

Differentiating between mixed-effects and latent-curve approaches to growth modeling

Daniel McNeish¹ · Tyler Matta²

Published online: 24 October 2017
© Psychonomic Society, Inc. 2017

Abstract In psychology, mixed-effects models and latent-curve models are both widely used to explore growth over time. Despite this widespread popularity, some confusion remains regarding the overlap of these different approaches. Recent articles have shown that the two modeling frameworks are mathematically equivalent in many cases, which is often interpreted to mean that one’s choice of modeling framework is merely a matter of personal preference. However, some important differences in estimation and specification can lead to the models producing very different results when implemented in software. Thus, mathematical equivalence does not necessarily equate to practical equivalence in all cases. In this article, we discuss these two common approaches to growth modeling and highlight contexts in which the choice of the modeling framework (and, consequently, the software) can directly impact the model estimates, or in which certain analyses can be facilitated in one framework over the other. We show that, unless the data are pristine, with a large sample size, linear or polynomial growth, and no missing data, and unless the participants have the same number of measure-

ments collected at the same set of time points, one framework is often more advantageous to adopt. We provide several empirical examples to illustrate these situations, as well as ample software code so that researchers can make informed decisions regarding which framework will be the most beneficial and most straightforward for their research interests.

Keywords Mixed effect model · Latent growth model · Multilevel model · Hierarchical linear model · Structural equation modeling

In psychology, empirical questions concerned with change over time are ubiquitous. As such, statistical methods for modeling longitudinal data concomitant with such questions have become widely studied in the methodological literature. For modeling growth, researchers typically employ some type of random effects model such that a mean growth trajectory is estimated for all observations in the data but a unique growth curve is estimated for each individual in the data as well (Curran & Bauer, 2011). These types of models are generally referred to as *subject-specific models* (Zeger, Liang, & Albert, 1988), which are more commonly known in psychology as *growth curve models*. Growth models have many aliases but can be broadly grouped into two different classes of methods: the latent-curve (LC) approach, which treats the repeated measures as multivariate (also known as the “wide” data format) and tends to be fit with general structural equation modeling (SEM) software (Meredith & Tisak, 1990; Tucker, 1958; Willett & Sayer, 1994), and the mixed-effect (ME) approach, which treats the repeated measures as univariate (also known as the “long” data format) and is generally fitted with

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13428-017-0976-5>) contains supplementary material, which is available to authorized users.

✉ Daniel McNeish
dmcneish@asu.edu

¹ Department of Psychology, Arizona State University, PO Box 871104, Tempe, AZ 85287, USA

² Centre for Educational Measurement, University of Oslo, Oslo, Norway

regression software (Bryk & Raudenbush, 1987; Laird & Ware, 1982; Rao 1965).¹

Over the past 20 years, methodological research has shown that the LC and ME approaches are actually nuanced twists on the same idea; they have been shown to converge, and to be mathematically equivalent in many cases (e.g., Bauer, 2003; Curran, 2003; Ledermann & Kenny, 2017; Mehta & Neale, 2005). Skrondal and Rabe-Hesketh (2004) went so far as to propose a general latent-variable modeling framework that effectively unified the LC and ME approaches. Historically, the treatment of this topic has focused on conditions for which the two growth-modeling frameworks converge. Bauer (2003) and Curran (2003) showed types of models for which the parameters in one framework mapped directly onto the parameters in the other. Chou, Bentler, and Pentz (1998) demonstrated that identical results can be obtained in either framework under certain conditions. Convergence has been shown similarly in the cases of missing data (Ferrer, Hamagami, & McArdle, 2004), interaction tests (Preacher, Curran, & Bauer, 2006), and subject-specific estimates (Mehta & West, 2000). Clearly, there is much overlap between the two modeling frameworks.

Although certainly there is a notable mathematical similarity between the LC and ME approaches, they are not always identical, and there remains confusion regarding when divergence occurs. We believe that this ambiguity does not lie in the mathematical model but is largely a result of software implementation, whereby some extensions are made easier under LC software, whereas others are more straightforward under the ME software. However, the recent literature has lost track of the extent to which the overlap extends, often resulting in claims that the modeling frameworks will produce identical results. For example, Hox (2010) states that “when equivalent multilevel regression analysis and latent curve modeling are applied to the same data set, the results are identical” (p. 243).

There is indeed much overlap between the frameworks, and estimates from both frameworks indeed do converge, given ideal data. However, in empirical studies, in which the data are often nonnormal or missing, are collected at nonuniform intervals, or have small samples, nontrivial differences between the frameworks will result in different conclusions, or at least in a varying difficulty with which the models can be fit. The goal of this article is to provide a clear overview of the LC and ME approaches to growth curve analysis and to highlight where these two approaches differ *practically*, rather than theoretically. We highlight some of the major differences in the implementation of these models in commonly used

statistical software, and include demonstrations of the differences via empirical examples. We conclude by offering a summary of the situations in which the methods are not equally advantageous to implement, and we provide recommendations regarding when each framework may provide particular advantages for empirical researchers fitting models to data. Thus, even though some studies have demonstrated the broad convergence of growth modeling frameworks, we feel that it is important to also demonstrate their meaningful divergence, as well.

Overview of the mixed-effect approach

The ME approach accounts for the fact that individuals are measured repeatedly over time by modeling the intercept and/or the time coefficients as random (Laird & Ware, 1982; Stiratelli, Laird, & Ware, 1984). This allows researchers to estimate a mean trajectory for the entire sample, as well as subject-specific deviations from the mean for each person in the data. The mean trajectory parameters for the whole sample are commonly referred to as “fixed effects.” These fixed effects—which are most commonly the intercept, time (and any functions thereof, such as polynomial terms), and any time-varying covariates—can also be included in the *random-effects* portion of the model. Random effects capture how much the estimates for a particular person differ from the fixed-effect estimate, which allows the growth trajectory to differ for each person.

More formally, with continuous outcomes, the linear ME model can be written in Laird and Ware’s (1982) notation as

$$y_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{u}_i + \boldsymbol{\varepsilon}_i, \quad (1)$$

where y_i is an $m_i \times 1$ vector of responses for person i , m_i is the number of observations for person i , \mathbf{X}_i is an $m_i \times p$ design matrix for the predictors in person i , p is the number of predictors (which includes the intercept), $\boldsymbol{\beta}$ is a $p \times 1$ vector of fixed regression coefficients, \mathbf{Z}_i is an $m_i \times r$ design matrix for the random effects of person i , r is the number of random effects, \mathbf{u}_i is an $r \times 1$ vector of random effects for person i , where $\mathbf{u}_i \sim \text{MVN}(\mathbf{0}, \mathbf{G})$, and $\boldsymbol{\varepsilon}_i$ is a matrix of residuals of the observations in person i , where $\boldsymbol{\varepsilon}_i \sim \text{MVN}(\mathbf{0}, \mathbf{R}_i)$ and $\text{Cov}(\mathbf{u}_i, \boldsymbol{\varepsilon}_i) = \mathbf{0}$.

As an alternative to the Laird and Ware (1982) matrix notation in Eq. 1, it is common in behavioral science literatures to see ME models written in Raudenbush and Bryk (2002; hereafter, RB) notation, as well. In RB notation, Eq. 1 can be written as

$$\begin{aligned} Y_{ij} &= \beta_{0i} + \beta_{1i}Time_{ij} + r_{ij}, \\ \beta_{0i} &= \gamma_{00} + \gamma_{01}X_{1i} + \dots + \gamma_{0k}X_{ki} + u_{0i}, \\ \beta_{1i} &= \gamma_{10} + \gamma_{11}X_{1i} + \dots + \gamma_{1k}X_{ki} + u_{1i}, \end{aligned} \quad (2)$$

¹ We adopt the ME and LC set of terminology advanced by Cudeck (1996), which was one of the first articles to contrast these methods, to differentiate between these two frameworks. However, we recognize that these analytical approaches carry many different monikers. For example, Skrondal and Rabe-Hesketh (2004) refer to these exact same approaches as *factor models* and *random coefficient models*, while Curran (2003) uses the terms *multilevel model* and *structural equation model*.

where Y_{ij} is the response for the i th person at the j th time, β_{0i} is the intercept for the i th person, β_{1i} is the slope for the i th person, γ are fixed-effect parameters that do not have an i subscript because they are constant for all people, $\mathbf{u}_i = (u_{0i}, u_{1i})^T$ is a vector of random effects for the i th person, Time_{ij} is the j th time point for the i th person, and r_{ij} is the residual for the i th person at the j th time. The covariance matrix for the random effects (u) is often represented by \mathbf{T} rather than \mathbf{G} in this notation. Similarly, the covariance matrix of the residuals (r) are represented by \mathbf{V} rather than \mathbf{R} .

Popular software programs used to fit linear ME models include, but are not limited to, SAS Proc Mixed, SPSS MIXED, Stata xtmixed, the lme4 package in R, and the HLM software program.

Overview of the latent-curve approach

The LC approach for growth essentially follows the same premise as ME models, except that growth is formulated in a general SEM framework rather than as an extension of the regression framework. Specifically, LC models are confirmatory factor analysis (CFA) models with an imposed factor mean structure and particular constraints to yield estimates of growth. The basic idea of the LC framework for growth is identical to that of the ME framework—there is some overall mean trajectory for the entire sample, but each individual receives random-effect estimate(s) to capture how their particular growth curve differs from the overall trajectory. The main conceptual difference is that the random effects are specified as latent variables in a CFA, rather than as randomly varying regression coefficients; however, these two notions can be shown to be mathematically equivalent (see, e.g., Curran, 2003).

