. I will refer to this model as the *constrained saturated model.* Comparing the model-implied moment matrices between the hypothesized model and the constrained saturated model shows that the two models are formally nested.

To demonstrate the equivalency of the two approaches to fitting the model, I reanalyze data previously presented by Snijders and Bosker (1999, p. 46) on the language proficiency of 8th-grade (about 11 year-old) students. Between 4 and 35 students were sampled from each of 131 elementary schools in the Netherlands. To meet the constraints of a balanced design, only schools with language scores on at least three male and three female students were included in the analysis ($N = 116$ schools); when more male or female students were available, three scores were randomly selected (bringing the total number of students to 696). Language scores ranged from 8 to 58, and showed little skew or kurtosis. This model, and those that will be presented subsequently, was fit using PROC MIXED in SAS, a standard software package for conducting multilevel analyses, and M*plus,* a commonly used SEM software package, in both cases using the full-information maximum likelihood fitting function.[3]

Figure 2, drawn after the recommendations of Gelman, Pasarica, & Dodhia (2002), compares the parameter estimates and standard errors obtained from the two modeling approaches for each of three estimated models. The two left-most entries in each panel correspond to the estimates for the current model with balanced data (the

TABLE 1
The Covariance Matrix and Mean Vector Implied by the Multilevel Linear Model and the Constrained Saturated Model for the Balanced Design Example

	Covariance Matrix and Mean Vector Implied by the Multilevel Linear Model					
Variance/covariance	$lang_{m1}$	$lang_{m2}$	$lang_{m3}$	$lang_{f1}$	$lang_{f2}$	$lang_{f3}$
$lang_{m1}$	$\hat{\tau}_{00} + \hat{\sigma}$					
$lang_{m2}$	$\hat{\tau}_{00}$	$\hat{\tau}_{00} + \hat{\sigma}$				
$lang_{m3}$	$\hat{\tau}_{00}$	$\hat{\tau}_{00}$	$\hat{\tau}_{00} + \hat{\sigma}$			
$lang_{f1}$	$\hat{\tau}_{00} + \hat{\tau}_{10}$	$\hat{\tau}_{00} + \hat{\tau}_{10}$	$\hat{\tau}_{00} + \hat{\tau}_{10}$	$\hat{\tau}_{00} + 2\hat{\tau}_{10} + \hat{\tau}_{11} + \hat{\sigma}$		
$lang_{f2}$	$\hat{\tau}_{00} + \hat{\tau}_{10}$	$\hat{\tau}_{00} + \hat{\tau}_{10}$	$\hat{\tau}_{00} + \hat{\tau}_{10}$	$\hat{\tau}_{00} + 2\hat{\tau}_{10} + \hat{\tau}_{11}$	$\hat{\tau}_{00} + 2\hat{\tau}_{10} + \hat{\tau}_{11} + \hat{\sigma}$	
$lang_{f3}$	$\hat{\tau}_{00} + \hat{\tau}_{10}$	$\hat{\tau}_{00} + \hat{\tau}_{10}$	$\hat{\tau}_{00} + \hat{\tau}_{10}$	$\hat{\tau}_{00} + 2\hat{\tau}_{10} + \hat{\tau}_{11}$	$\hat{\tau}_{00} + 2\hat{\tau}_{10} + \hat{\tau}_{11}$	$\hat{\tau}_{00} + 2\hat{\tau}_{10} + \hat{\tau}_{11} + \hat{\sigma}$
M	$\hat{\beta}_0$	$\hat{\beta}_0$	$\hat{\beta}_0$	$\hat{\beta}_0 + \hat{\beta}_1$	$\hat{\beta}_0 + \hat{\beta}_1$	$\hat{\beta}_0 + \hat{\beta}_1$

	Covariance Matrix and Mean Vector for the Constrained Saturated Model					
Variance/covariance	$lang_{m1}$	$lang_{m2}$	$lang_{m3}$	$lang_{f1}$	$lang_{f2}$	$lang_{f3}$
$lang_{m1}$	$\hat{a}$					
$lang_{m2}$	$\hat{b}$	$\hat{a}$				
$lang_{m3}$	$\hat{b}$	$\hat{b}$	$\hat{a}$			
$lang_{f1}$	$\hat{c}$	$\hat{c}$	$\hat{c}$	$\hat{d}$		
$lang_{f2}$	$\hat{c}$	$\hat{c}$	$\hat{c}$	$\hat{e}$	$\hat{d}$	
$lang_{f3}$	$\hat{c}$	$\hat{c}$	$\hat{c}$	$\hat{e}$	$\hat{e}$	$\hat{d}$
M	$\hat{f}$	$\hat{f}$	$\hat{f}$	$\hat{g}$	$\hat{g}$	$\hat{g}$

Note. Elements designated by the same symbols or letters are constrained to be equal. Symbols defined in text.
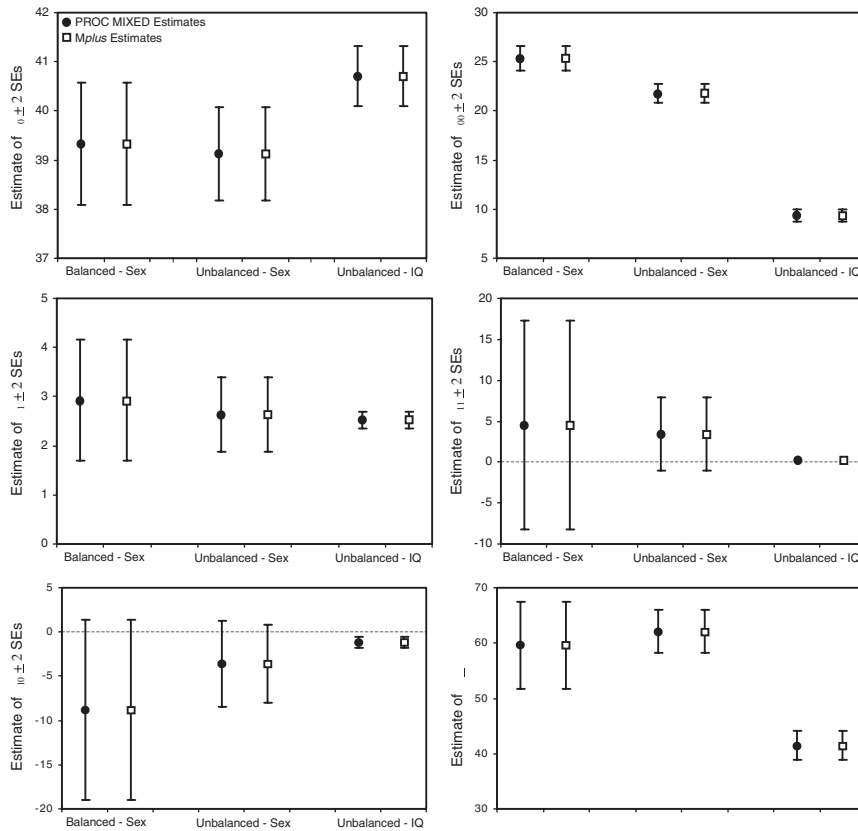
FIGURE 2. *A comparison of the parameter estimates (± 2 standard errors) obtained by fitting three different multilevel linear models to the language data of Snijders & Bosker (1999) using the SEM software Mplus and the multilevel software SAS PROC MIXED. The first model includes the categorical covariate sex and was restricted to balanced data by selecting three male and three female students per school, the second model includes the categorical covariate sex and was fit to the full sample of students using the missing data approach to unbalanced data, and the third model includes the continuous covariate verbal IQ and was fit to the full sample of students using school-varying factor loadings to accommodate the unbalanced data.*

remaining estimates will be discussed shortly). As can be seen, the parameter estimates and standard errors are virtually identical over the two modeling approaches.[4] In both cases, there is a small but significant sex difference in language scores favoring female students by an average of about 3 points, and no significant school-level variability in this effect. The log-likelihoods obtained by the two approaches are also equivalent within rounding error. To test the overall fit of the model to the data, this value was compared to the log-likelihood of the constrained saturated model, yielding

$\chi^2(1) = 1.18$; $p = .72$. Thus we may conclude that the hypothesized model fits the observed data quite well. Note that such a test of model fit is routine in SEM analyses, but atypical using traditional approaches to multilevel modeling.

To summarize to this point, when the design is balanced an MLM can be represented as an SEM quite easily. The example provided above simply generalizes a relationship between the two models that is well-understood in certain contexts, such as growth curve modeling and the modeling of dyadic data, where balanced designs are common. The primary complexity involved in estimating the MLM as an SEM when the design is balanced is that the chi-square test of model fit must be modified if the ordering of the observations is arbitrary. When the design is unbalanced, additional complexities arise, but given recent analytical and software developments they are no longer an impediment to parameterizing MLMs as SEMs, as I now illustrate.

