Growth Modeling Using Random Coefficient Models: Model Building, Testing, and Illustrations

PAUL D. BLIESE
Walter Reed Army Institute of Research

ROBERT E. PLOYHART
George Mason University

In this article, the authors illustrate how random coefficient modeling can be used to develop growth models for the analysis of longitudinal data. In contrast to previous discussions of random coefficient models, this article provides step-by-step guidance using a model comparison framework. By approaching the modeling this way, the authors are able to build off a regression foundation and progressively estimate and evaluate more complex models. In the model comparison framework, the article illustrates the value of using likelihood tests to contrast alternative models (rather than the typical reliance on tests of significance involving individual parameters), and it provides code in the open-source language R to allow readers to replicate the results. The article concludes with practical guidelines for estimating growth models.

Of all the types of data that organizational researchers are likely to encounter, longitudinal data are arguably the type that presents the most analytic challenges. The complexity associated with the analysis of longitudinal data is both conceptual and methodological. On the conceptual side are questions revolving around the specific goals of the analysis and the nature of the change process (Chan, 1998, 2002). For instance, in some cases, the goal of a longitudinal analysis may simply be to determine whether there are mean differences in treatment conditions over time; in other cases, a researcher may be more interested in patterns of change over time. In still other cases, the goal may be a combination of understanding both mean differences and changes in patterns. Clearly, the first case is a much simpler analytic exercise than the latter two; yet even in the simple case, researchers must further contend with methodological issues that are unique to longitudinal data.

The methodological challenges can be daunting. For instance, because longitudinal data are collected from single entities (usually persons) over multiple times, it is likely that there will be some degree of nonindependence in the responses. That is, multiple responses from an individual will tend to be correlated, and this correlation violates the

Organizational Research Methods, Vol. 5 No. 4, October 2002 362-387
DOI: 10.1177/109442802237116

statistical assumption of independence underlying many common data analytic techniques (Kenny & Judd, 1986). To further complicate matters, it is also likely that responses temporally close to each other (e.g., Responses 1 and 2) will be more strongly correlated than responses temporally far apart (e.g., Responses 1 and 4). Finally, it is likely that responses will tend to become either more or less variable over time. For instance, individuals starting jobs may initially have a high degree of variability in performance, but over time, variance in job performance may diminish. A well-fitting longitudinal model needs to account for each of these methodological factors.

A final hurdle stems from the fact that even when researchers have a clear vision of their conceptual goals and a clear understanding of how the methodological issues need to be addressed, they may simply lack the procedural knowledge required to translate theory into analyses. Although this hurdle may seem trivial compared to the conceptual and methodological issues, we believe that in many ways, the procedural hurdle inhibits the application of longitudinal analyses as much as do the conceptual and methodological barriers.

In short, longitudinal research questions require a different way of thinking about design and analysis than do cross-sectional designs, even to the extent that the data sets must be structured differently. Longitudinal designs have important conceptual and methodological issues that are simply not present in cross-sectional research (e.g., equating time as an independent variable, correlated errors) and are, thus, not fully described in most statistical texts.

Given these challenges, it is not surprising that there are relatively few complex longitudinal analyses presented in the organizational literature (for exceptions, see Chan, Ramey, Ramey, & Schmitt, 2000; Chan & Schmitt, 2000; Deadrick, Bennett, & Russell, 1997; Garst, Frese & Molenaar, 2000; Hofmann, Jacobs, & Baratta, 1993; Lance, Vandenberg, & Self, 2000; Ployhart & Hakel, 1998).[1] We suspect that organizational researchers have access to more longitudinal data than is reflected in research publications, and we also believe that longitudinal analyses have considerable promise for testing organizational theory.

Consequently, the goals of this article are to discuss and illustrate a longitudinal analysis technique generally referred to as *growth modeling*. Growth modeling involves looking at how individuals (or units, groups, organizations, etc.) change over time and whether there are differences in patterns of change. For example, if one examined job performance over time, one might be interested in whether there were identifiable differences in patterns of performance over time (e.g., some improve, some get worse, some stay the same). Most of the early applications of growth models were in developmental psychology contexts (hence the term *growth* modeling) but have subsequently had a significant impact on the analysis of longitudinal data in many substantive domain areas.