In SEM matrix notation, the LC model for growth is written as

$$\begin{aligned} \mathbf{Y}_i &= \mathbf{\Lambda} \boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i, \\ \boldsymbol{\eta}_i &= \boldsymbol{\alpha} + \mathbf{\Gamma} \mathbf{X}_i + \boldsymbol{\zeta}_i, \end{aligned} \tag{3}$$

where $\mathbf{\Lambda}$ is a matrix of factor loadings that can be, but are not always, prespecified to fit a specific type of growth trajectory; $\boldsymbol{\eta}_i$ is a vector of subject-specific growth factor (intercept and slope) values for person i ; $\boldsymbol{\varepsilon}_i$ is a vector of residuals for person i , which are distributed $\text{MVN}(\mathbf{0}, \boldsymbol{\Theta})$; $\boldsymbol{\alpha}$ is a vector of growth factor means; $\mathbf{\Gamma}$ is a matrix of coefficients for the predicted effect of time-invariant covariates on the latent growth trajectory factors; \mathbf{X}_i is a matrix of time-invariant covariates for person i ; and $\boldsymbol{\zeta}_i$ is a vector of random effects for person i and is distributed $\text{MVN}(\mathbf{0}, \boldsymbol{\Psi})$. Figure 1 shows a path diagram of a hypothetical unconditional (i.e., no covariates) LC model with four time points and the values of $\mathbf{\Lambda}$ constrained to feature linear

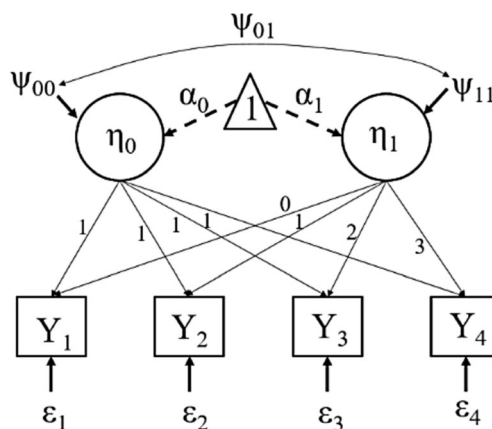


Fig. 1 Hypothetical latent growth model path diagram with four time points and linear growth. Numeric values indicate parameters that are constrained. Y indicates the observed variable values at each time point, η are growth factors, ψ are the disturbance co/variances of the latent growth parameters, α are the latent factor means, and ε are residuals (whose unlabeled variances would be θ)

growth. Software commonly used to fit growth models in the LC framework includes *Mplus*, LISREL, AMOS, the lavaan R package, and SAS Proc Calis.

A brief comparison of the ME and LC approaches

As has been well-explained (e.g., Curran, 2003; Singer & Willett, 2003; Skrondal & Rabe-Hesketh, 2004), the ME approach has much in common with the LC approach. Comparing Eq. 2 to Eq. 3,

- $\boldsymbol{\alpha}$ and $\mathbf{\Gamma}$ are related to $\boldsymbol{\gamma}$,
- $\boldsymbol{\Psi}$ is related to \mathbf{T} ,
- $\boldsymbol{\eta}_i$ is related to $\boldsymbol{\beta}_i$,
- $\boldsymbol{\Theta}$ is related to \mathbf{V} .

To highlight the similarity between the ME and LC equations, in the case in which each person is measured at the same time points (a.k.a. *time-structured data*), the matrix expression in Eq. 3 simplifies to

$$\begin{aligned} Y_{ij} &= \eta_{0i} + \eta_{1i} \text{Time}_j + \varepsilon_{ij}, \\ \eta_{0i} &= \alpha_{\beta 0} + \gamma_{01} X_{1i} + \dots + \gamma_{0k} X_{ki} + \zeta_{0i}, \\ \eta_{1i} &= \alpha_{\beta 1} + \gamma_{11} X_{1i} + \dots + \gamma_{1k} X_{ki} + \zeta_{1i}, \end{aligned} \tag{4}$$

where Y_{ij} is the response for the i th person at the j th time, η_{0i} is the latent intercept for the i th person, η_{1i} is the latent slope for the i th person, γ are the paths from time-invariant predictors, $\boldsymbol{\zeta}_i = (\zeta_{0i}, \zeta_{1i})^T$ is a vector of random effects for the i th person, Time_j is the j th time point, and ε_{ij} is the residual for the i th person at the j th time. Without much examination, it can be seen that Eq. 4 is essentially identical to Eq. 2, except for the different parameter labels arising from a different modeling framework. The equivalence

between the ME and LC frameworks, both mathematical and functional, however, is not perfect.

Differences in implementation

Prior to delving into specific details about the key differences between the implementations of the ME and LC frameworks, Table 1 presents a simplified summary of the arguments that will be presented for the remainder of this article, along with recommendations for which type of model is best suited to deal with particular scenarios. These recommendations are not definitive, because they assume that each situation is to be accommodated in isolation—in reality, it is common for more than one of these conditions to be present in a single analysis (e.g., multigroup analysis of nonlinear growth with a small sample size). In such cases, there may not be a perfect solution to adequately model all facets of the data in question, so researchers may need to weigh the pros and cons of each of the frameworks. These recommendations also apply

predominantly to continuous outcomes, and different recommendations may be desirable in the presence of discrete outcome variables.

To generally summarize Table 1, the ME approach tends to be most useful for straightforward models (e.g., a simple growth trajectory with one outcome variable), with complex data structures such as smaller samples, time-unstructured data, or multiple levels of nesting requiring more flexibility. Conversely, the LC approach is best suited to complex models with straightforward data structures, such as growth models embedded in larger models, assessments of global model fit, unconstrained time-varying covariates, and complex variance functions. Elaboration of these issues will be the focus of the remainder of this article. SAS and *Mplus* code from the subsequent examples is provided in an [online supplement](#).

Small sample sizes

Although a smaller sample size does not alter the mathematical equivalence of the models, it does affect the estimation

Table 1 Summary of recommendations for choosing between ME and LC frameworks in practice

Context	Summary and Recommendation
Small samples	The ME approach offers more options to accommodate growth models with few people. Bayesian methods are alluring in the LC framework, although researchers must be especially careful with the prior distributions.
Time-unstructured data	The ME approach minimally distinguishes between time-structured and time-unstructured data with respect to software implementation. The LC approach is viable with time-unstructured data in most cases, provided that specific conditions are met and additional steps are taken.
Nonlinear growth	If researchers know that growth will be nonlinear but are not concerned with or do not know the particular function, the latent-basis model in the LC approach offers an easy option that yields a straightforward interpretation. For specific forms of nonlinear growth, the ME approach should be preferred, because it can estimate the model directly and does not need to linearize the model prior to estimation, as is needed in structured LC models.
Multiple-group models	Either framework can accommodate multigroup models as they are commonly applied. If partial constraints are desired, the LC framework is far more flexible and can more easily accommodate situations in which a specific subset of parameters should be constrained or freely estimated.
Model fit	Global model fitting is a unique advantage of the LC framework. Researchers should be judicious about the indices they report and should keep in mind that growth models in an LC framework have an implied-mean structure (e.g., SRMR tends to perform poorly). Researchers should also be mindful that these criteria can perform undesirably with small samples, if uncorrected.
Residual structures	Differences between the frameworks are fairly minor. The LC framework is more flexible but requires all relations to be manually programmed or constrained. The ME framework simplifies the process with preprogrammed structures but is less flexible.
Three-level models	Three-level random-effects models are a natural extension of the ME framework and present no difficulty. A three-level random-effects model is quite difficult to fit in the LC framework, but if the third level is treated with a fixed-effects approach, then three levels are no more difficult to fit than a standard LC model.
Growth model embedded in a larger model	Embedding a growth model within a larger model or extending a growth model beyond linear growth with observed variables is straightforward to do in an LC framework. Options exist in the ME framework, as well, but to various degrees can be difficult to program.
Missing data	Most ME software will listwise delete observations with missing predictors, because the likelihood is conditional. LC software can accommodate missing predictors with FIML because the joint likelihood is typically used. Software for either framework can impute missing values.
Time-varying covariates	The ME framework constrains the effect to be constant across all time points and constrains the correlation between the covariate and the growth factors to be 0. These constraints can be lifted in the LC framework. Both frameworks allow for random effects of time-varying covariates.

options that researchers have at their disposal (McNeish, 2016b). Methodological research on clustered data in general (of which growth models form a subset) has shown that full maximum likelihood estimation leads to downwardly biased estimates of random-effect variances, fixed-effect standard errors, and growth factor mean standard errors when a study includes fewer than about 50 people (e.g., W. J. Browne & Draper, 2006; McNeish & Stapleton, 2016). This underestimates the variability in individual growth curves and can vastly inflate the operating Type I error rate of time-invariant predictors and growth factor means. However, full maximum likelihood is the primary frequentist method by which models in the LC framework are estimated. Robust estimation based on Huber–White so-called *sandwich estimators* also does not solve the issue pertaining to standard error estimation with small sample sizes (Maas & Hox, 2004).

In the ME framework, this issue has been well-studied and has largely been addressed via restricted maximum likelihood estimation (REML) to address the bias in the random-effect variances (Harville, 1977), and Kenward–Roger correction to address the bias in standard error estimates and inflated operating Type I error rates (Kenward & Roger, 1997, 2009). Recent studies have shown that these methods perform well with sample sizes as low as the single digits if reasonable model complexity is observed (Ferron, Bell, Hess, Rendina-Gobioff, & Hibbard, 2009). Unfortunately, as has been noted in McNeish (2016a), these methods do not have analogues in the LC framework,² and small-sample issues continue to be vastly underresearched in the LC as compared to the ME framework. McNeish (2016b) discusses in full detail the difficulty of deriving a REML-type estimator for the LC framework. Without delving into these details, the general issue is that the restricted likelihood function involves additional computations that are manipulations of the fixed-effect design matrix (\mathbf{X} in Eq. 1). This matrix does not exist in the LC framework, because its elements are allocated to the α , Λ , and \mathbf{X} matrices in Eq. 3. Additionally, as will be discussed shortly, the freedom within the LC approach allows the elements in the Λ matrix to be estimated (in the ME framework these elements are variables, not parameters, and cannot be estimated), which further complicates the derivation of a broadly applicable REML-type estimator in the LC framework.