Fitting MLMs with Unbalanced Data as SEMs

Recall that the primary reason that a balanced design was required to make the translation from the MLM to the SEM was so that $\mathbf{X}_j$ could be simplified to $\mathbf{X}$ in Equation 14. Given this, $\mathbf{X}$ could easily be incorporated in the SEM in the factor loading matrix $\mathbf{\Lambda}_y$. In the case that the data are unbalanced, however, this simplification cannot be made. How then, can the model be fit as an SEM? Mehta & West (2000) considered this issue for growth models with unbalanced data. They reviewed two approaches for handling unbalanced data in the SEM framework, both of which are generalizable to a wide variety of MLMs. The first approach is to act as though the design is balanced, but that certain observations are missing. A full-information case-based likelihood is then used to estimate the model using only the available observations. This approach is available in most SEM software but can be cumbersome to implement. A second and more natural approach is to accommodate $\mathbf{X}_j$ directly by allowing the factor loading matrix $\mathbf{\Lambda}_y$ to vary over schools. Again, a case-based likelihood function is evaluated, where the model-implied means and covariances are updated for each case given their specific $\mathbf{\Lambda}_y$.[5] Both approaches have distinct advantages and disadvantages, so I demonstrate each in turn using empirical examples.

Treating Unbalanced Data as Missing

In the preceding analysis, 3 male and 3 female students were sampled from each school so that the design would be balanced. However, in actuality, between 2 and 18 male students and 0 and 19 female students were available for analysis over the 131 schools. This situation is easily handled in the multilevel modeling framework where the marginal model for the outcomes naturally varies over schools per Equation 14. The multilevel representation of the model is thus unchanged from Equations 16 and 17. By contrast, the representation of the model as an SEM must be changed to accommodate the unbalanced data.

To fit the model to the unbalanced data, I first act as though the design is balanced and that 18 male and 19 female students were sampled from each school (the maximum number from each sex). The total number of possible observations from each

school is then 37. Second, if a school has fewer than 18 males or 19 females, I code the remaining observations as missing. Again placing the male language scores before the female language scores, the measurement model can be modified to

$$
\begin{pmatrix}
lang_{m1j} \\
lang_{m2j} \\
\vdots \\
lang_{m18j} \\
lang_{f1j} \\
lang_{f2j} \\
\vdots \\
lang_{f19j}
\end{pmatrix}
=
\begin{pmatrix}
1 & 0 \\
1 & 0 \\
\vdots & \vdots \\
1 & 0 \\
1 & 1 \\
1 & 1 \\
\vdots & \vdots \\
1 & 1
\end{pmatrix}
\begin{pmatrix}
\pi_{0j} \\
\pi_{1j}
\end{pmatrix}
+
\begin{pmatrix}
r_{m1j} \\
r_{m2j} \\
\vdots \\
r_{m18j} \\
r_{f1j} \\
r_{f2j} \\
\vdots \\
r_{f19j}
\end{pmatrix},
\tag{23}
$$

and the latent variable model remains as before. I continue to assume constant variance for the residuals and a full and freely estimated covariance matrix for the random effects (latent factors). A path diagram of the model is provided in Figure 3.

Note that by setting up the model this way, a single Λ_y common to all schools is sufficient. The model is then fit by constructing a single $\hat{\Sigma}(\theta)$, with 37 rows and columns, and a single $\hat{\mu}(\theta)$, with 37 rows. As the likelihood function is evaluated for each case *j,* rows and columns are dropped from these matrices to leave only the
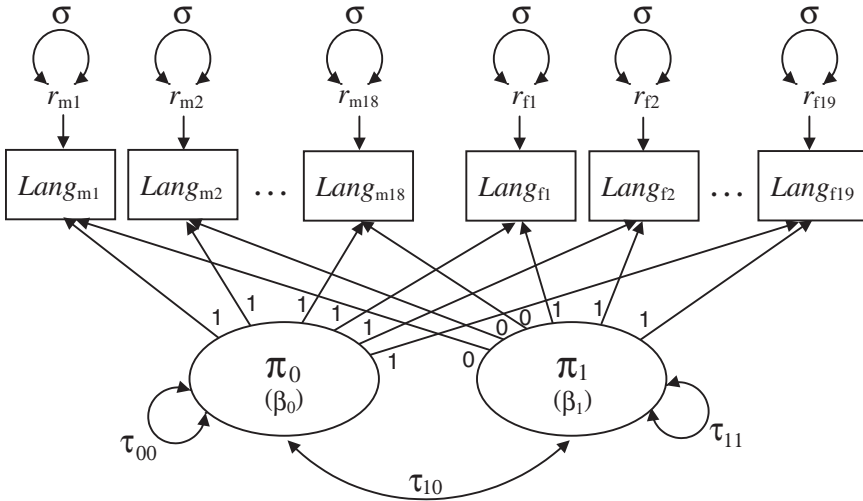

FIGURE 3. *A path diagram of a multilevel linear structural equation model with a categorical Level 2 covariate where unbalanced data are treated as missing.*

elements of the covariance matrix and mean vector that correspond to the available observations for the case (Arbuckle, 1996; Wothke, 2000). Thus, although $\mathbf{\Lambda}_y$ is constant, the implied marginal distribution for the observations [summarized by $\hat{\mathbf{\Sigma}}(\mathbf{\theta})$ and $\hat{\mathbf{\mu}}(\mathbf{\theta})$] is nevertheless updated case-wise to accommodate the unbalanced design. The SEM approach and multilevel approach are hence analytically equivalent.

To demonstrate this point, the model was once again fit to the data of Snijders and Bosker (1999), but this time including the language scores of all the students (a total of 2,287 students). The results of fitting the model as an MLM and as an SEM are presented in the center of each panel of Figure 2. Note that the parameter estimates, standard errors, and log-likelihoods are again essentially equivalent (within rounding error). In comparison to the prior model, the standard errors for this model are much smaller, reflecting the larger sample of students included in the analysis. To test the fit of the model, the constrained saturated model was also fit to the data, and the resulting likelihood-ratio test again indicated good fit for the hypothesized model [$\chi^2(1) = 2.41; p = .12$].

There are two primary advantages to this strategy for handling unbalanced data. First, a test of overall model fit is possible because a single $\hat{\mathbf{\Sigma}}(\mathbf{\theta})$ and $\hat{\mathbf{\mu}}(\mathbf{\theta})$ can be used to summarize the data. Second, it is easy to allow the residual variances of the observed variables to differ over values of the Level 1 covariate(s) (e.g., constraining the residual variance to different values for male and female students). The primary drawback of this strategy is that there are many models for which a single $\hat{\mathbf{\Sigma}}(\mathbf{\theta})$ and $\hat{\mathbf{\mu}}(\mathbf{\theta})$ cannot be constructed (at least not without substantial difficulty). For instance, suppose that the covariate of interest was verbal intelligence. Unlike sex, verbal intelligence is a continuous covariate that can, in principle, take on an infinite number of different values. In this case, it would be difficult to treat the design as balanced, and the number of "missing" values would far exceed the number of observed values, increasing the computational difficulty of estimating the model. What is needed in such cases is the ability to allow the factor loading matrix to vary over the Level 2 units. This constitutes the second strategy for accommodating unbalanced data in SEM, which I now demonstrate.

Treating Unbalanced Data with Case-Varying Factor Loadings

Allowing the factor loading matrix to take on different fixed values for each case *j* directly parallels the multilevel approach to unbalanced data (Neale et al., 1999). That is, $\mathbf{X}_j$ is represented as $\mathbf{\Lambda}_{yj}$. To demonstrate this approach, I extend the language proficiency example, this time using verbal intelligence (IQ) as the Level 1 covariate. The MLM consists of the Level 1 (student-level) model

$$lang_{ij} = \pi_{0j} + \pi_{1j}IQ_{ij} + r_{ij}, \tag{24}$$

where $\mathbf{\Sigma}_r = \sigma\mathbf{I}$. The Level 2 (school-level) model is identical in form to Equations 17 and 18, only this time the coefficients reflect the regression of language ability on IQ (not sex).