Because there are currently several excellent discussions of growth modeling in the organizational literature (see, for instance, Chan, 1998, 2002; Chou, Bentler, & Pentz, 1998; Lance, Meade, & Williamson, 2000; Lance, Vandenberg, & Self, 2000; Ployhart & Hakel, 1998; Ployhart, Holtz, & Bliese, 2002; Willet & Sayer, 1994), it is important to highlight the unique goals of this article. First, we discuss growth modeling using a random coefficient modeling (RCM) framework instead of a structural equation modeling (SEM) framework. Many of the existing discussions of growth modeling assume that the reader has familiarity with SEM (e.g., Chan, 1998, 2002; Lance, Vandenberg, & Self, 2000; Ployhart & Hakel, 1998; Willett & Sayer, 1994). In contrast, by approaching growth modeling from an RCM framework, we are able to illustrate a model-building approach originating from a basic regression framework. This, in turn, will make the material accessible to a broader audience. Note that our approach is not to argue for RCM over SEM. Similarities and differences, as well as strengths and weaknesses, of modeling growth curves via random coefficient models versus structural equation models can be found in detail elsewhere (e.g., Gottman, 1995; Hand & Crowder, 1996; Willett & Sayer, 1994), and in most cases, results from both approaches are nearly identical. Rather, our purpose is to illustrate how one might use RCM for the analysis of longitudinal growth data.

The second reason this article is unique is that discussions of RCM in the organizational literature typically discuss RCM in reference to multilevel analyses (e.g., Bliese, 2002b; Hofmann, Griffin, & Gavin, 2000). Although the basic multilevel approach extends naturally to longitudinal research designs (e.g., Bryk & Raudenbush, 1987), there are important differences between applying RCM in multilevel and applying it in longitudinal designs. The differences are the result of the fact that the level-1 variable (time) has a chronological ordering in growth models, whereas level-1 variables typically have no structure in multilevel models. This critical difference raises two new issues that must be addressed in growth models. First, one must carefully consider how to treat time as a predictor to test various growth functions (linear, quadratic, etc). Second, one must pay more attention to the proper form of the error variance-covariance matrix. In growth models, the error variance-covariance matrix typically displays much more complex patterns than do multilevel models.

The third contribution of this article is that it proposes a model building approach because such approaches are arguably preferable in arriving at correct statistical inferences (see, for instance, Gonzales & Griffin, 2001). Although model building has been covered in some previous work (e.g., Bryk & Raudenbush, 1992; Hofmann, 1997), the correct order for building longitudinal RCM models is only generally described (e.g., Bryk & Raudenbush, 1992) or described primarily in the context of multilevel examples. This may contribute to misunderstandings in how to correctly build RCM growth models. Some evidence of this is shown by the fact that most applications of RCM to longitudinal data report only a single model and the parameters within that model. Yet from statistical and theory-testing perspectives, an arguably better approach is to compare competing statistical models that represent alternative theories about the nature of the data. We outline a general strategy for model building and show how to test and compare models in the building process. We believe that this approach will help users of RCM obtain the most information from their data.

Fourth, although there is literature that discusses growth modeling from an RCM framework, much of this work is currently in need of revision. Analytic packages change quickly; thus, limitations associated with RCM noted by Willet and Sayer (1994) and Chou et al. (1998) are really limitations associated with specific software packages not with the method per se. As an example, the previous literature has noted that RCM packages are not capable of modeling a wide variety of different error variance-covariance structures. This limitation, however, was unique to the early versions of software (such as hierarchical linear modeling [HLM]; Bryk, Raudenbush, & Congdon, 1994) and is not a shortcoming with RCM programs in general.

Finally, this article differs from most others in that the actual code is provided along with the examples. This is not an entirely unique approach, as Singer (1998) provided code in her illustration of growth modeling using PROC MIXED in SAS; however, we expand Singer's idea by illustrating growth modeling using open-source software. Specifically, we use the computing environment R and the NLME (Nonlinear and Linear Mixed Effects models) package (Pinheiro & Bates, 2000). Both R and the NLME package are freely available on the Web (http://cran.r-project.org) and run on common operating systems. The software is supported by the community efforts of statisticians and freely available to users in academia and industry. A side advantage associated with R is that the code required to conduct growth models is succinct; therefore, it can be easily integrated into the text. Thus, readers who are not directly interested in R, but are interested in the other aspects of the article, will encounter relatively few distractions.