Furthermore, different test statistics are utilized in each framework. In the SEM software programs used in the LC approach, parameters are tested with Z or χ^2 statistics. In ME software programs, the fixed-effect parameters are tested with t or F statistics, which are more appropriate with smaller sample sizes, because they do not assume infinite denominator degrees

of freedom, as is the case with the Z or χ^2 statistic (however, the appropriate degrees of freedom for such tests are widely debated; see, e.g., Schaalje, McBride, & Fellingham, 2002). Although it is known that the Z distribution is the limiting distribution of the t distribution, and similarly that χ^2 is the limiting distribution of the F distribution, using asymptotic test statistics with smaller samples can artificially inflate the operating Type I error rate, even if all parameters are estimated without bias (e.g., Schaalje et al., 2002).

Although small-sample-size issues may seem somewhat trivial, small-sample inference is an increased priority in growth models, because of the data collection difficulties in many contexts. That is, longitudinal data are difficult and expensive to collect, and meta-analytic reviews have shown that about one-third of growth model studies feature sample sizes in the double or single digits (e.g., Roberts & DelVecchio, 2000). McNeish (2016a) further went on to show that Bayesian estimation of small-sample LC growth models (a leading alternative to REML) does not necessarily alleviate these small-sample concerns, unless careful consideration is given to the prior distributions (see van de Schoot, Broere, Perryck, Zondervan-Zwijenburg, & van Loey, 2015, for specifics on setting the prior distributions for small-sample growth models), and also showed that small-sample methods developed in the ME context can yield estimates superior to those in the LC context with uninformative prior distributions.

Potthoff and Roy (1964) example A classic example of growth modeling with small samples comes from Potthoff and Roy (1964), who investigated the distance between the pituitary gland and the pteryomaxillary fissure in 27 children between the ages of 8 and 14 (these data are available in the *SAS 9.3 Users' Guide*, Example 58.8). Person 20 and Person 24 were removed from the data due to high Cook's D values for both the fixed effects and the covariance parameters, for a total sample size of 25 (for more information on diagnostics with these data, see Example 58.6 in the *SAS 9.3 User's Guide*). The growth in this distance is linear, and the sex of the child is included as a time-invariant predictor of both the intercept and the slope. Both the intercept and slope have random effects that do not covary, and the residual error structure is modeled as a homogeneous diagonal which was chosen on the basis of the Bayesian information criterion.

Table 2 compares the estimates from the ME framework, as estimated with REML and Kenward–Roger correction, and the LC framework, estimated with full maximum likelihood. As can be seen in Table 2, the LC intercept and slope variance estimates are noticeably lower than the ME estimates. Additionally, the substantive conclusions are different between the different methods. Because the ME random-effect variance estimates are larger and were subjected to small-sample corrections for the standard error estimates and the degrees of freedom for t tests, the p values are larger and hover

² Note that Cheung (2013) did devise a REML estimator for growth models in the LC framework; however, it can only be implemented in cases in which the LC and ME frameworks correspond exactly. Thus, it relies on model transformation rather than being a true REML estimator for LC models, broadly defined.

Table 2 Comparison of Potthoff and Roy (1964) data estimates in the ME framework, estimated with REML and a Kenward–Roger correction, and in the LC framework, estimated with full maximum likelihood

Parameter	Mixed Effects		Latent Curve	
	Est.	<i>p</i>	Est.	<i>p</i>
Intercept	21.209	–	21.209	–
Time	0.959	<.001	0.959	<.001
Sex	1.737	.045	1.737	.028
Sex × Time	0.416	.054	0.416	.036
<i>Variance Components</i>				
Var (Intercept)		3.572		3.289
Var (Time)		0.091		0.074
Var (Residual)		0.848		0.842

right around the .05 mark. On the other hand, the *p* values associated with the LC estimates are based on standard error estimates that are known to be downwardly biased while also using a questionably appropriate asymptotic sampling distribution, and therefore they are clearly under .05 with these data. Granted, the change in statistical significance is a nuance of these particular data and will not be universal. However, it serves to highlight that, despite identical coefficient estimates, the variance and standard errors are estimated differently between the frameworks with smaller samples.

Time-unstructured data

Longitudinal research designs are often thought of as multiwave, *time-structured* studies, meaning that all subjects provide a measure at the same time or at every wave. Nonetheless, practically speaking, data commonly deviate from the time-structured format (Singer & Willett, 2003; Sterba, 2014). When people in a longitudinal study are measured at different time points, the data are referred to as *time-unstructured*, which can be especially common when the time variable of interest is chronological age (especially in younger children, with whom it is important to record age to the month rather than the year). Because time is most often thought of as continuous, coarsening it by treating time as equivalent for all people when it is not can have an adverse impact on the parameter estimates (Aydin, Leite, & Algina, 2014; Singer & Willett, 2003). The degree to which this coarsening will impact estimates depends on how variably spaced the measurement occasions are (Coulombe, Selig, & Delaney, 2016; Singer & Willett, 2003).

Because the ME framework processes data in the “long,” univariate format and Time is an explicit predictor in the model, variably spaced measurement occasions and varying numbers of measurement occasions do not present a challenge (McCoach, Rambo, & Welsh, 2013; Sterba, 2014). That is,

both **X** and **Z** in Eq. 1 have an *i* subscript, indicating that each subject has its own values for time. For Eq. 2, which shows the ME specification in RB notation, Time has an *i* subscript, which allows each person to have unique time values in the data. In contrast, the LC framework processes data multivariately, meaning that each time point requires a unique column in the data (i.e., SEM software requires data to be in the “wide” format). That is, **Λ** in Eq. 3 does not have an *i* subscript, indicating that all subjects are expected to have the same time values. With time-unstructured data in which each person has potentially unique time points, structuring the data in the LC framework can become problematic.

Table 3 provides the data operationalized for the ME framework. *ID* is the subject ID, *Outcome* is the variable whose change is of interest, *Time* is the time in months when the outcome was collected, and *Time Group* is the *Time* rounded to the nearest year. Because the ME framework treats repeated measures univariately, each measure can simply be matched to the appropriate time. However, in the LC model, in which the model takes a multivariate specification, the process is not so simple, because (1) it is unclear to which values parameters should be constrained, and (2) it is unclear how to organize the columns within the data.

To prepare the data for an LC model, researchers must make some decisions. The first and simplest option, depicted in Table 4, is to collapse some categories to make the data coarser—that is, to use *Time Group* instead of *Time*. This would then remove the confusion regarding how to set the slope loading constraints and how to set up the data, because each person has observed data at the same time points; however, the downside is that researchers may lose potentially useful information, and large shifts could bias many of the parameter estimates (Blozis & Cho, 2008; Singer & Willett, 2003).

Table 3 Representation of repeated measures organized in a univariate, “long” format, where each person occupies multiple rows and the repeated measures are housed in a single column

ID	Outcome	Time	Time Group
1	2	0	0
1	3	2.6	3
1	3	6.1	6
1	4	8.9	9
2	3	0.1	0
2	4	6.2	6
2	5	9.2	9
⋮	⋮	⋮	⋮
<i>N</i>	2	0.2	0
<i>N</i>	3	2.6	3
<i>N</i>	4	5.9	6
<i>N</i>	4	9.1	9

Table 5 illustrates a second option, creating an outcome variable for each of the possible time points and treating differing time points as missing data (Wu, West, & Taylor, 2009). This option typically relies on some overlap being present in the time points at which measurements occurred, to avoid convergence issues associated with many missing-data patterns and large amounts of missing data. A third option would be to treat subjects with the same measurement occasions as a group and to conduct a multigroup analysis (Preacher, Wichman, MacCallum, & Briggs, 2008). A drawback of this approach is that it remains feasible only when there are few possible combinations of time points—if each person has vastly different time points, it becomes unruly to have several groups or several sparse outcome variables.

The last option would be to use definition variables, which entails constraining the parameter values to an observed variable value from the data (Hamagami, 1997; Mehta & Neale, 2005; Mehta & West, 2000; Serang, Grimm, & McArdle, 2016; Sterba, 2014). Generally, this involves creating additional variables in the data that contain the time at which each measurement occurred. Through definition variables, the loadings can be constrained to the value in these variables, allowing each person’s slope loadings to represent the unique times at which their data were collected. In matrix notation, Λ would equal

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ t_{0i} & t_{1i} & t_{2i} & t_{3i} \end{bmatrix}^T,$$

where t are the person-specific time points at which the data were collected (person-specific due to the i subscript). This can be especially useful and straightforward if each person has the same number of measurements but the measurements occur at different times. To provide some background on software implementation, in *Mplus* one can specify person-specific values for the loadings from the slope factor of the observed variables by using the TSCORES command with an AT option (see Example 6.12 in the *Mplus User’s Guide*; Muthén & Muthén, 2012).