This model can be represented equivalently as an SEM with the following measurement model

$$
\begin{pmatrix} lang_{1j} \\ lang_{2j} \\ \vdots \\ lang_{Ij} \end{pmatrix} = \begin{pmatrix} 1 & IQ_{1j} \\ 1 & IQ_{2j} \\ \vdots & \vdots \\ 1 & IQ_{Ij} \end{pmatrix} \begin{pmatrix} \pi_{0j} \\ \pi_{1j} \end{pmatrix} + \begin{pmatrix} r_{1j} \\ r_{2j} \\ \vdots \\ r_{Ij} \end{pmatrix}, \tag{25}
$$

where I represents the total number of students in school $j,$ and the values IQ_{1j}, $IQ_{2j}, \ldots IQ_{Ij}$ in $\mathbf{\Lambda}_y$ are updated as the likelihood function is evaluated for each school j. Note that because $\mathbf{\Lambda}_y$ is permitted to vary over j, there is no particular need to sort the observations, unlike the missing data approach. Consistent with the multi-level model, the covariance matrix of the residuals is constrained to $\mathbf{\Theta}_\epsilon = \mathbf{\Sigma}_r = \sigma\mathbf{I}$. The latent variable model is of the same form as Equations 20 and 21. A path diagram for this model is presented in Figure 4.

This model was again fit to the data of Snijders & Bosker (1999), including 2,287 students from 131 schools. To make the results of the analysis more interpretable, the verbal IQ scores (obtained from the ISI test) were standardized prior to the analysis.
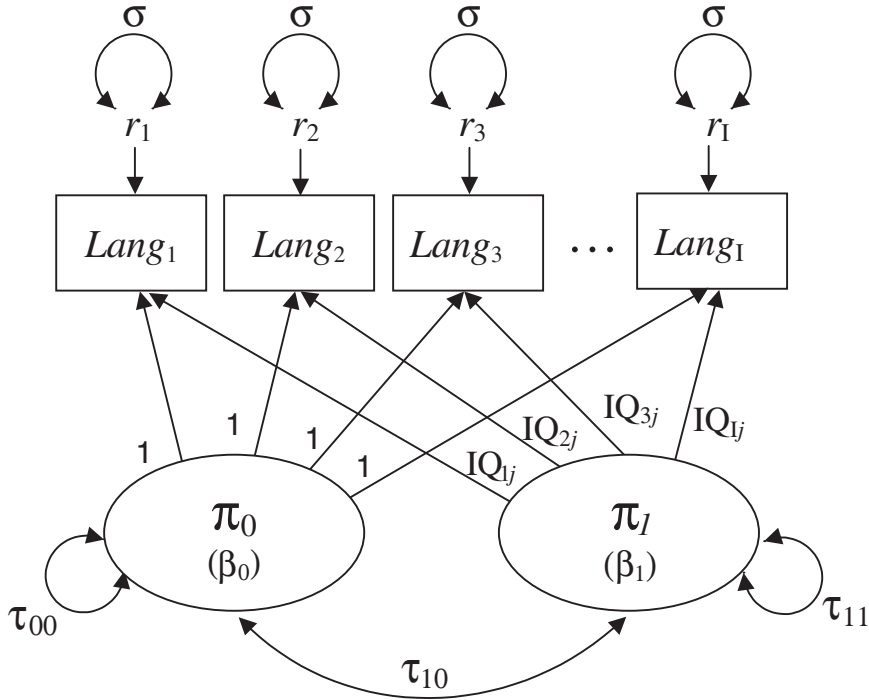

FIGURE 4.   *A path diagram of a multilevel linear structural equation model with a continuous Level 2 covariate accommodated by allowing the factor loading matrix to vary over Level 2 units.*

The estimates and standard errors obtained by fitting the model as an MLM and as an SEM are displayed in Figure 2 as the right-most entries of each panel. The parameter estimates are again essentially identical (as are the log-likelihoods of the models), and the standard errors differ only slightly. Both models indicate that, on average, each standard deviation increase in a student's IQ is associated with a 2.5 unit increase in language scores. Though not readily apparent from Figure 2, significant school-level differences were also detected in the size of this effect. Finally, a basic comparison of the residual variance estimates from each of the fitted models suggests that IQ explains much more of the individual variability in language scores than sex. Note, however, that it is not possible to formally test the fit of the IQ model using the conventional likelihood-ratio statistic because this model cannot be summarized by a single covariance matrix and mean vector (Raudenbush, 2001). This constitutes the primary drawback of using case-varying factor loadings to accommodate unbalanced data rather than the missing data approach when the latter is feasible.

The issue of how to treat unbalanced data when fitting multilevel models in SEM is more than anything a data management problem. As I have shown, recent developments have made it possible to fit MLMs with categorical and continuous Level 1 covariates as SEMs, even when the data is unbalanced. What I have not yet addressed is the ability to incorporate upper-level predictors of the random coefficients in the model. It is to this topic that I now turn.

Adding Upper-Level Predictors to the Model

Upper-level predictors were initially excluded from the model to simplify the problem of translating the MLM to an SEM. Specifically, the absence of upper-level predictors implied that the design matrix for the fixed and random effects would be identical ($\mathbf{X}_j = \mathbf{Z}_j$) in Equation 12. This design matrix was then easily translated into the factor loading matrix of an SEM, allowing the fixed and random effects to be captured as factor means and disturbances. The problem with adding Level 2 covariates to the model is that this simplification can no longer be made. The fixed effects will out-number the random effects, and the design matrix $\mathbf{X}_j$ will thus require more columns than the design matrix $\mathbf{Z}_j$. There are two analytically equivalent solutions to this problem. The first was formulated by Rovine & Molenaar (1998, 2000, 2001), and follows the linear mixed-effects model most literally. The second extends the approach used with conditional latent curve models where upper-level predictors are incorporated directly as observed variables. These alternatives are contrasted in the next section.

The Latent Variable Solution

Rovine & Molenaar (1998, 2000, 2001) parameterized the linear mixed-effects model in Equation 12 as an SEM by defining two types of latent variables, one to represent the fixed effects and another to represent the random effects. Using the "y-side" of the SEM for consistency, I designate the "fixed-effects factors" as $\boldsymbol{\eta}_1$ and the "random-effects factors" as $\boldsymbol{\eta}_2$, where the complete vector of endogenous latent variables is then $\boldsymbol{\eta}' = \boldsymbol{\eta}_1' \mid \boldsymbol{\eta}_2'$. The means for $\boldsymbol{\eta}_1$ are freely estimated but

their variances and covariances are constrained to zero. In contrast, the means of $\boldsymbol{\eta}_2$ are constrained to zero but their variances and covariances are freely estimated. The design matrices for the fixed and random effects can then be represented simultaneously in the factor loading matrix as $\boldsymbol{\Lambda}_{yj} = \mathbf{X}_j | \mathbf{Z}_j$.[6] The means of $\boldsymbol{\eta}_1$ are then equivalent to $\boldsymbol{\beta}$ and the factor variances and covariances of $\boldsymbol{\eta}_2$ are equal to $\mathbf{T}$. Note that cross-level interactions can be represented in $\boldsymbol{\eta}_1$ with corresponding columns in $\mathbf{X}_j$ equal to the product of the Level 1 and Level 2 covariates.

To demonstrate this technique, consider the simple example given earlier of three male and female students sampled per school, but suppose now that class size has been added as a Level 2 predictor of the sex effect (π_{1j}). The reduced-form model is then

$$y_{ij} = [\beta_0 + \beta_{10}sex_{ij} + \beta_{11}(sex_{ij} * size_j)] + [u_{0j} + u_{1j}sex_{ij}] + r_{ij}, \quad (26)$$

where the fixed and random effects are segregated in the bracketed terms. The measurement model of the equivalent SEM can then be written

$$\begin{pmatrix} lang_{m1_j} \\ lang_{m2_j} \\ lang_{m3_j} \\ lang_{f1_j} \\ lang_{f2_j} \\ lang_{f3_j} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & size_j & 1 & 1 \\ 1 & 1 & size_j & 1 & 1 \\ 1 & 1 & size_j & 1 & 1 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ - \\ \eta_4 \\ \eta_5 \end{pmatrix} + \begin{pmatrix} r_{m1_j} \\ r_{m2_j} \\ r_{m3_j} \\ r_{f1_j} \\ r_{f2_j} \\ r_{f3_j} \end{pmatrix}, \quad (27)$$

where the factor loading matrix has been partitioned into $\mathbf{X}_j | \mathbf{Z}_j$ and the latent variable vector has been partitioned into $\boldsymbol{\eta}'_1 | \boldsymbol{\eta}'_2$, representing the fixed and random effects, respectively. Note that the third column of the factor loading matrix is specified to capture the cross-level interaction of class size with sex in the prediction of language ability. Due to the inclusion of this effect, the model would have to be estimated using a case-varying factor loading matrix.