Growth Modeling Background and Illustration

Conceptually, our goal is to describe growth models by incrementally expanding on a fundamental regression model. To make this process concrete, we will use an example with data collected from the U.S. Army. In this data, job satisfaction was assessed from 495 respondents at three time intervals 6 months apart. As background, the data were collected over the course of a year as part of an evaluation of how soldiers and leaders adapted to significant technological changes in the work environment. The data are available from the first author for readers interested in replicating the results.

One important aspect of this data is that it has missing values, as is common in most longitudinal research. Specifically, only 421 of the 495 respondents have complete job satisfaction data for the three time periods. In the following analyses, we assume that the data are missing at random (though we do not test this assumption). Data missing at random does not lead to bias in the parameter estimates (e.g., DeShon, Ployhart, & Sacco, 1998; Little & Schlenker, 1995). It will be evident that one of the strengths of RCM is that the missing data pose no particular problems in terms of estimation. Basically, the parameter estimates are based on the available information.

Regression Approach

A reasonable first step in the analysis of the data described above would be to determine whether job satisfaction increased or decreased over the three data collection times. With a little data manipulation, one could use regression to examine job satisfaction trends. To do so, it would be necessary to convert a multivariate data set where each row contained complete data for an individual into a univariate or stacked form where each row represented one time period for an individual. Table 1 provides an example of a univariate data structure.

Notice that each respondent has three rows of data, so there are 1,485 rows rather than 495. Furthermore, time is coded such that Time 0 is the initial time, Time 1 is 6 months later, and Time 2 is 1 year after the initial data collection. By coding the first time period as 0, the intercept refers to the value of the dependent variable at the initial data collection time. With the data in univariate format, we can estimate a simple model where job satisfaction is regressed on time. The code to estimate such a model in R makes use of the linear model (lm) function:

```
> lm(JSAT~TIME,data=univbct,na.action=na.omit)
```

Table 1
Longitudinal Data in a Univariate Form

Subnum	*Time*	*Jsat*
1	0	3.4
1	1	3.6
1	2	3.3
2	0	2.5
2	1	3.5
.		
.		
1,485	2	3.4

Briefly, the command has several components. The formula in the model is `JSAT~TIME,` specifying that job satisfaction is regressed on time. The `data=` gives the name of the data set containing the variables. Finally, the `na.action=na.omit` component instructs lm to omit the rows with missing data and estimate the model with the nonmissing data. Selected results from this analysis are presented in Table 2.

The results indicate that the mean level of job satisfaction at the initial data collection time (Time 0) was 3.22. The results reveal no significant relationship between time and job satisfaction even with the high degree of power associated with more than 1,400 observations.

This example shows that we can technically use regression to analyze longitudinal data. However, if we view statistics as a principled argument (Abelson, 1995), we would quickly have to conclude that the regression approach is fraught with shortcomings. Methodologically, the most serious problem is that the analysis fails to account for the fact that each individual provided three pieces of information, so responses are nonindependent. In this case, ignoring nonindependence leads to standard errors that are too large, thereby decreasing the likelihood of detecting significant results that actually exist (see Bliese, 2002b; Kenny & Judd, 1986). Later we show that accounting for the nonindependence in a random intercept model reveals that time is significantly related to job satisfaction. In addition, there are methodological issues that arise because the predictor (time) has a systematic, chronological structure. Clearly, these issues are not captured in this simple regression model. We return to this topic when we address within-group error variance-covariance matrices in Step 4 of the modeling illustration.

Finally, the regression analysis provides very little information given the relative complexity of the data. Presumably, a researcher would want to know things such as (a) What are the characteristics of people who have high versus low job satisfaction? (b) Does job satisfaction increase for everyone, or do some people increase whereas others decrease? and (c) If respondents show different patterns of change in job satisfaction over time, what predicts these differences? Thus, the regression procedure described here cannot test questions relating to individual differences in growth. To address these methodological and conceptual shortcomings, we need to build on our regression model and begin to specify a growth model.

Table 2
Ordinary Least Squares Model Regressing Job Satisfaction on Time

Variable	*Estimate*	SE	t *Value*	p
Intercept	3.222	0.039	82.223	.000
TIME	0.052	0.030	1.696	.090

Basic Growth Model

To facilitate our discussion of growth models, it will be convenient to introduce some notation. The notation that we will use is from Bryk and Raudenbush (1992). The notation is similar to regression notation, although we will use different symbols to represent intercepts and slopes.