Burchinal and Appelbaum (1991) example As an example, consider the data from Burchinal and Appelbaum (1991), who were interested in measuring the number of speech errors

Table 4 Representation of repeated measures, organized in a multivariate “wide” format, using a coarsened version of *Time Group* such that each repeated measure occupies its own column

ID	Tgroup0	Tgroup3	Tgroup6	Tgroup9
1	2	3	3	4
2	3	NA	4	5
⋮	⋮	⋮	⋮	⋮
N	2	3	4	4

made by 43 children between 2 and 8 years old. As may be expected, as children became older, they tended to make fewer and fewer speech errors. However, because the time variable in these data was chronological age in months, the data are time-unstructured; every person essentially has unique measurement occasions, and the intervals between measurement occasions are unique for each person. As an added difficulty, 26% of the sample have four measurement occasions, 44% have five measurement occasions, and 30% have six measurement occasions. Figure 2 shows a trend plot for these data over time with a superimposed mean curve, which seems to demonstrate that a quadratic term may be needed. The data also feature a variable for the intelligibility of the child’s speech, which will be used as a covariate of the intercept and of the linear and quadratic slopes.

In the ME framework, no special approach needs to be taken to accommodate the time-unstructured nature of the data, and the standard approach is unfazed.³ In the LC framework, three major problems arise immediately—each person has unique measurements, there is no overlap in the time points between people, and the numbers of measurements for different people are not equal. The definition-variable approach was not able to converge, likely due to the complexity introduced by the number of measurement occasions being different for each person and the small sample size. The multiple-group modeling strategy is also untenable with such widely varying measurement occasions, which only leaves collapsing age as a possibly strategy. The original data were coded to the month, so we instead coarsened the time variable to the nearest year. Table 6 shows the estimates of the ME and coarsened LC approaches. The results are in the same general vicinity, but the parameter estimates are noticeably different. In particular, the LC variance estimates are much smaller, presumably because some of the variability has been truncated because of the time coarsening.

Nonlinear growth

Although linear growth trajectories are commonly used in empirical studies, many growth processes are inherently nonlinear and call for different types of models in order to adequately model growth over time (Cudeck & du Toit, 2002; Grimm & Ram, 2009; Grimm, Ram, & Hamagami, 2011; Preacher &

³ With the potential caveat that the sample size is rather small, as we discussed in the previous section. To render the estimates more comparable, the ME estimates were obtained with full maximum likelihood and without any small-sample corrections. Readers may also note that the growth trajectory is more likely exponential than quadratic. As will be discussed shortly, there are also differences between the frameworks with regard to fitting nonlinear trajectories. Choosing a model that differs from the given recommendations in multiple respects could confound the point we are hoping to make about time-unstructured data. Therefore, we chose to approximate the nonlinearity with a quadratic term, to preserve the comparability of the models in all respects other than the time-unstructured data.

Table 5 Representation of repeated measures, organized in a multivariate format using the original time metric *Time*. Notice the large amount of missing data, indicated by NA

ID	T0	T0.10	T0.20	T2.35	T2.60	T5.90	T6.10	T6.20	T8.90	T9.10	T9.20
1	2	NA	NA	3	NA	NA	3	NA	4	NA	NA
2	NA	3	NA	NA	NA	NA	NA	4	NA	5	NA
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
<i>N</i>	NA	NA	2	NA	NA	4	NA	NA	NA	NA	4

Hancock, 2015; Sterba, 2014). The estimation and specification of nonlinear models can be vastly different between the ME and LC frameworks, with many different types of models being uniquely estimable in only one framework. We will not be able to fully cover all of the nuances of nonlinear models in a single section, since many full-length articles have been dedicated solely to this topic (e.g., Blozis & Harring, 2016b); however, we will attempt to highlight the most salient of the differences that are likely to arise in empirical research.

A notable difference between the ME and LC frameworks is their ability to handle models with nonlinear parameters. To explain, nonlinear models can be nonlinear in their *variables* or their *parameters*. A model with a second-order polynomial term [e.g., $E(Y) = \beta_0 + \beta_1\text{Time} + \beta_2\text{Time}^2$] is linear in its parameters, since the β coefficients only appear after addition signs, but nonlinear in its variables, because Time enters the model with a quadratic term. Conversely, an exponential growth model [e.g., $E(Y) = \alpha_F + (\alpha_0 - \alpha_F) \exp(-\alpha_R\text{Time})$] is nonlinear in its parameters, because the rate parameter (e.g., $-\alpha_R$) appears in the exponential expression.

ME software is able to accommodate either of these specifications without changing the interpretation of the model

parameters (e.g., SAS Proc Mixed for models nonlinear in their variables, SAS Proc NLMIXED for models nonlinear in their parameters; Blozis & Harring, 2016b). SEM software can accommodate models that are nonlinear in their variables but cannot directly accommodate models that are nonlinear in their parameters (Blozis & Harring, 2016b). A common method to fit models that are nonlinear in their parameters using SEM software is through a structured latent-curve model (SLCM; Blozis, 2004; M. W. Browne, 1993), which linearizes the nonlinear portions of the model using Taylor series expansions (for details on this procedure, see Blozis & Harring, 2016a). Although many consider the SLCM to be the LC equivalent of nonlinear ME models, as detailed by Blozis and Harring (2016b), the linearization process changes the interpretation of the model and of the random effects (the interpretation of the factor means takes a population-averaged rather than a subject-specific interpretation, which are not equivalent in nonlinear models; Fitzmaurice, Laird, & Ware, 2004; Zeger et al., 1988). In SLCMs, subject-specific curves are not required to follow the same functional form as the mean curve—the only restriction is that the sum of the subject-specific curves must equal the mean trajectory. In nonlinear ME models, every subject-specific curve follows the same functional form as the

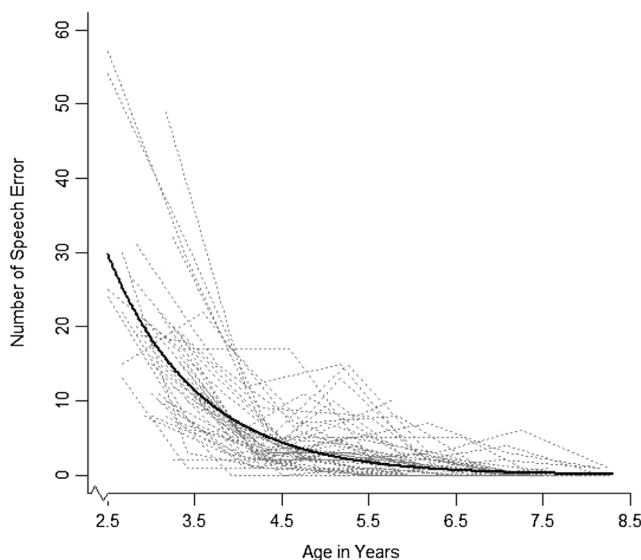


Fig. 2 Plot of speech error learning curves for all 43 children in the Burchinal and Appelbaum (1991) data, with a mean curve superimposed in black

Table 6 Comparison of model estimates for Burchinal and Appelbaum (1991) data for ME and a LC that coarsens time

Parameter	Mixed Effects		Coarsened Latent Curve	
	Est.	<i>p</i>	Est.	<i>p</i>
Intercept	90.32	–	105.78	–
Time	– 26.81	<.01	– 30.50	<.01
Time ²	2.00	<.01	2.20	<.01
Intelligibility	– 7.21	<.01	– 12.39	<.01
Time × Intelligibility	2.04	.02	3.60	.02
Time ² × Intelligibility	– 0.15	.07	– 0.26	.04
<i>Variance Components</i>				
Var (Intercept)		112.82		91.70
Var (Time)		2.16		1.59
Cov (Intercept, Time)		– 15.93		– 12.48
Var (Residual)		20.48		21.73

mean curve—thus, differences between people stem only from the subject-specific random-effect estimates, not from potentially different trajectories altogether.

An advantage of modeling nonlinear growth in the LC framework is the so-called *latent-basis model* (McArdle, 1986; Meredith & Tisak, 1990) or *free curve* (Wood & Jackson, 2013). In the LC framework, time is a parameter of the model (the values of the loadings from the slope factor) rather than a variable in the data. As a result, one can estimate the values of these slope loadings rather than fixing them to specific values. To give the slope factor scale, two loadings must be constrained, and a common choice is to constrain the loading of the first time point to 0 and the loading of the last time point to 1, while estimating the loadings of all intermediate time points. When the growth trajectory is specified a priori, for an unconditional linear growth model with time-structured data,

$$\mathbf{A} = \mathbf{X}_i = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix}^T.$$

However, in a latent-basis model,

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & \lambda_{21} & \lambda_{31} & 1 \end{bmatrix}^T,$$

where λ are estimated parameters; this is not feasible to do in linear ME models, because the elements of \mathbf{X}_i are variables in the data and therefore not estimable. Grimm, Ram, and Estabrook (2016) have noted, however, that the latent-basis model can be coerced to fit in the ME framework as a nonlinear mixed-effect model. They provide code for fitting a latent-basis model in SAS Proc NLMIXED on their page 243, through the clever use of data steps. This code is restricted to having the residual variance constrained across time points, however.

With this approach, the slope factor mean is an estimate of the amount of total growth that occurs over the entire observation window, provided that the first loading is constrained to 0 and the last loading is constrained to 1. In this case, the estimated slope loadings correspond to the percentages of the total growth that has occurred up to and including that specific time point (e.g., an estimated slope loading of .65 at Time 3 means that the outcome at Time 3 is 65% of the total growth. It does not mean that 65% of the growth process has occurred at Time 3, however, unless the curve is monotonic).⁴ This approach can be especially advantageous if researchers know that growth is nonlinear but are not sure or not interested

in testing specific types of nonlinear growth trajectories (e.g., exponential or Gompertz; Raykov & Marcoulides, 2012). The model is also linear in its parameters, so the estimation and interpretation of the parameters is straightforward (Wu & Lang, 2016). Figure 3 shows a hypothetical path diagram of an unconditional latent-basis model with four time points. The parameters that make the latent-basis model different from a standard LC model are featured in bold underlined text. It is important to understand, however, that the functional form of the curve is not specified a priori, but is determined from the data. The next subsection demonstrates the differences between nonlinear ME models, SLCMs, and latent-basis models by means of an empirical example.