The mean vector and covariance matrix of the latent variables are then

$$\boldsymbol{\alpha}' = (\beta_0 \quad \beta_{10} \quad \beta_{11} \quad 0 \quad 0) \quad (28)$$

$$\boldsymbol{\Psi} = \begin{pmatrix} 0 & & & & \\ 0 & 0 & & & \\ 0 & 0 & 0 & & \\ 0 & 0 & 0 & \tau_{00} & \\ 0 & 0 & 0 & \tau_{10} & \tau_{11} \end{pmatrix}, \quad (29)$$

illustrating that the latent factors in $\boldsymbol{\eta}_1$ capture the fixed effects (with nonzero means but zero variances) and the factors in $\boldsymbol{\eta}_2$ capture the random effects (with zero means but nonzero variances and covariances).

The Observed Variable Solution

The observed variable solution to the problem generalizes the approach used to incorporate predictors in latent growth curve models. Specifically, Level 2 predictors are incorporated as exogenous observed variables (in contrast to drawing the values of the predictors into the factor loading matrix). These observed variables are assumed to be fixed and measured without error, so they may be substituted for ξ in the latent variable model (Equation 1). The direct effects of the Level 2 predictors on η (representing the random intercepts and slopes) are then captured as regression coefficients in the Γ matrix. As Curran, Bauer & Willoughby (now in press) have pointed out, with this parameterization of the model, the cross-level interaction of the standard multilevel linear model is computed as the indirect effect of the predictor on the outcome via the latent factors, or the product of Γ and Λ_{yj}. Note that with this approach the random coefficients can still be represented as latent factors with both means and variances (i.e., no partitioning of η is required).

To illustrate this approach, consider the same model as above, where class size is added as a Level 2 predictor of the sex effect on language ability. Using the observed variables approach, the measurement model can be left unchanged from Equation 19. To estimate the model, the latent variable model in Equation 20 is simply modified to

$$\begin{pmatrix} \pi_{0j} \\ \pi_{1j} \end{pmatrix} = \begin{pmatrix} \beta_0 \\ \beta_{10} \end{pmatrix} + \begin{pmatrix} 0 \\ \beta_{11} \end{pmatrix} (size_j) + \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix}. \tag{30}$$

Note that class size is represented directly as an observed predictor variable rather than via the factor loadings. This allows the cross-level interaction to be captured as a coefficient in the Γ matrix rather than as a latent variable mean. The Ψ matrix remains unchanged from Equation 21.

Comparing the Two Approaches

While true to the form of the linear mixed-effects model, there are three primary drawbacks to the latent variable approach of Rovine & Molenaar (1998, 2000, 2001). First, representing fixed and random effects as separate sets of latent factors is nonintuitive and can greatly expand the size of the model. Second, the values of the upper-level predictors (like the lower-level predictors) must be incorporated into the factor loading matrix, increasing the difficulty of parameterizing the model and potentially slowing the evaluation of the likelihood function. This may also make it difficult to conduct tests of overall model fit (i.e., use the missing data approach for unbalanced data). Third, by representing the upper-level covariates in the factor loading matrix, there is little possibility to elaborate the model for the predictors. The observed variable solution resolves these problems. It provides a parsimonious and intuitive representation of the model. Moreover, because the upper-level covariates are treated as observed variables it is easy to extend the model by including a measurement model for the predictors or mediational pathways between predictors.

Both approaches are readily generalizable to three-level models and beyond, though the parameterization of the model becomes increasingly complex.

In the following section I present an empirical example of a three-level model that includes a mediational model for the predictors at the highest level using the observed variables approach. This example provides a vehicle for demonstrating some of the advantages of estimating MLMs as SEMs.

Extending the Model

To this point I have considered how MLMs may be parameterized and estimated as SEMs even when the data are unbalanced. The only advantage I have shown for doing so is that under certain circumstances an inferential test of overall model fit can be obtained. However, the primary motivation for estimating MLMs as SEMs is to take full advantage of the flexibility the SEM framework, including the ability to estimate measurement models for the outcomes or predictors and the ability to estimate the paths in a causal chain directly. I now consider an empirical example that involves the estimation of a three level model where the first level is a factor model for psychometric scale data. I follow this by adding a mediational model for the upper-level predictors of the latent factors.

Multilevel Factor Analysis

The data is drawn from the High School and Beyond Administrator and Teacher Survey: Teacher Questionnaire conducted in 1984 and includes 10,365 teachers from 456 schools. I center my analysis on the Teacher Control scale, as described by Pallas (1988), which includes 9 items, reproduced in Table 2, which were designed to assess teacher perceptions of control of school policy (4 items) and the classroom (5 items). All items were rated on six-point Likert scales, and for the purposes of this analysis were grand-mean centered at zero. The data is highly unbalanced at the teacher level: The number of teachers sampled per school ranged from 1 to 30. There was also a small amount of missing data at the item-level (1% of responses). All available data were included in the analyses (92,334 responses).

It is interesting to note that the original measurement properties of the Teacher Questionnaire were assessed by conducting exploratory factor analysis without regard to the nested structure of the data (Stern & Williams, 1986, p. 241–245). More recently, taking a confirmatory approach, Raudenbush, Rowan, & Kang (1991) reanalyzed the Teacher Survey data to demonstrate how measurement models could be incorporated in conventional multilevel analyses by using a three-level model. It is useful to contrast this approach with the estimation of the model as an SEM.

Consider a two-factor model for the Control Scale. In the multilevel approach, the first level of the model constitutes the measurement model and would be written

$$y_{ijk} = \pi_{1jk}x_1 + \pi_{2jk}x_2 + r_{ijk}, \tag{31}$$

where *i, j* and *k* index items, teachers, and schools, respectively, and x_1 and x_2 are specified to hold the values of the factor loadings for the two latent factors,

TABLE 2
Items Selected for Analysis from the High School and Beyond Administrator and Teacher Survey: Teacher Questionnaire (1984)

Questions Assessing Teacher Perceptions of Control of School Policy

How much influence do teachers have over school policy in each of the areas below?
1. Determining student behavior codes
2. Determining the content of in-service programs
3. Setting policy on grouping students in classes by ability
4. Establishing the school curriculum

Questions Assessing Teacher Perceptions of Control over Classroom

How much control do you feel you have in your classroom over each of the following areas of your planning and teaching?
5. Selecting textbooks and other instructional materials
6. Selecting content, topics and skills to be taught
7. Selecting teaching techniques
8. Disciplining students
9. Determining the amount of homework to be assigned

Note. All items answered on a 6-point Likert scale.

represented as π_{1jk} and π_{2jk}. Note that this approach reverses the procedure illustrated here for parameterizing MLMs as SEMs. Here the latent factors are represented as random regression coefficients and the factor loadings as the values of the Level 1 covariates. The key difference is that in the multilevel approach the values of the covariates are assumed to be known, so the factor loadings contained in x_1 and x_2 must be fixed *a priori*. Conventionally, values of 1s and 0s are assigned to designate whether the item loads on the factor or not. This practice parallels the use of "dummy" covariates to toggle between outcomes in multivariate MLMs (Goldstein, 1995).

Using this approach, the 2-factor model for the 9 items may be written

$$
\begin{pmatrix} y_{1jk} \\ y_{2jk} \\ y_{3jk} \\ y_{4jk} \\ y_{5jk} \\ y_{6jk} \\ y_{7jk} \\ y_{8jk} \\ y_{9jk} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \pi_{1jk} \\ \pi_{2jk} \end{pmatrix} + \begin{pmatrix} r_{1jk} \\ r_{2jk} \\ r_{3jk} \\ r_{4jk} \\ r_{5jk} \\ r_{6jk} \\ r_{7jk} \\ r_{8jk} \\ r_{9jk} \end{pmatrix},
\tag{32}
$$

or

$$\mathbf{y}_{jk} = \mathbf{X}_{jk}\boldsymbol{\pi}_{jk} + \mathbf{r}_{jk},$$

where π_1 represents the "control over school policy" factor and π_2 represents the "control over classroom" factor. I assume that the residuals are independent with different variances, or that $\boldsymbol{\Sigma}_r = DIAG(\sigma_1, \ldots, \sigma_9)$.