Consider, first, our previous example where we regressed job satisfaction on time. In Bryk and Raudenbush (1992) notation, this model is specified as

$$Y_i = \pi_0 + \pi_1(\text{Time}_i) + r_i \tag{1}$$

$$\pi_0 = \beta_{00} \tag{2}$$

$$\pi_1 = \beta_{10}, \tag{3}$$

where Y_i represents a single job satisfaction score (one row in Table 1), π_0 represents the intercept (which was 3.22), π_1 represents the slope for time (which was .05), and r_i represents the residual error. The terms π_0 and π_1 are not typical regression notation; however, Equations (1) through (3) show that they are directly equivalent to the more familiar β_{00} and β_{10} terms, respectively. In combined form, Equations (1) through (3) yield the familiar regression equation, $Y_i = \beta_{00} + \beta_{10}(\text{Time}_i) + r_i$.

The most significant shortcomings with the regression model are that it (a) assumes that all individuals start the same and change at the same rate (that is, it is a fixed-effects model) and (b) does not account for the fact that each individual provides multiple pieces of data. In our example, Y_1, Y_2, and Y_3 all come from one respondent, and so these three values of Y are likely to be related to each other in terms of absolute levels and in terms of patterns of change over time. We can explore the idea that absolute levels and patterns of change are unique for each individual by estimating a simple regression model using only the first three rows of data (the first individual's data). If we do so, we find that the intercept (initial job satisfaction value) for the first respondent is 1.22 and the slope associated with time is .67. The first respondent has lower initial job satisfaction than does the average respondent, and he or she has a faster increase in job satisfaction over time than do the other respondents. Figure 1 provides a visual representation of the intercept and slope differences for the first 12 respondents in the sample. Respondent 1 is in the lower left-hand corner. As the figure shows, there appears to be considerable individual variability in levels and patterns of job satisfaction over time, so it is important to consider this information in the analyses.

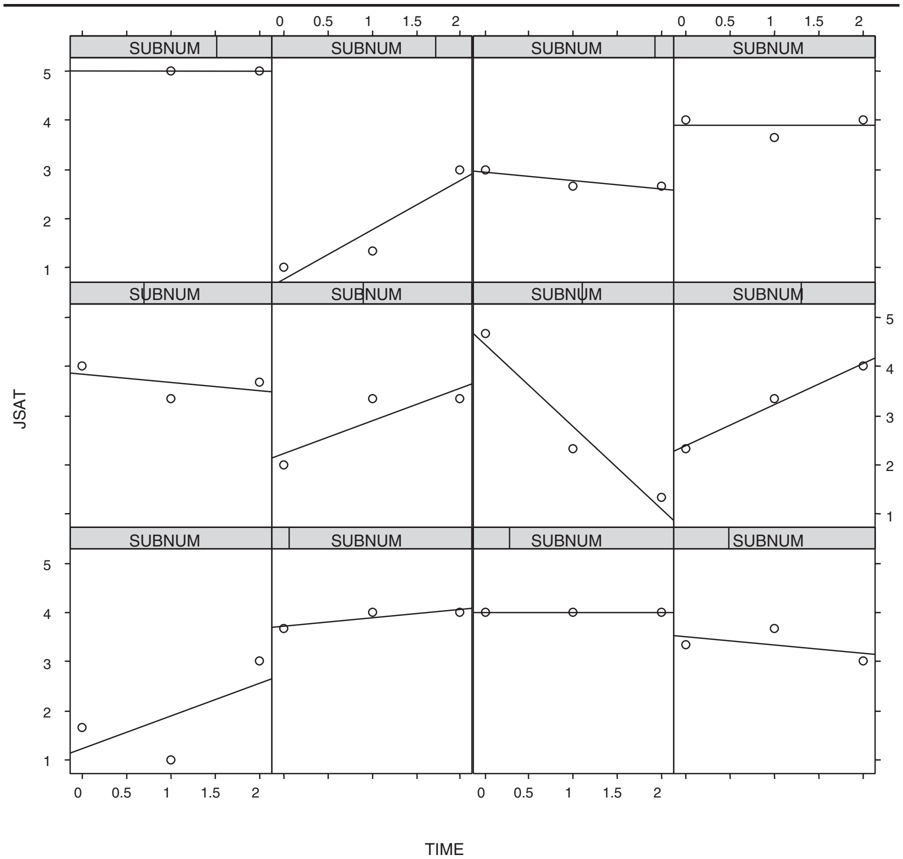

Figure 1: Relationship Between Time and Job Satisfaction for 12 Participants

One could employ a two-step process where the slopes and intercepts from the individual models were used in subsequent analyses. For instance, one might expect that age would be related to both slopes and intercepts. Specifically, older respondents might tend to have higher initial job satisfaction (a positive relationship with the intercept) and no change in job satisfaction over time (nonsignificant relationship with the time slope). If we estimated 495 regression equations and exported the intercept and slope estimates, we could look at the relationship between respondent age and the slope and intercept values and test our hypothesis.