ECLS-K example We used a reduced version of the ECLS-K dataset (Tourangeau, Nord, Lê, Sorongon, & Najarian, 2009), which has recently been used with nonlinear growth models in a study by Cameron, Grimm, Steele, Castro-Schilo, and Grissmer (2015). ECLS-K follows students from the fall semester of kindergarten to the spring semester of Grade 8. The data are expansive in terms of both the number of observations and the number of variables. In this example analysis, we model the vertically scaled reading scores taken at five time points (kindergarten fall, Grade 1 spring, Grade 3 spring, Grade 5 spring, and Grade 8 spring) for the 2,145 students with complete data. For simplicity, we will model the growth unconditionally, such that there are no covariates in the model. Three separate models will be fit to the data: a nonlinear Michaelis–Menten ME model, a structured latent Michaelis–Menten curve model (in the LC framework), and a latent-basis model (also in the LC framework). Because these models may be unfamiliar to some readers, we will discuss the basic details of each.

The Michaelis–Menten model has its origins in biochemistry and in modeling rates of chemical reactions, although it has been found to be useful in behavioral science applications due to its highly interpretable parameterization (Cudeck & Haring, 2007; Haring, Kohli, Silverman, & Speece, 2012). The model has three parameters: an intercept that estimates people's outcome variable when Time = 0, an asymptote parameter that estimates the upper limit for the outcome variable as Time $\rightarrow \infty$, and a midpoint parameter that estimates the value on the Time scale at which people are halfway between their intercept value and their upper asymptote. In a growth model context, each of these parameters typically has a random effect, which allows the estimates to be subject-specific. The unconditional model can be written as

$$\begin{aligned} y_{it} &= \beta_{0i} + \frac{(\beta_{Ai} - \beta_{0i})\text{Time}}{\beta_{Mi} + \text{Time}} + r_{it}, \\ \beta_{0i} &= \gamma_0 + u_{0i}, \\ \beta_{Ai} &= \gamma_A + u_{Ai}, \\ \beta_{Mi} &= \gamma_M + u_{Mi}, \end{aligned} \quad (5)$$

⁴ A notable limitation of the latent-basis model is that it makes a *proportionality assumption* (Wu & Lang, 2016). This means that, although the total amount of growth can vary for each individual in the data, the proportions of growth at each time point are assumed to be equal. That is, if the slope loading at Time 3 is estimated to be .65, then the model assumes that 65% of the total outcome is achieved at Time 3 for all subjects in the data. For more detail about this assumption and its ramifications, readers are referred to Wu and Lang (2016).

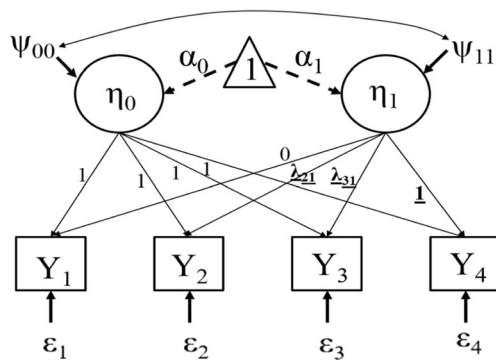


Fig. 3 Hypothetical latent-basis model path diagram with four time points. Numeric values indicate parameters that are constrained. Y indicates the observed variable values at each time point, η are growth factors, ψ are the disturbance co/variances of the latent growth parameters, α are the latent factor means, λ are the estimated factor loadings, and ε are residuals (whose unlabeled variances would be θ)

where the parameters with a “0” subscript relate to the intercept, an “A” subscript to the upper asymptote, and an “M” subscript to the midpoint.

Note in Eq. 5 that the Michaelis–Menten model is nonlinear in its parameters, because the midpoint parameter appears in the denominator with other parameters appearing in the numerator. Therefore, the exact model cannot be fit with SEM software, although a variation of it can—namely, the SLCM. Although the mathematical details are beyond the scope of this article, the general idea is to place constraints (which are based on partial derivatives of the mean function of Eq. 5) on particular parameters in order to linearize an inherently nonlinear model into a linear form (see the online [supplementary materials](#) for calculations of how the constraints are derived specific to the Michaelis–Menten model, or see Blozis & Harring, 2016a, or M. W. Browne, 1993, for full general details). As we mentioned previously, this will fit a model that is similar to Eq. 5 but not identical, because the linearization process changes aspects of the model.

The latent-basis model is similar to that in Fig. 2, except that there are now five time points. The loadings of the first and last time points will be constrained to 0 and 1, respectively. No special considerations are necessary, because the data have a large sample size and are time-structured (i.e., all people are measured at the same time points).

Table 7 shows the parameter estimates for the three different types of nonlinear models. First, compare the Michaelis–Menten estimates between the nonlinear ME model and the SLCM specifications. Although it is common to consider these models interchangeable, the variance estimates are noticeably different even though the quality of the data is rather high (e.g., large sample, normally distributed outcomes based on IRT scores, no missing data). The difference stems from the ME model retaining the subject-specific interpretation of the parameters, whereas the linearization process in the SLCM has a marginal interpretation (Blozis & Harring, 2016b). As we noted

previously, the change in the random effects and subject-specific curves is stark, which can be seen by the vastly different variance–covariance parameters for the three random effects.

To demonstrate an advantage of the LC framework, consider the latent-basis model in the rightmost column of Table 7. Although it is not as theoretically grounded as the Michaelis–Menten model, the latent-basis model still yields interesting information. Immediately, it can be seen that the average Reading score for fall of kindergarten is about 37, and the total growth from kindergarten to spring of Grade 8 is about 137 points. The slope loadings are informative for determining the percentage of the total growth at each point—here it can be seen that most growth occurs up to the spring of Grade 3, and then growth begins to taper off. The latent basis model is not designed to make in-depth conclusions about certain nonlinear trajectories (e.g., in this model, it tells nothing of the upper asymptote or the midpoint), but the model itself is quite simple to program with SEM software, can be a useful option when researchers know that growth is nonlinear but do not know the specific shape, and is much simpler computationally (the ME nonlinear model in Table 7 took approximately 90 min to converge, and the SLCM requires calculus to compute the partial derivatives for the constraints on the loadings). With representative samples, the latent-basis model can often approximate the appropriate confirmatory trajectory across the observation window. For instance, Fig. 4 compares the expected values from the example analysis of the Michaelis–Menten ME model and the latent-basis model with a smooth; the Michaelis–Menten SLCM is omitted because the mean trajectory is quite close to that of the ME model (though the variances are not). Both models yield almost identical trajectories, despite the relative ease of specifying and estimating the latent-basis model.

Model fit assessment

Although the statistical models in the ME and LC frameworks can be shown to be mathematically equivalent in most cases and yield essentially equivalent estimates of growth trajectories under ideal settings, model fit assessment procedures are noticeably different between the methods. There are two main approaches when assessing model fit: (1) comparing competing models to choose the most parsimonious one, and (2) comparing a model’s implied covariance to the sample covariance, to broadly assess the similarity of the model-implied and observed covariance and mean structures.⁵

Both the ME and the LC frameworks provide indices and tests to directly compare nested and nonnested models—likelihood ratio tests for nested models, and the Akaike and

⁵ Alternative frameworks for assessing model fit that we do not address here are cross-validation and, under Bayesian estimation, posterior predictive checks.

Table 7 Comparison of different frameworks for nonlinear growth

Parameter	Michaelis–Menten		Latent Basis	
	Mixed Effects Est.	Structured Latent Curve Est.	Latent Curve Parameter	Est.
Intercept	36.38	36.35	Intercept Mean	36.77
Asymptote	257.11	258.32	Slope Mean	136.90
Midpoint	4.89	4.87	Loading Kind. Fall	0.00
			Loading Grade 1 Spr.	0.33
			Loading Grade 3 Spr.	0.71
			Loading Grade 5 Spr.	0.87
			Loading Grade 8 Spr.	1.00
<i>Covariance Parameters</i>				
Var (Intercept)	82.98	100.63	Var (Intercept)	100.63
Var (Asymptote)	512.17	285.75	Var (Slope)	508.02
Var (Mid-Point)	1.36	2.87	Corr (Int., Slope)	0.02
Corr (Int., Asy.)	– 0.43	– 0.46	Residual Kind. Fall	0.00
Corr (Int. Mid.)	– 0.85	– 0.62	Residual Grade 1 Spr.	233.30
Corr (Asy., Mid.)	0.02	0.07	Residual Grade 3 Spr.	169.65
Residual Kind. Fall	32.86	0.00	Residual Grade 5 Spr.	66.89
Residual Grade 1 Spr.	272.23	256.83	Residual Grade 8 Spr.	231.16
Residual Grade 3 Spr.	154.02	153.97		
Residual Grade 5 Spr.	88.73	76.50		
Residual Grade 8 Spr.	198.94	233.65		

The ME model was estimated with Gaussian quadrature with ten quadrature points. The residual error structure was specified to be a heterogeneous diagonal in all three models. Significance is not reported because the sample size was large. Bold values were constrained and thus not estimated. The residual variance in fall of kindergarten was constrained to zero in the models fit in the LC framework in order to avoid a Heywood case, because the estimate was slightly negative