The teacher-level model is then

$$\pi_{1jk} = \beta_{1k} + u_{1jk}$$

$$\pi_{2jk} = \beta_{2k} + u_{2jk},$$

(33)

where β_{1k} and β_{2k} capture the averages of the latent factors within schools and u_{1jk} and u_{2jk} represent deviations of the factor scores of individual teachers from their school averages, having covariance matrix

$$\mathbf{T}_\pi = \begin{pmatrix} \tau_{\pi 11} & \\ \tau_{\pi 21} & \tau_{\pi 22} \end{pmatrix}.$$

(34)

Similarly, the school-level model is

$$\beta_{1k} = \varepsilon_{1k}$$

$$\beta_{2k} = \varepsilon_{2k},$$

(35)

where no intercepts are needed because the data are centered, and ε_{1k} and ε_{2k} represent variability in the mean factor scores of the schools, with covariance matrix

$$\mathbf{T}_\beta = \begin{pmatrix} \tau_{\beta 11} & \\ \tau_{\beta 21} & \tau_{\beta 22} \end{pmatrix}.$$

(36)

Of particular interest in this model is the proportion of factor variance that is due to between-school differences versus within-school differences. This is expressed by the intra-class correlation coefficient

$$ICC = \frac{\tau_\beta}{\tau_\beta + \tau_\pi},$$

(37)

or the between-school variance over the total variance of the factor. A high ICC suggests that there are systematic differences between schools in teacher perceptions of control and that these differences might be productively modeled as a function of school characteristics.

The same model can be estimated as an SEM, with the important exception that the factor loadings can be estimated and need not be fixed *a priori.* That is, the design matrix $\mathbf{X}_{jk}$ from Equation 32 can be redefined as

$$\mathbf{X}_{jk} = \begin{pmatrix} \lambda_1 & 0 \\ \lambda_2 & 0 \\ \lambda_3 & 0 \\ \lambda_4 & 0 \\ 0 & \lambda_5 \\ 0 & \lambda_6 \\ 0 & \lambda_7 \\ 0 & \lambda_8 \\ 0 & \lambda_9 \end{pmatrix}. \tag{38}$$

To identify the model and to set the metric for the latent variables, the factor loadings λ_1 and λ_5 will be fixed to 1 (see Bollen, 1989).

To parameterize the model as an SEM, I utilize the missing data approach for unbalanced data. The data is sorted so that each row corresponds to a school. The first nine entries on each row correspond to the responses of the first teacher in that school on the nine items of the Control scale, the second nine entries are for the second teacher, and so on for up to $J = 30$ possible teachers or $9 \times 30 = 270$ variable entries. The model is then parameterized in the same way as a second-order latent factor model, similar to a second-order latent curve model (see Sayer & Cumsille, 2002). The matrix notation for the model is unwieldy, so I do not present it here, but a path diagram is given in Figure 5.

There are two primary differences between the path diagram in Figure 5 and those presented previously. First, there is a measurement model for the outcomes, or the two control factors. This extends the SEM to a three-level model where the lowest level is the measurement model for the items, with the difference that here the factor loadings for the items are estimated. A second key difference is that the model in Figure 5 is multivariate; that is, there are two outcome factors. The key to parameterizing these aspects of the model is that the measurement model for the two sets of indicators is repeated over the J teachers of the model with equality constraints on the factor loadings ($\lambda_1, \ldots, \lambda_9$), residual variances of the items ($\sigma_1, \ldots, \sigma_9$), and within-school variances and covariances of the factor scores for the teachers ($\tau_{\pi 11}, \tau_{\pi 21}$ and $\tau_{\pi 22}$). The common variance among teachers within a school on the two factors is then drawn into the second-order factors β_{1k} and β_{2k} with variances and covariances equal to $\tau_{\beta 11}, \tau_{\beta 21}$ and $\tau_{\beta 22}$. No means or intercepts were included in this model because the items were centered at zero prior to the analysis.

First, I contrast the multilevel approach to measurement models with the SEM approach. Using the Control Scale data, a three-level multilevel linear model was
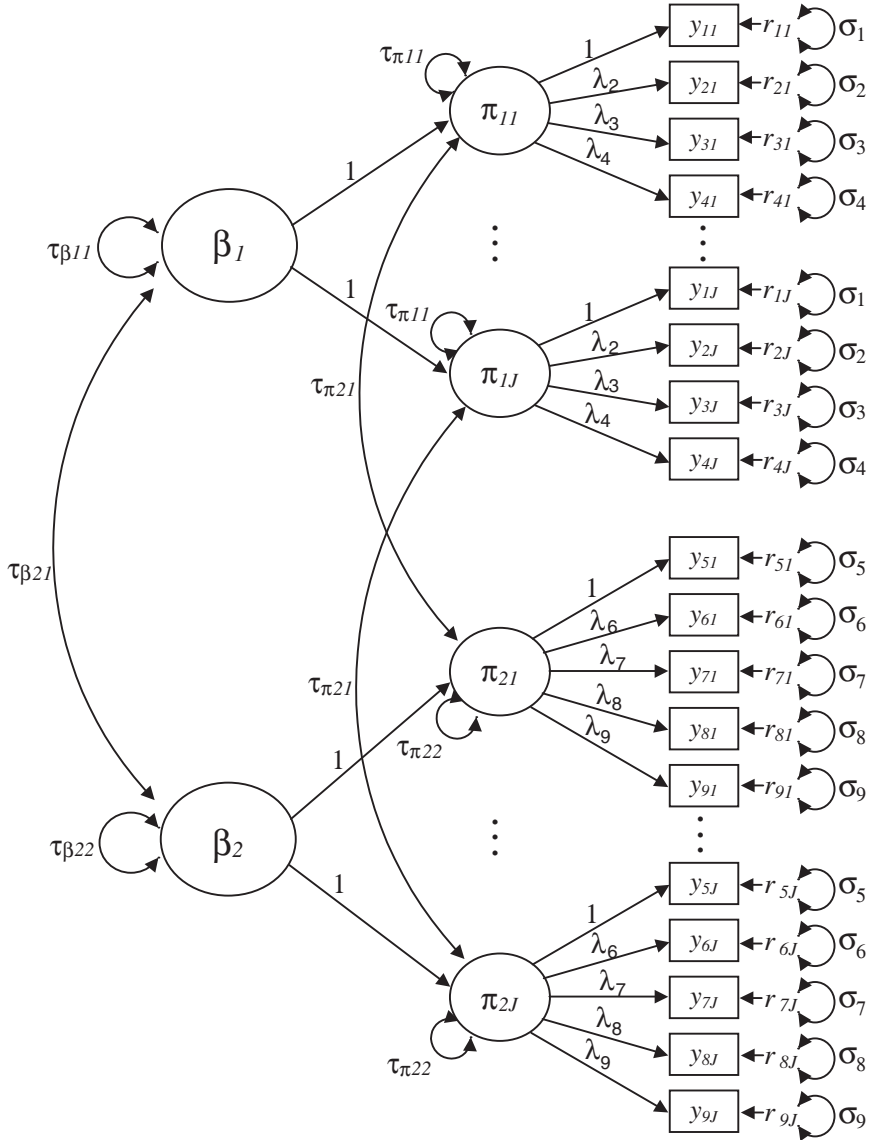
FIGURE 5. *A path diagram of a 2-Factor multilevel linear structural equation model with estimated factor loadings and unbalanced data treated as missing.*

estimated where the factor loadings for the items were set to 1s and 0s as is conventional for this modeling approach. An SEM was then estimated, parameterized according to the path diagram in Figure 5; that is, with freely estimated loadings for the nonscaling indicators. The estimated loadings from the SEM analysis, presented in Figure 6, indicated that factor loadings of 1s are adequate for some items (items 2 and 6) but not others (items 3, 4, 7, 8 and 9). The apparent consequence of using unit loadings in the MLM was to distort the estimation of the factor variances, as can be seen in Figure 7 which compares the estimated within- and between-school factor variances and covariances. For instance, the total variance of the control over policy factor was estimated at .89 using unit loadings compared to .73 with estimated loadings. Similarly, the total variance in the control over classroom factor was estimated at .34 using unit loadings compared to .80 with estimated loadings. Of particular interest were the ICCs, which for the SEM analysis were estimated as .28 for the policy factor and .19 for the classroom factor. These are sizeable values, reflecting the disattenuation of the ICC due to the removal of measurement error (Raudenbush, Rowan & Kang, 1991).