One major problem with this approach, however, is that one could easily engage in this analysis and be left trying to explain nothing more than random variation. That is, when we look at Figure 1, we need to keep in mind that each individual's growth trajectory is based on three data points at most. With so few points describing an individual's trajectory, one would expect considerable variability in both intercept and slope estimates across individuals, even if individual values were all randomly drawn from the same population. Consequently, a critical first step in modeling growth trajectories is to conduct statistical tests to determine whether the variability one observes is evidence of (a) random fluctuation or (b) meaningful individual differences. Unfortunately, the practice of estimating separate regression equations and using the parameter estimates from the equations sheds no light on the issue of random versus meaningful variation. Thus, there is a need for another approach.

Growth modeling in RCM lets one apply the logic of estimating separate regression equations without actually doing so. One can think of growth modeling (and multilevel modeling in general) as a compromise between two approaches (Kreft & de Leeuw, 1998). The compromise is between (a) a regression model that ignores the fact that observations are nested within individuals (our first example) and (b) a series of regression models that estimate a separate model for each individual (our second example).

The specific form of the compromise is that the intercepts and slopes estimated in Equation (1) are permitted to randomly vary among individuals. If the variation in intercepts is statistically found to be large, the model assumes that individuals differ in terms of their dependent variables (job satisfaction in our example). In contrast, if the initial variation in the intercept is small, the model assumes that there are no differences between individuals. Likewise, if the variation in the slope is statistically found to be large, the model assumes that individuals' growth patterns differ; however, if variation in the slope is small, then the model assumes that the rate of change is basically constant across individuals.

In terms of notation, we need to modify Equations (1) through (3) to specify a basic growth model. First, we need to add an index to keep track of the individual respondent j in each equation; second, we need to add a residual term to Equation (2) and Equation (3). Adding the residual term u_{0j} to Equation (2) allows us to estimate the difference between an individual's intercept and the overall or combined intercept. Adding the residual term u_{1j} to Equation (3) allows us to estimate the difference between an individual's slope value and the overall or combined slope value. This modified model is now specified as

$$Y_{ij} = \pi_{0j} + \pi_{1j}(\text{Time}_{ij}) + r_{ij} \tag{4}$$

$$\pi_{0j} = \beta_{00} + u_{0j} \tag{5}$$

$$\pi_{1j} = \beta_{10} + u_{1j}. \tag{6}$$

To be concrete, we know that the overall intercept in our example is $\beta_{00} = 3.22$ and that the intercept for the first individual in the sample is $\pi_{01} = 1.22$; thus, u_{01} for the first individual is 2.00. Likewise, the overall slope is $\beta_{10} = .05$, and the slope for the first individual is $\beta_{11} = .67$, so u_{11} for the first individual is .62. In practice, RCM programs would not actually estimate u_{01} as being 2.00, nor u_{11} as being .62, because RCM uses overall sample information to adjust estimates; nonetheless, the example conceptually illustrates what occurs.

In terms of modeling, we are not really interested in the residuals per se; rather, we are interested in the variance of the residuals. In terms of notation, the intercept variance is defined as τ_{00}, the slope variance is defined as τ_{11}, and the covariance between the slope and the intercept is τ_{01}. This latter value can be used to estimate the correlation between intercept values and slopes. For instance, individuals with high initial levels of job satisfaction might have small changes in job satisfaction over time or vice versa. The final variance term in the basic growth model is σ^2, the variance term associated with the within-group residual r_{ij}.

The elements of the model defined in Equations (4) though (6) can be combined into a single equation:

$$Y_{ij} = [\beta_{00} + \beta_{10}(\text{Time}_{ij})] + [u_{0j} + u_{1j}(\text{Time}_{ij}) + r_{ij}]. \tag{7}$$

The result looks very much like a regression model with two extra error terms, u_{0j} and $u_{1j}(\text{Time}_{ij})$, representing the residual intercept and residual slope terms, respectively.