Bayesian information criteria for nonnested models. If the same estimation method is used, these quantities will be identical between the frameworks. However, assessing model fit by comparing the sample covariance and model-implied covariance is much different for the two frameworks. The LC framework has a host of indices and statistical tests that can help determine how well the model-implied covariance and mean structure reproduce the sample covariance and sample means (Coffman and Millsap, 2006; Chou et al., 1998; Liu,

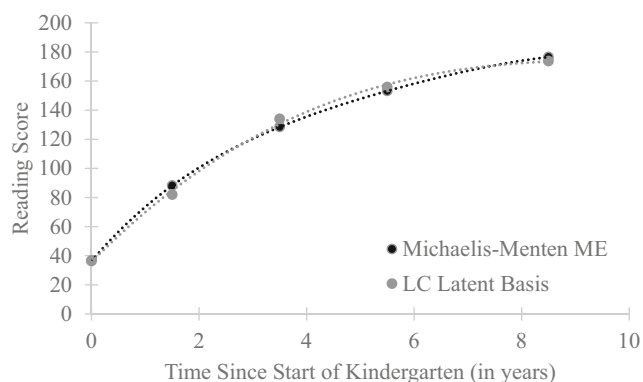


Fig. 4 Comparison of mean trajectories for the Michaelis–Menten ME (black) and LC latent-basis (gray) models

Rovine, & Molenaar, 2012; Wu et al., 2009). ME models, on the other hand, can use visual tools to determine how well the model-implied covariance reflects the sample covariance (Hedeker & Gibbons, 2006), or can use variance-explained measures (Vonesh, Chinchilli, & Pu, 1996; Xu, 2003) or approximate R^2 extensions (Edwards, Muller, Wolfinger, Qaqish, & Schabenberger, 2008; Johnson, 2014; Nakagawa & Schielzeth, 2013), but these methods are less commonly reported and less mainstream than the indices in the LC framework. The flexibility of the ME approach often adds to this difficulty. For example, formally comparing model-implied matrices to observed matrices can be cumbersome for time-unstructured data, and ME R^2 methods can be difficult to implement when random slopes are included, which is common for growth models (e.g., Snijders & Bosker, 1994).

Although the fit indices in the general SEM framework have been well-studied, because LC growth models require both the covariance structure and the mean structure, the protocol for assessing fit differs from conventional SEM fit assessment strategies and guidelines—namely because most SEM fit assessment recommendations are based on models that only model the covariance (e.g., CFA models). In LC growth models, the commonly reported SRMR index performs quite poorly and rarely identifies misfit in the mean

structure portion of the model (Leite & Stapleton, 2011; Wu & West, 2010). This can be concerning because a major part of any growth model analysis is to report how people are generally changing over time. Additionally, with such relative fit indices as the comparative fit index (CFI) or the Tucker–Lewis index (TLI), the baseline model that is used as a source of comparison is inappropriate. Specifically, Widaman and Thompson (2003) noted that, in the CFI and TLI formulas, the baseline model must be nested within the analysis model. However, in LC growth models, the standard baseline model in the software is often the independence model, in which each observed variable has an estimated variance and the mean structure (if present) is saturated (Bentler & Bonett, 1980). This independence model is *not* nested within LC growth models, though, making the CFI and TLI models reported from mainstream SEM software inappropriate in the case of growth models. Instead, Widaman and Thompson (2003) recommended the linear latent-growth model with the intercept variance constrained to zero as an alternative baseline model that is nested in the more expansive LC model.

As we noted in an earlier section, small samples are quite common in data suitable for growth modeling in general. Small samples not only affect the estimation of parameters in LC models, but also the calculation of data-model fit statistics and indices (Bentler & Yuan, 1999; Herzog & Boomsma, 2009; Kenny & McCoach, 2003; Nevitt & Hancock, 2004). Without belaboring the specific details, the general issue is that, with smaller samples, the maximum likelihood test statistic (commonly referred to as T_{ML}) does not follow the appropriate chi-square distribution with the associated degrees of freedom. As a result, this test statistic tends to have vastly deflated p values, which leads to the over-rejection of models even if the models are perfectly specified, until the sample size approaches about 100 people (Kenny & McCoach, 2003). Furthermore, because most LC fit indices are based on some manipulation of the chi-square test statistic (e.g., RMSEA, CFI, TLI), these indices are also affected and yield unnecessarily unfavorable assessments of fit. Fortunately, three different heuristic small-sample corrections have been proposed, by Bartlett (1950), Swain (1975), and Yuan (2005), that adjust the chi-square test statistic so that it more closely approximates the appropriate chi-square distribution. Simulation studies have shown that these corrections work well in the general context of SEMs (Fouladi, 2000; Herzog & Boomsma, 2009; Nevitt & Hancock, 2004). A caveat of these small-sample corrections, however, is that they require a complete dataset (e.g., either the dataset must have no missing observations or the data must have been rectangularized with a suitable missing-data method, such as multiple imputation). For a discussion of issues related to these small-sample corrections with missing data in growth models, readers are referred to McNeish and Harring (2017).

Potthoff and Roy (1964) example To demonstrate, we revisited the previous Potthoff and Roy (1964) example, which has 25 participants. With a sample of this size, it is unlikely that the maximum likelihood chi-square test will be overpowered, and it is therefore the best measure with which to assess fit. Because the model has few degrees of freedom and a small sample size, the root-mean square error of approximation is not desirable to report, either (Kenny, Kaniskan, & McCoach, 2015). Without any correction, the maximum likelihood test shows that the hypothesis that the model-implied mean and covariance structures equal the observed mean vector and covariance matrix should be rejected, and we would conclude that the model does not fit well, $\chi^2(11) = 23.152, p = .017$. However, with a smaller sample this test tends to over-reject well-fitting models. Both the Bartlett- and Yuan-corrected statistics that have been recommended for LC models with small samples and complete data show that the null hypothesis should not be rejected and that the model provides a reasonable fit to the data, $\chi^2_{Bart}(11) = 19.133, p = .059$; $\chi^2_{Yuan}(11) = 19.454, p = .053$. Equivalent measures of global fit in the ME framework are not well-developed, and assessing fit in the ME framework is largely relegated to significance tests of specific parameters or model comparisons (Wu et al., 2009).

Putting it all together

In previous sections we showed how the LC and ME framework differed, one facet at a time. This approach was taken for its didactic simplicity, and more advanced readers may note that the less beneficial framework may still be viable in these circumstances, with the aid of some creative programming. However, as we mentioned earlier, multiple facets on which the LC and ME frameworks differ can also be present simultaneously, which can make the selection of the appropriate framework even more crucial, because the programming tricks required to combat multiple differences simultaneously can be unruly.

Consider the data analyzed by Prosser, Rasbash, and Goldstein (1991), appearing in Rabe-Hesketh and Skrondal (2012), of children's weights (in kilograms) over the first few years of their lives. This particular dataset has a small sample ($n = 68$), nonlinear growth (babies gain weight rapidly at first, but weight gains begin to taper off), unstructured time (the Time variable is chronological age, so each child has a unique value), and a multiple-group component (males and females typically are different sizes at this stage of development). We fit the model in the ME framework using SAS Proc Mixed with linear and quadratic terms for growth, random but uncorrelated intercepts and linear slopes, and a homogeneous diagonal error structure. The model was estimated with REML and a Kenward–Roger correction. Because one of

the questions related to these data concerns differences between the sexes, we fit a series of multigroup models, with (a) all parameters between boys and girls constrained to be equal, (b) only error variances free between the two groups, (c) error and random-effect variances free between the two groups, and (d) all parameters free between the two groups. Through restricted likelihood ratio tests (for tests involving fixed effects in the growth trajectory, we used full likelihood ratio tests), the fully unconstrained model fit best, and the parameter estimates are shown in Table 8. In this table, it can be seen that boys weigh more at birth, their weight grows more quickly, and there is much more variation in the linear growth rate for boys than for girls, but there is much more variation in weight at birth for girls than for boys.

Fitting the same model in the LC framework using *Mplus* is much more difficult. First, each of the 68 people in the data are measured at unique time points, so we coarsened the time variable by rounding to the nearest 3 months. Even after this coarsening, there were 53 different response time patterns (out of 68 people). Because age only ranged from shortly after birth to about 2.5 years, coarsening the data further (e.g., to the nearest year) would be unreasonably broad. Recall that SEM software treats different response time patterns as missing data, so the estimation failed to converge because the amount of missing data was quite large relative to the sample (about 80% of the data matrix).

This model is mathematically equivalent in either the ME or LC framework, and theoretically it could be fit in either framework. However, the ME framework is far more straightforward, because one need not worry about unstructured time, and the smaller sample size is less problematic. To connect this example to our earlier discussion, the example has a complex data structure and a straightforward model, so the ME possesses the clear advantage. Though this is not discussed in a dedicated section, the multiple-group component can be

seamlessly applied in the ME framework, as well. As such, when estimating the model with SAS Proc Mixed, the model is essentially the same as any other standard growth model fit using this software, and the model converged without any issues. Using the LC framework is far more troublesome here, because the small sample and unstructured time were highly problematic—because unstructured time essentially turns into a large missing-data problem, the model could not even produce estimates. Switching to a Bayesian framework and treating the missing data as parameters in the model (i.e., the fully Bayesian approach to missing data) did not help, because the Markov chain Monte Carlo chains could not converge. Attempts to further simplify the model, such as removing the quadratic term and removing one of the random effects, were similarly unsuccessful. Thus, even though the models are mathematically identical in both frameworks, the ME framework is far more advantageous for fitting this model because (1) the model can easily produce results and (2) REML estimation and a Kenward–Roger correction are more trustworthy for these data, anyway. To summarize this example and the thesis of this article succinctly, mathematical similarity does not mean that models are functionally equivalent.