To validate the results of the SEM analysis empirically, a second MLM was estimated, only this time the factor loadings were set to their estimated values from
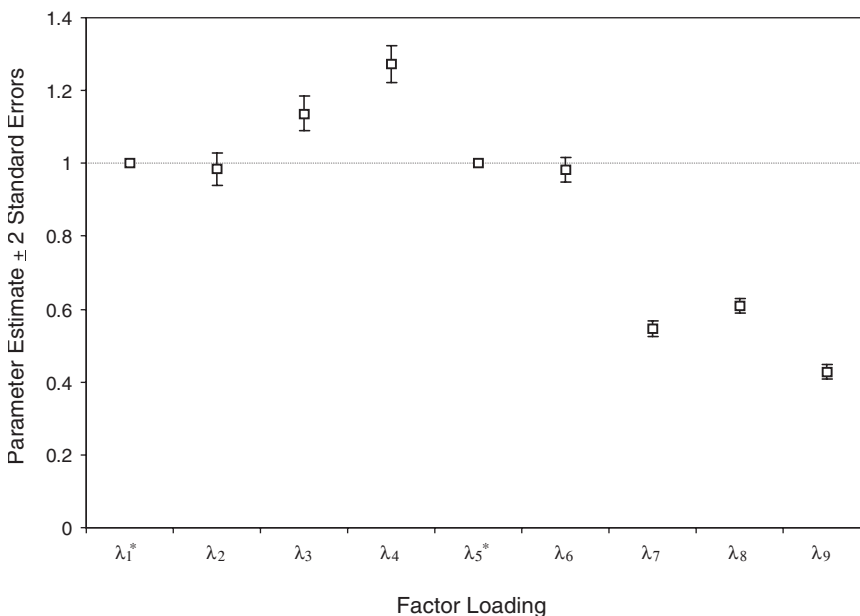

FIGURE 6. *Factor loading estimates ± 2 standard errors obtained by fitting the model displayed in Figure 5 in Mplus. Starred factor loadings were fixed to 1 to set the scale of the latent factors. The dotted line shows the reference unit loading value of 1 assumed in many multilevel linear modeling approaches to multilevel factor analysis.*
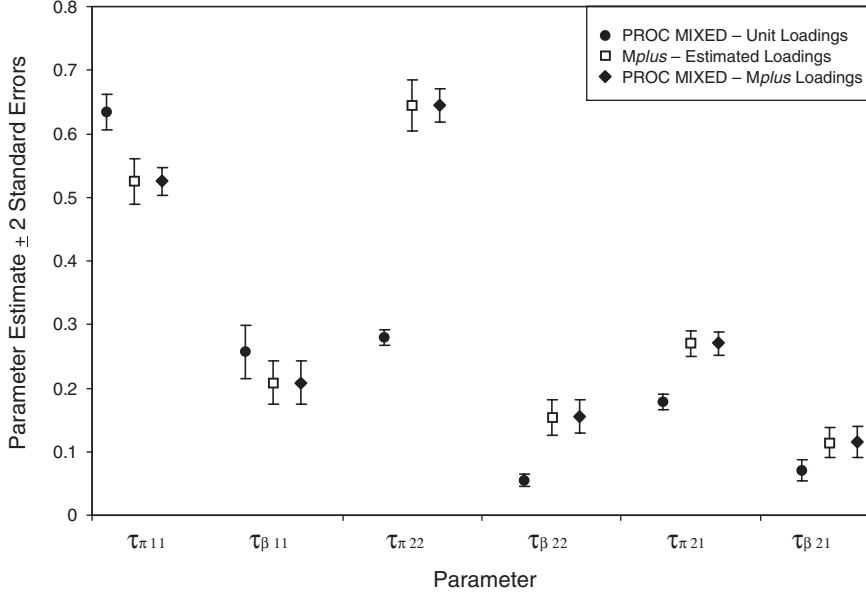
FIGURE 7. *Parameter estimates ± 2 standard errors of the within-school and between-school variances and covariances of the latent factors of the model displayed in Figure 5 as obtained by fitting the model in PROC MIXED using traditional unit loadings, in Mplus, freely estimating loadings λ_2 through λ_4 and λ_6 through λ_9, and in PROC MIXED fixing the loadings to the values estimated in Mplus.*

the SEM analysis. Note that by making this change the remaining parameter estimates for the MLM are almost identical to the estimates from the SEM, as shown in Figure 7. It is worth noting, however, that the standard errors for the parameter estimates are not equivalent. The standard errors for the MLM fail to take account of the uncertainty in the estimation of the factor loadings and are thus slightly smaller than they should be.

Finally, the fit of the model was tested. For this data set the covariance matrix for the constrained saturated model consists of

$$
\hat{\boldsymbol{\Sigma}}_{(270 \times 270)} = \begin{pmatrix} \hat{\boldsymbol{\Sigma}}_{\mathbf{T}(9 \times 9)} & & & \\ \hat{\boldsymbol{\Sigma}}_{\mathbf{B}(9 \times 9)} & \hat{\boldsymbol{\Sigma}}_{\mathbf{T}(9 \times 9)} & & \\ \vdots & \ddots & \ddots & \\ \hat{\boldsymbol{\Sigma}}_{\mathbf{B}(9 \times 9)} & \cdots & \hat{\boldsymbol{\Sigma}}_{\mathbf{B}(9 \times 9)} & \hat{\boldsymbol{\Sigma}}_{\mathbf{T}(9 \times 9)} \end{pmatrix},
\tag{39}
$$

where $\hat{\boldsymbol{\Sigma}}_{\mathbf{T}}$ is the within-teachers covariance matrix for the 9 items and is constrained to be equal over the 30 blocks on the diagonal of $\hat{\boldsymbol{\Sigma}}$ (corresponding to the 30 possible teachers within each school). Note that $\hat{\boldsymbol{\Sigma}}_{\mathbf{T}}$ represents the total variance

159

and covariance in the set of nine items over teachers and hence takes on the same values as one would obtain by calculating a covariance matrix on the total sample of teachers aggregated over schools. The covariance between items among different teachers of the same school is captured in $\hat{\mathbf{\Sigma}}_B$, which accounts for the dependence introduced by systematic between-school differences in teacher responses. No mean vector is necessary since means were not modeled. Unfortunately, the large size of this model, involving 90 free parameters, a 270×270 covariance matrix, and over 92,000 responses, exceeded the memory resources of two SEM software packages (M*plus* version 2.1 and Mx version 1.52a).[7]

Thus, to test the fit of the hypothesized model, both the multilevel factor model and constrained saturated model were fit to a reduced data set, consisting only of the scores from the first seven teachers of each school. Although not ideal, this provided a good indication of the fit of the hypothesized model. The likelihood-ratio test calculated by fitting the models to the reduced sample was $\chi^2(68) = 1352$; $p < .001$, indicating quite poor fit. Note that this information on the poor fit of the model typically would not be available if a multilevel framework were used to fit the model. Within the SEM tradition, however, a significant χ^2 test of model fit often initiates the search for an alternative, hopefully better-fitting, theoretical model. For instance, an alternative 2-factor model might group items into curriculum-oriented questions (e.g., items 4, 5, 6, 7) and discipline-oriented questions (e.g., items 1 and 8). For didactic purposes, however, I retain the current model (as depicted in Figure 5) in the subsequent analysis despite its poor fit.

Prediction and Mediation

I now extend the multilevel factor analysis presented previously to a full SEM including school-level predictors, following the observed variables approach to incorporating upper-level predictors. A mediational model is posited where the effect of one school-level predictor, whether the school is public or private, is partially explained by another school-level variable, school size. The school-level data was obtained from the High School and Beyond First Follow-Up: School Questionnaire administered in 1982, two years prior to the Administrator and Teacher Survey. Only data for the 456 schools in the preceding analysis were included (using all available teacher data). The two predictors of interest were whether the school was public (82%) or private (18%), and school size, which ranged from 21 to 5,342 with a mean of 1,254. To ease computations, school size was rescaled by dividing by 100, so that each unit represents 100 students. There was missing data for 13 schools for the public/private variable and missing data on school size for 42 schools. Because the SEM approach utilizes a joint likelihood function for the predictors and outcomes, missing data can be accommodated for both types of variables.[8] All 456 schools were thus retained in the analysis.

One interesting question is whether school size completely explains the difference between public and private schools in teacher perceptions of control, or only partially explains the difference. To test between these possibilities, two models

were estimated. The path diagram in Figure 8 presents the predictive relationships of the *full model,* allowing both direct and indirect effects of school status on teacher perceptions of control (note that the teacher-level and item-level of the model are not reproduced in Figure 8 since they are unchanged from Figure 5). In the *reduced model,* the dashed arrows representing the direct effects of school status independent of school size were removed. The full and reduced models are nested, allowing for a formal test of the hypothesis that the difference between public and private schools in teacher perceptions of control is entirely due to differences in school size. The likelihood-ratio test, $\chi^2(2) = 70$; $p < .001$, rejected this hypothesis in favor of the view that school size only partially mediates the effect of employment at a public versus private school.