Discussion

Although we have presented several differences in the implementations of these modeling frameworks, it is important to note that the aforementioned differences are not the only dimensions along which the ME and LC frameworks are distinguishable. To list other, less pervasive differences:

- The ME approach is more easily extensible to data with higher levels of clustering (e.g., growth in students' scores over time, since students are clustered within schools). The nested random effects needed for this type of model are difficult to implement in the LC approach as a three-level hierarchy (Curran, Obeidat, & Losardo, 2010). However, if no research questions exist at the third level of hierarchy, and the interest is simply in controlling for this level and ensuring that the standard error estimates accurately reflect the data structure, a fixed-effects approach at the third-level is quite easy to implement in an LC framework (see, e.g., McNeish & Wentzel, 2017).
- Although this can be software-specific, most ME software programs will only perform full information maximum likelihood (FIML) estimation if any missing values are contained to the outcome variables, and listwise deletion will be used if the predictor variables have missing values (Allison, 2012; Enders, 2010). Both approaches can utilize multiple imputation to deal with missing data; however, LC models have no issues when using FIML with missing predictors, which can help researchers avoid the

Table 8 Final estimates for the Prosser, Rasbash, and Goldstein (1991) data and deviance for four multiple-group models in the ME framework

Predictor	Boys Estimate	Girls Estimate
Intercept	3.68	3.35
Age	8.02	7.29
Age ²	− 1.74	− 1.54
Var (Int)	0.29	0.61
Var (Age)	0.47	0.09
Var (Residual)	0.39	0.24
Model	Deviance	No. Parameters
All unconstrained	494.2	12
All variances free	506.3	9
Error variances free	513.9	7
All constrained	517.0	6

somewhat complicated problem of imputing longitudinal data (Newman, 2003).

- LC software makes it quite easy to extend the growth model to a fully structural model (e.g., for distal outcomes, categorical latent variables for identifying latent classes, second-order growth models for latent outcomes, latent change score models, and outcomes at different levels; Curran et al., 2010; Hancock, Kuo, & Lawrence, 2001; McArdle, 2009; Sterba, 2014). These types of analyses can be conducted in some ME software, as well (notably, the GLLAMM programs featured in Stata). Put another way, ME software typically considers the growth model to be the primary modeling interest, whereas general SEM software for fitting LC models can naturally extend the models so that the growth model is only one portion of a larger overall model.
- The choice of the residual covariance structures in the ME framework is often limited by preprogrammed software options. For example, SPSS Mixed has 17 possible options (Peugh & Enders, 2005), SAS Proc Mixed has 37 options (SAS Institute, 2008), and Stata xtmixed has eight options (StataCorp, 2013). In the LC framework, researchers have the flexibility to specify any structure they wish. The caveat for this added flexibility is that the LC framework requires the structure to be programmed manually. This makes some common structures, such as autoregressive(1) or compound symmetry, a little tricky, whereas they are preprogrammed in ME software. However, a structure such as

$$\begin{bmatrix} \sigma_1^2 & \sigma_{12} & 0 & 0 \\ \sigma_{21} & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix}$$

would be quite simple to specify in the LC framework, whereas ME software has no corresponding preprogrammed structure.

- The ME framework in an extension of regression models, and methods for inspecting adherence to the model assumptions are much more developed and have wider availability in ME software applications. Such an endeavor is not widely discussed or reported in the application of growth models in the LC framework, and software support for such assumption checks tends to be lacking.
- Interactions with time tend to be easier to specify in the ME framework, because Time is featured as a variable in the model. In the LC framework, Time is not a variable because time is implied via constraints or estimates of the slope loadings. There are few differences between the

frameworks if the LC slope loadings are constrained to a linear trajectory, but the interpretation of interactions with Time can be less straightforward in the LC framework with other growth trajectories (e.g., Li, Duncan, & Acock, 2000).

Though it is true that growth models in the LC and ME frameworks are almost always mathematical equivalent, mathematical equivalence does not necessarily imply that the choice of a modeling framework is arbitrary and merely a matter of personal preference. Unless the data are pristine, in that the sample is large, growth is linear, there are no missing data, and measurements are completely balanced, advantages and disadvantages will be associated with selecting one framework or another. To break down the comparison into its simplest form, the ME framework tends to be better suited for complex data structures and straightforward models, whereas the LC framework tends to be better suited for straightforward data structures and complex models. In general, the ME framework and software are less flexible for including growth as part of a broader model, but they provide a greater wealth of specialized options for the types of models that can be accommodated—namely, when the growth process is the primary modeling interest. On the other hand, researchers have paid more attention to extending the LC framework to the general latent-variable modeling framework, resulting in more software options to incorporate a dizzying array of modeling extensions, although LC tends to be less attuned to circumstances that require more specialized methods (e.g., small-sample correction, or estimators for models nonlinear in their parameters). Fitting growth models to data can be a complex task in and of itself. Either the ME or LC framework is capable of addressing a wide variety of models for growth; however, as we hopefully have demonstrated here, researchers may be able to facilitate the modeling process to a certain degree simply by choosing to equivalently model their data within a different framework whose implementation more closely aligns with their research goals.

References

- Allison, P. D. (2012). *Handling missing data by maximum likelihood*. Statistics and Data Analysis Keynote Presentation at the SAS Global Forum, Orlando, FL.
- Aydin, B., Leite, W. L., & Algina, J. (2014). The consequences of ignoring variability in measurement occasions within data collection waves in latent growth models. *Multivariate Behavioral Research*, 49, 149–160. <https://doi.org/10.1080/00273171.2014.887901>
- Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Statistical Psychology*, 3, 77–85.
- Bauer, D. J. (2003). Estimating multilevel linear models as structural equation models. *Journal of Educational and Behavioral Statistics*, 28, 135–167.

- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588–606. <https://doi.org/10.1037/0033-2909.88.3.588>
- Bentler, P. M., & Yuan, K. H. (1999). Structural equation modeling with small samples: Test statistics. *Multivariate Behavioral Research*, *34*, 181–197.
- Blozis, S. A. (2004). Structured latent curve models for the study of change in multivariate repeated measures. *Psychological Methods*, *9*, 334–353.
- Blozis, S. A., & Cho, Y. I. (2008). Coding and centering of time in latent curve models in the presence of interindividual time heterogeneity. *Structural Equation Modeling*, *15*, 413–433.
- Blozis, S. A., & Harring, J. R. (2016a). On the Estimation of nonlinear mixed-effects models and latent curve models for longitudinal data. *Structural Equation Modeling*, *23*, 904–920.
- Blozis, S. A., & Harring, J. R. (2016b). Understanding individual-level change through the basis functions of a latent curve model. *Sociological Methods & Research*, advanced online publication, <https://doi.org/10.1177/0049124115605341>.
- Browne, M. W. (1993). Structured latent curve models. In C. M. Cuadras & C. R. Rao (Eds.). *Multivariate analysis: Future directions 2* (pp. 171–198). Amsterdam: North-Holland.
- Browne, W. J., & Draper, D. (2006). A comparison of Bayesian and likelihood-based methods for fitting multilevel models. *Bayesian Analysis*, *1*, 473–514.
- Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, *101*, 147–158.
- Burchinal, M., & Appelbaum, M. I. (1991). Estimating individual developmental functions: Methods and their assumptions. *Child Development*, *62*, 23–43.
- Cameron, C. E., Grimm, K. J., Steele, J. S., Castro-Schilo, L., & Grissmer, D. W. (2015). Nonlinear Gompertz curve models of achievement gaps in mathematics and reading. *Journal of Educational Psychology*, *107*, 789–804.
- Cheung, M. W. L. (2013). Implementing restricted maximum likelihood estimation in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *20*, 157–167.
- 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.
- Coffman, D. L., & Millsap, R. E. (2006). Evaluating latent growth curve models using individual fit statistics. *Structural Equation Modeling*, *13*, 1–27.
- Coulombe, P., Selig, J., & Delaney, H. (2016). Ignoring individual differences in times of assessment in growth curve modeling. *International Journal of Behavioral Development*, *40*, 76–86.
- Cudeck, R. (1996). Mixed-effects models in the study of individual differences with repeated measures data. *Multivariate Behavioral Research*, *31*, 371–403.
- Cudeck, R., & du Toit, S. H. (2002). A version of quadratic regression with interpretable parameters. *Multivariate Behavioral Research*, *37*, 501–519.
- Cudeck, R., & Harring, J. R. (2007). Analysis of nonlinear patterns of change with random coefficient models. *Annual Review of Psychology*, *58*, 615–637.
- Curran, P. J. (2003). Have multilevel models been structural equation models all along?. *Multivariate Behavioral Research*, *38*, 529–569.
- Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, *62*, 583–619.
- Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development*, *11*, 121–136.
- Edwards, L. J., Muller, K. E., Wolfinger, R. D., Qaqish, B. F., & Schabenberger, O. (2008). An R^2 statistic for fixed effects in the linear mixed model. *Statistics in Medicine*, *27*, 6137–6157.
- Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford Press.
- Ferrer, E., Hamagami, F., & McArdle, J. J. (2004). Modeling latent growth curves with incomplete data using different types of structural equation modeling and multilevel software. *Structural Equation Modeling*, *11*, 452–483.
- Ferron, J. M., Bell, B. A., Hess, M. R., Rendina-Gobioff, G., & Hibbard, S. T. (2009). Making treatment effect inferences from multiple-baseline data: The utility of multilevel modeling approaches. *Behavior Research Methods*, *41*, 372–384.
- Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2004) *Applied longitudinal analysis*. Hoboken: Wiley.
- Fouladi, R. T. (2000). Performance of modified test statistics in covariance and correlation structure analysis under conditions of multivariate nonnormality. *Structural Equation Modeling*, *7*, 356–410.
- Grimm, K. J., & Ram, N. (2009). Nonlinear growth models in M plus and SAS. *Structural Equation Modeling*, *16*, 676–701.
- Grimm, K. J., Ram, N., & Estabrook, R. (2016). *Growth modeling: Structural equation and multilevel modeling approaches*. New York: Guilford Press.
- Grimm, K. J., Ram, N., & Hamagami, F. (2011). Nonlinear growth curves in developmental research. *Child Development*, *82*, 1357–1371.
- Hamagami, F. (1997). A review of the Mx computer program for structural equation modeling. *Structural Equation Modeling*, *4*, 157–175.
- Hancock, G. R., Kuo, W. L., & Lawrence, F. R. (2001). An illustration of second-order latent growth models. *Structural Equation Modeling*, *8*, 470–489.
- Harring, J. R., Kohli, N., Silverman, R. D., & Speece, D. L. (2012). A second-order conditionally linear mixed effects model with observed and latent variable covariates. *Structural Equation Modeling*, *19*, 118–136.
- Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association*, *72*, 320–338.
- Hedeker, D., & Gibbons, R. D. (2006). *Longitudinal data analysis*. Hoboken: Wiley.
- Herzog, W., & Boomsma, A. (2009). Small-sample robust estimators of noncentrality-based and incremental model fit. *Structural Equation Modeling*, *16*, 1–27.
- Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York: Routledge.
- Johnson, P. C. (2014). Extension of Nakagawa & Schielzeth's R^2 GLMM to random slopes models. *Methods in Ecology and Evolution*, *5*, 944–946.
- Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research*, *44*, 486–507. <https://doi.org/10.1177/0049124114543236>
- Kenny, D. A., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling*, *10*, 333–351.
- Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, *53*, 983–997.
- Kenward, M. G., & Roger, J. H. (2009). An improved approximation to the precision of fixed effects from restricted maximum likelihood. *Computational Statistics and Data Analysis*, *53*, 2583–2595.
- Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, *38*, 963–974.
- Ledermann, T., & Kenny, D. A. (2017). Analyzing dyadic data with multilevel modeling versus structural equation modeling: A tale of two methods. *Journal of Family Psychology*, *31*, 442–452.

- Leite, W. L., & Stapleton, L. M. (2011). Detecting growth shape misspecifications in latent growth models: An evaluation of fit indexes. *The Journal of Experimental Education*, 79, 361–381.
- Li, F., Duncan, T. E., & Acock, A. (2000). Modeling interaction effects in latent growth curve models. *Structural Equation Modeling*, 7, 497–533.
- Liu, S., Rovine, M. J., & Molenaar, P. (2012). Selecting a linear mixed model for longitudinal data: repeated measures analysis of variance, covariance pattern model, and growth curve approaches. *Psychological Methods*, 17, 15–30.
- Maas, C. J., & Hox, J. J. (2004). Robustness issues in multilevel regression analysis. *Statistica Neerlandica*, 58, 127–137.
- McArdle, J. J. (1986). Latent variable growth within behavior genetic models. *Behavior Genetics*, 16, 163–200.
- McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*, 60, 577–605.
- McCoach, D. B., Rambo, K. E., & Welsh, M. (2013). Assessing the growth of gifted students. *The Gifted Child Quarterly*, 57, 56–67.
- McNeish, D. (2016a). On using Bayesian methods to address small sample problems. *Structural Equation Modeling*, 23, 750–773.
- McNeish, D. (2016b). Using data-dependent priors to mitigate small sample size bias in latent growth models: A discussion and illustration using Mplus. *Journal of Educational and Behavioral Statistics* 41, 27–56.
- McNeish, D., & Harring, J. R. (2017). Correcting model fit criteria for small sample latent growth models with incomplete data. *Educational and Psychological Measurement*, advance online publication, <https://doi.org/10.1177/0013164416661824>
- McNeish, D., & Stapleton, L.M. (2016). The effect of small sample size on two level model estimates: A review and illustration. *Educational Psychology Review*, 28, 295–314.
- McNeish, D., & Wentzel, K.R. (2017). Accommodating small sample sizes in three level models when the third level is incidental. *Multivariate Behavioral Research*, 52, 200–215.
- Mehta, P. D., & Neale, M. C. (2005). People are variables too: Multilevel structural equations modeling. *Psychological Methods*, 10, 259–284.
- 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.
- Muthén, L. K., & Muthén, B. (2012). *Mplus user's guide* (Version 7). Los Angeles: Muthén & Muthén.
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4, 133–142.
- Nevitt, J., & Hancock, G. R. (2004). Evaluating small sample approaches for model test statistics in structural equation modeling. *Multivariate Behavioral Research*, 39, 439–478.
- Newman, D. A. (2003). Longitudinal modeling with randomly and systematically missing data: A simulation of ad hoc, maximum likelihood, and multiple imputation techniques. *Organizational Research Methods*, 6, 328–362.
- Peugh, J. L., & Enders, C. K. (2005). Using the SPSS mixed procedure to fit cross-sectional and longitudinal multilevel models. *Educational and Psychological Measurement*, 65, 717–741.
- Potthoff, R. F., & Roy, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika*, 51, 313–326.
- Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, 31, 437–448.
- Preacher, K. J., & Hancock, G. R. (2015). Meaningful aspects of change as novel random coefficients: A general method for reparameterizing longitudinal models. *Psychological Methods*, 20, 84–101.
- Preacher, K. J., Wichman, A. L., MacCallum, R. C., & Briggs, N. E. (2008). *Latent growth curve modeling*. Thousand Oaks: Sage.
- Prosser, R. J., Rasbash, J., & Goldstein, H. (1991). *ML3: Software for 3-level analysis. User's guide for V.2*. London: Institute of Education, University of London.
- Rabe-Hesketh, S., & Skrondal, A. (2012). *Multilevel and longitudinal modeling using Stata* (3rd ed.). College Station: Stata Press.
- Rao, C. R. (1965). The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. *Biometrika*, 52, 447–458.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models*. Thousand Oaks: Sage.
- Raykov, T., & Marcoulides, G. A. (2012). *A first course in structural equation modeling*. Mahwah: Routledge.
- Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, 126, 3–25. <https://doi.org/10.1037/0033-2909.126.1.3>
- SAS Institute. (2008). *SAS/STAT 9.2 user's guide*. Cary: Author.
- Schaalje, G. B., McBride, J. B., & Fellingham, G. W. (2002). Adequacy of approximations to distributions of test statistics in complex mixed linear models. *Journal of Agricultural, Biological, and Environmental Statistics*, 7, 512–524.
- Serang, S., Grimm, K. J., & McArdle, J. J. (2016). Estimation of time-unstructured nonlinear mixed-effects mixture models. *Structural Equation Modeling*, 23, 856–869.
- Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.
- Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling*. Boca Raton: Chapman & Hall/CRC.
- Snijders, T. A., & Bosker, R. J. (1994). Modeled variance in two-level models. *Sociological Methods & Research*, 22, 342–363.
- StataCorp. (2013). *Stata 13 base reference manual*. College Station: Stata Press.
- Sterba, S. K. (2014). Fitting nonlinear latent growth curve models with individually varying time points. *Structural Equation Modeling*, 21, 630–647.
- Stiratelli, R., Laird, N., & Ware, J. H. (1984). Random-effects models for serial observations with binary response. *Biometrics*, 40, 961–971.
- Swain, A. J. (1975). *Analysis of parametric structures for variance matrices*. Unpublished doctoral dissertation, University of Adelaide, Australia.
- Tourangeau, K., Nord, C., Lê, T., Sorongon, A. G., & Najarian, M. (2009). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K): Combined user's manual for the ECLS-K Eighth-Grade and K–8 full sample data files and electronic codebooks* (Publication NCES 2009-004). Washington, DC: National Center for Education Statistics.
- Tucker, L. R. (1958). Determination of parameters of a functional relation by factor analysis. *Psychometrika*, 23, 19–23.
- van de Schoot, R., Broere, J. J., Perryck, K. H., Zondervan-Zwijenburg, M., & van Loey, N. E. (2015). Analyzing small data sets using Bayesian estimation: The case of posttraumatic stress symptoms following mechanical ventilation in burn survivors. *European Journal of Psychotraumatology*, 6, 25216. <https://doi.org/10.3402/ejpt.v6.25216>
- Vonsh, E. F., Chinchilli, V. M., & Pu, K. (1996). Goodness-of-fit in generalized nonlinear mixed-effects models. *Biometrics*, 52, 572–587. <https://doi.org/10.2307/2532896>
- 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.
- 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. <https://doi.org/10.1037/0033-2909.116.2.363>
- Wood, P. K., & Jackson, K. M. (2013). Escaping the snare of chronological growth and launching a free curve alternative: General deviance as latent growth model. *Development and Psychopathology*, 25, 739–754.
- Wu, W., & Lang, K. M. (2016). Proportionality Assumption in Latent Basis Curve Models: A Cautionary Note. *Structural Equation Modeling*, 23, 140–154.
- Wu, W., & West, S. G. (2010). Sensitivity of fit indices to misspecification in growth curve models. *Multivariate Behavioral Research*, 45, 420–452.
- 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.
- Xu, R. (2003). Measuring explained variation in linear mixed effects models. *Statistics in Medicine*, 22, 3527–3541.
- Yuan, K. H. (2005). Fit indices versus test statistics. *Multivariate Behavioral Research*, 40, 115–148.
- Zeger, S. L., Liang, K. Y., & Albert, P. S. (1988). Models for longitudinal data: A generalized estimating equation approach. *Biometrics*, 44, 1049–1060.