The parameter estimates for the mediational paths in the full model are also presented in Figure 8 (the parameter estimates for the remainder of the model changed little from Figures 6 and 7 and are not reported). An examination of these values reveals the same general pattern for the two control factors, but also subtle differences. First, teachers in private schools perceived themselves to have more control over school policy and their classrooms than teachers in public schools. Specifically, on a six-point scale, teacher perceptions of control over school policy were on average .6 units higher in private schools than public schools [where the total effect is calculated as the direct effect plus the indirect effect, or $.51 + (-7.50)(-.01)$]. This difference accounted for 28% of the between-school variance in the Policy
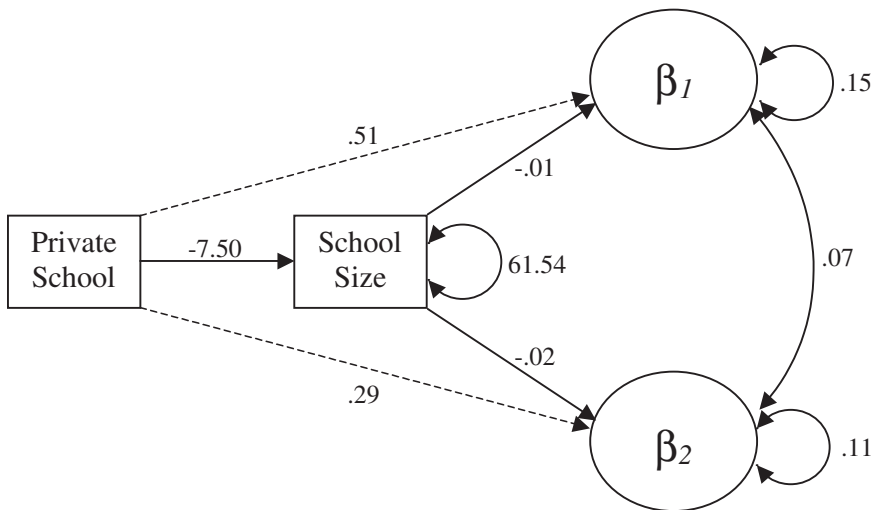

FIGURE 8.   *A path diagram of the latent variable model for a 2-Factor multilevel structural equation model where the effect of the School-Level predictor Private School is mediated by School Size (The teacher and item levels of the model are identical to Figure 5 so are not displayed). Paths are labeled with their estimated parameter values (all displayed parameter estimates significantly differ from zero at* p *< .05).*

factor. Similarly, teachers in private schools perceived themselves to have greater control over their classrooms by an average of .4 units [calculated as .29 + (−7.50)(−.02)], a difference that explained 33% of the between-school variance in the Classroom factor.

Second, the results supported the hypothesis that the differences between public and private schools are partially due to differences in school size. On average, 750 fewer students were enrolled at private schools than public schools. Further, for each 100 students enrolled in a school, teacher perceptions of control over school policy dropped by a tenth of a point, while teacher perceptions of control over the classroom dropped by two-tenths of a point (again on six-point scales). Given these values, only 14% of the total difference between public and private school teachers' perceptions of control over school policy can be attributed to differences in school size. In contrast, a much larger percentage, 34%, of the total difference between public and private school teachers' perceptions of control over the classroom can be explained by the fact that private schools tend to have fewer students. Thus the results suggest that while school size partially mediates the effect of employment in a private versus public school for both factors, its role is much more prominent for teacher perceptions of control over the classroom.[9] Note that these same results generally could not be obtained directly using a conventional multilevel approach, though indirect procedures are available for calculating some of them (Krull & MacKinnon, 1999, 2001; Raudenbush & Sampson, 1999).

Limitations and Alternative Approaches

The approach illustrated here for estimating MLMs as SEMs integrates many of the strengths of the two modeling frameworks. Three key extensions of the basic multilevel linear model made possible by the SEM framework are formal tests of model fit, measurement models for the outcomes, and mediational models among predictors. Many other extensions are also possible. These include measurement models for the predictors, models for multiple samples (Jöreskog, 1971), and finite mixture models (Arminger & Stein, 1997; Jedidi, Jagpal, & DeSarbo, 1997; Muthén & Shedden, 1999), among other possibilities. There is nothing in principle that prevents the estimation of similar models within the multilevel framework, but currently many of these features are only available within the SEM.

Likewise, many features of the multilevel modeling framework are currently unavailable within the SEM. For instance, many multilevel *nonlinear* models cannot be parameterized as SEMs, since structural equation modeling is predicated on a linear system of equations.[10] Similarly, generalized linear mixed-effects models with link functions for binary, ordinal, or count response scales cannot currently be estimated as SEMs. Traditionally, SEM has taken a much different approach to handling these different response scales. Specifically, it is assumed that binary and ordinal data represent crude categorizations of underlying continuous variables. The relationships of these underlying continuous variables are captured in a polychoric correlation or covariance matrix, and the model is fit to this matrix with special modifications to the fit function (Muthén, 1984).

Other aspects of the two modeling frameworks also differ. For instance, not all estimators are available in each modeling framework (e.g., the restricted maximum likelihood estimator commonly used with multilevel models currently has no counterpart in SEM). Also, while Wald tests of the variance components are typical of the SEM approach and some MLMs, other MLMs utilize a chi-square test of the residuals to account for the skewed sampling distributions of the parameters (Bryk & Raudenbush, 1992, p. 55). These differences largely reflect *modeling traditions* rather than the constraints of the analytic models themselves. It is hoped that by illustrating the commonalities between the two approaches, advances made in each modeling framework will inform the other, as has been the case in the growth modeling literature where the identity between the two approaches has been known for many years. As one example, it is common for MLMs to be presented with formal assumptions of normality for the residuals and random effects, as was done here (e.g., Raudenbush, 2001; Verbeke & Molenberghs, 2000). However, it is widely known in the SEM community that the maximum likelihood estimator retains many of its desirable properties under the less stringent assumption of no excessive kurtosis for the endogenous observed variables (Browne, 1984).

Multilevel latent variable models such as the two factor model presented above have been the topic of many other recent analytic developments. There are two primary alternative approaches to the one illustrated here. The first, originating with Goldstein & MacDonald (1988), involves partitioning the total population covariance matrix Σ_T (aggregated over all individuals) into two additive components: Σ_W representing within-groups covariance in the set of items (or the covariance of individual deviations from their group means) and Σ_B representing between-groups covariance (or the covariance of the group means of the items). Muthén (1994) demonstrated how sample estimates of the within- and between-groups covariance matrices can be obtained and proposed a quasi-maximum likelihood estimator (MUML) for estimating the model on the two component covariance matrices simultaneously. The primary problem with MUML is that it assumes a balanced design to make the estimation of the model more tractable. More recently, however, full-information maximum likelihood (FIML) routines have become available that can accommodate unbalance designs (Bentler & Liang, 2003; du Toit & du Toit, in press; Muthén & Muthén, 2002). In comparison, the approach illustrated here also accommodates unbalanced designs but models Σ_T and Σ_B directly, with no prior decomposition of Σ_T, as can be seen in Equation 39. It thus provides a more natural analog to conventional MLMs. Further research is needed to more thoroughly compare these alternatives both analytically and in terms of their relative performance in finite samples. It is important to note, however, that the within- and between-groups approach to multilevel SEM, unlike the one illustrated here, cannot currently be combined with some of the other features of the SEM framework (e.g., multiple groups analysis, mixture modeling).

A second alternative approach to multilevel SEM was recently proposed by Rabe-Hesketh and colleagues in their Generalized Linear Latent and Mixed Models (GLLAMM) STATA macro (Rabe-Hesketh, Pickles, & Taylor, 2000;

163

Rabe-Hesketh, Pickles, & Skrondal, 2001; Rabe-Hesketh, Skrondal, & Pickles, in press). The key advantage of GLLAMM is that it permits link functions to be used for outcomes on different response scales, such as binary or count data. However, because the derivatives in the likelihood function are approximated, it is very slow and is not recommended for continuous data (Rabe-Hesketh, Pickles, & Skrondal, 2001). GLLAMM may best be viewed as complementary to other approaches to multilevel SEM, to be used when outcomes are not continuous. The rapid pace of these developments attests to the continuing convergence of multilevel modeling and SEM, and the new modeling possibilities that an integrative framework may offer to applied researchers.

Notes

[1] Any pair of nested models can be compared by the likelihood-ratio test (subject to certain regularity conditions), making possible formal contrasts between alternative models.

[2] This differs from the traditional CFA model in which factor loadings are estimated and relect the regression of the observed variables on the latent variable.

[3] SAS code and M*plus* input files for all of the example analyses can be downloaded from the author at http://www4.ncsu.edu/~djbauer. Note that these same analyses could be conducted in other multilevel or SEM software programs offering the same capabilities.

[4] Tabled values for the parameter estimates, standard errors, and log-likelihoods of this and subsequent models may also be downloaded from the author at http://www4.ncsu.edu/~djbauer.

[5] Currently, only two SEM software programs, Mx and M*plus* have this capability.

[6] Although Rovine & Molenaar (1998, 2000, 2001) do not do so, to achieve maximum generality, I have subscripted the factor loading matrix by *j* to indicate that it may vary case-wise. This subscript could be removed for balanced data or if using the missing data approach to unbalanced data.

[7] This reflects the fact that SEM software was not initially designed to fit models of this type and so does not always efficiently analyze partitioned matrices as in Equation 39. Future developments are likely to overcome this difficulty (L. Muthén, personal communication).

[8] This contrasts with the conditional likelihood function typically used in the multilevel framework which requires complete data on the predictors.

[9] Of course, these conclusions must be tempered by the poor fit of the underlying factor model.

[10] To be more precise, if the response is a nonlinear function of the random effects (e.g., a negatively accelerated exponential function with random parameters) the model cannot currently be parameterized as an SEM. Alternatively, if the nonlinearity is only in the fixed effects part of the model, then the model can be parameterized as an SEM by using nonlinear constraints (see du Toit & Cudeck, 2001). Further, many nonlinear functions can be reasonably approximated by poly-

nomial functions (which are linear in their parameters and hence can be estimated as SEMs).

References

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides, and R. E. Schumacker (Eds.), *Advanced structural equation modeling* (pp. 243–277). Mahwah, NJ: Erlbaum.

Arminger, G., & Stein, P. (1997). Finite mixtures of covariance structure models with regressors. *Sociological Methods and Research, 26,* 148–182.

Bauer, D. J., & Curran, P. J. (June, 2002). *Estimating multilevel linear models as structural equation models.* Paper presented at the annual meeting of the Psychometric Society, Chapel Hill, NC.

Bentler, P. M., & Liang, J. (2003). Two-level mean and covariance structures: Maximum likelihood via an EM algorithm. In S. P. Reise & N. Duan (Eds.), *Multilevel modeling: Methodological advances, issues, and applications* (pp. 53–70). Mahwah, NJ: Erlbaum.

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

Browne, M. W. (1984). Asymptotic distribution free methods in analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology, 37,* 62–83.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: applications and data analysis methods.* Newbury Park, CA: Sage Publications.

Chou, C-P., Bentler, P. M., & Pentz, M. A. (1998). Comparisons of two statistical approaches to study growth curves: The multilevel model and the latent curve analysis. *Structural Equation Modeling, 5,* 247–266.

Curran, P. J., Bauer, D. J., & Willoughby, M. T. (in press). Testing and probing main effects and interactions in latent trajectory models. *Psychological methods.*

du Toit, S. H. C., & Cudeck R. (2001). The analysis of nonlinear random coefficient regression models with LISREL using constraints. In Cudeck, R., du Toit, S. H. C. and Sorbom, D. (Eds.) *Structural equation modeling: Present and future* (pp. 259–278). Lincolnwood, IL: Scientific Software International.

du Toit, S. H. C., & du Toit, M. (in press). Multilevel structural equation modeling. In J. de Leeuw & I. G. G. Kreft (Eds.), *Handbook of quantitative multilevel analysis.* Boston, MA: Kluwer Academic Publishers.

Gelman, A., Pasarica, C., & Dodhia, R. (2002). Let's practice what we preach: Turning tables into graphs. *The American Statistician, 56,* 121–130.

Goldstein, H. I. (1995). *Multilevel statistical models.* Kendall's Library of Statistics 3. London: Arnold.

Goldstein, H. I., & McDonald, R. P. (1988). A general model for the analysis of multilevel data. *Psychometrika, 53,* 455–467.

Jedidi, K., Jagpal, H. S., & DeSarbo, W. S. (1997). Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science, 16,* 39–59.

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36,* 409–426.

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8.* Mooresville, IN: Scientific Software Inc.

Krull, J. L., & MacKinnon, D. P. (1999). Multilevel mediation modeling in group-based intervention studies. *Evaluation Review, 23,* 418–444.

Krull, J. L., & MacKinnon, D. P. (2001). Multilevel modeling of individual and group level mediated effects. *Multivariate Behavioral Research, 36,* 249–277.

Laird, N. M., & Ware, J. H. (1982). Random effects models for longitudinal data. *Biometrics, 38,* 963–974.

MacCallum, R. C., Kim, C., Malarkey, W., & Kielcolt-Glaser, J. (1997). Studying multivariate change using multilevel models and latent curve models. *Multivariate Behavioral Research, 32,* 215–253.

Mehta, P. D., & West, S. G. (2000). Putting the individual back into individual growth curves. *Psychological Methods, 5,* 23–43.

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

Meredith, W., & Tisak, J. (1984, July). *"Tuckerizing" growth curves.* Paper presented at the annual meeting of the Psychometric Society, Santa Barbara, CA.

Muthén, B. O. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika, 49,* 115–132.

Muthén, B. O. (1994). Multilevel covariance structure analysis. *Sociological Methods & Research, 22,* 376–398.

Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. In P. Marsden (Ed.), *Sociological Methodology 1995,* (pp. 215–216).

Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics, 55,* 463–469.

Muthén, L. K., & Muthén, B. O. (2002). *Mplus user's guide* [Computer Manual, Version 2.1]. Los Angeles, CA: Muthén & Muthén.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (1999). *Mx: Statistical Modeling* [Computer Manual, 5th Edition]. Richmond, VA: Virginia Commonwealth University, Department of Psychiatry.

Newsom, J. T. (2002). A multilevel structural equation model for dyadic data. *Structural Equation Modeling, 9,* 431–447.

Pallas, A. M. (1988). School climate in American high schools. *Teachers College Record, 89,* 541–553.

Rabe-Hesketh, S., Pickles, A., & Skrondal, A. (2001). GLLAMM: A class of models and a Stata program. *Multilevel Modeling Newsletter, 13,* 17–23.

Rabe-Hesketh, S., Pickles, A., & Taylor, C. (2000). Generalized linear latent and mixed models. *Stata Technical Bulletin 53,* 47–57.

Rabe-Hesketh, S., Skrondal, A. & Pickles, A. (in press). Generalized multilevel structural equation modelling. *Psychometrika.*

Raudenbush, S. W. (2001). Toward a coherent framework for comparing trajectories of individual change. In L. M. Collins and A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 35–64). Washington, DC: American Psychological Association.

Raudenbush, S. W., Rowan, B., & Kang, S. J. (1991). A multilevel, multivariate model for studying school climate with estimation via the EM algorithm and application to U.S. high-school data. *Journal of Educational Statistics, 16,* 295–330.

Raudenbush, S. W., & Sampson, R. (1999). Assessing direct and indirect effects in multilevel designs with latent variables. *Sociological Methods & Research, 28,* 123–153.

Rovine, M. J., & Molenaar, P. C. M. (1998). A nonstandard method for estimating a linear growth model in LISREL. *International Journal of Behavioral Development, 22,* 453–473.

Rovine, M. J., & Molenaar, P. C. M. (2000). A structural modeling approach to a multilevel random coefficients model. *Multivariate Behavioral Research, 35,* 51–88.

Rovine, M. J., & Molenaar, P. C. M. (2001). A structural equations modeling approach to the general linear mixed model. In L. M. Collins and A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 65–96). Washington, DC: American Psychological Association.

Sayer, A. G., & Cumsille, P. E. (2002). Second-order latent growth models. In L. M. Collins and A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 179–200). Washington, DC: American Psychological Association.

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis.* London: Sage Publications.

Stern, J., & Williams, M. (Eds.). (1986). *The condition of education.* Washington, DC: U.S. Government Printing Office.

Verbeke, G., & Molenberghs, G. (2000). *Linear mixed models for longitudinal data.* New York: Springer-Verlag.

Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin, 116,* 363–381.

Wothke, W. (2000). Longitudinal and multi-group modeling with missing data. In T. D. Little, K. U. Schnabel and J. Baumert (Eds.) *Modeling longitudinal and multiple group data: Practical issues, applied approaches, and specific examples* (pp. 219–240). Hillsdale, NJ: Erlbaum.

Author

DANIEL J. BAUER is Assistant Professor, Psychology Department, North Carolina State University, Raleigh, NC 27695-7801; dan_bauer@ncsu.edu. His areas of specialization are quantitative methodology and developmental psychology.