The combined model is generally described in terms of the two components in the brackets. The elements in the brackets on the left are considered the "fixed" portion of the model, and the elements in the bracket on the right are considered the "random" portion of the model. The distinction between the fixed and random components of the basic growth model are important in specifying the model in RCM modeling programs and will become clear in the following examples.

In the next section, we conduct a step-by-step generation of a basic growth model using a model-contrasting approach. The goal of this illustration is to begin with a simple regression model and end up with a complete growth model by successively adding complexity. The sequence of steps that we follow as we build the regression model into a growth model differs in minor ways from the sequence of steps that we recommend for conducting growth modeling in the final part of the article (the reason is because we add some unnecessary steps in the illustration to facilitate understanding). Nonetheless, the steps illustrated in the following examples are consistent with our later recommendations and set the stage for the proposed model contrasting sequence outlined later.

Level 1: Estimating the Basic Model in R

As noted previously, growth modeling and multilevel modeling in R are conducted using the NLME library written by Pinheiro and Bates (2000). Within the NLME library, we will work primarily with the linear mixed-effects (lme) function although we will initially make use of the generalized least squares (gls) function in the illustrative example. Basic information on running R and on data manipulation can be found in Bliese (2002a) among other places.

Step 1

The first step in progressing from a regression model to a growth model via model comparisons is to establish a simple model without any random effects to serve as a baseline. One option for such a model is the familiar regression model identified in Equations (1) through (3). We will start with the simple regression model and progressively add complexity in terms of random effects. At each step, we will compare log-likelihood ratios (deviances) between models to aid decisions about including specific terms. Because we want to develop the most parsimonious model, we test whether adding more complexity to the model improves model fit above and beyond the existing terms in the model.

To use Equations (1) through (3) as the baseline model in the deviance tests, we need to ensure that the estimation techniques used in the baseline model match those used in subsequent models. Doing so ensures similarity in deviance terms.[2] Random coefficient modeling routines use maximum likelihood algorithms and return likelihood deviances; therefore, it is important to estimate the baseline model using maximum likelihood algorithms as well. To estimate Equations (1) through (3) using a maximum likelihood algorithm, we use the gls function.

```
> model.1<-gls(JSAT~TIME,data=univbct,na.action=na.omit)
```

The use of the assignment character `<-` saves the output of the gls function to an object named `model.1.` This object will be compared to more complex models in following analyses. A summary of `model.1` (not shown) reveals that the parameter estimates and standard errors are identical to those reported in Table 2 (from the ordinary least squared model) as would be expected. Recall that in this model, time is not significantly related to job satisfaction.

Step 2

Step 2 involves determining whether the fit of data can be improved by adding a random intercept term to the baseline model. That is, we are interested in seeing if we can improve our model by allowing individuals to vary in terms of their overall levels of job satisfaction. Figure 1 suggests that adding a random intercept will improve our model, but it is important to confirm our expectation. We specify a random intercept model using the lme function.

```
> model.2<-lme(JSAT~TIME,random=~1|SUBNUM,data=univbct,
na.action=na.omit)
```

The fixed portion of the model is `~TIME`, which can also be expressed as `~1 + TIME` with `1` representing the intercept. The random part of the model is expressed as `~1|SUBNUM`. The `1` in the random portion of the model represents the intercept. The `~1|SUBNUM` portion indicates that intercepts are permitted to randomly vary across respondents. Note that we continue to use the `na.action=na.omit` option in this model. As with the linear model, this option instructs lme to omit rows with missing data and estimate the model on whatever data are available. Conceptually, this means that some individuals will have intercept estimates based on three observations whereas others will have intercept estimates based on one or two observations. Importantly, however, the program does not exclude individuals who may be missing data at one or more time periods.

The random intercept model that we have estimated is $Y_{ij} = [\beta_{00} + \beta_{10}(\text{Time}_{ij})] + [u_{0j} + r_{ij}]$. This model differs from the full level-1 model presented in Equations (4) through (6) in that Equation (6) is more restrictive and is specified as $\pi_{1j} = \beta_{10}$ instead of as $\pi_{1j} = \beta_{10} + u_{1j}$. Essentially, in Model 2 we restrict the slope between time and job satisfaction to be constant for each individual, but we allow individuals to differ in terms of job satisfaction.

Model 1 and Model 2 differ only in terms of random effects, so they can be contrasted using likelihood tests.[3] That is, both models have the same predictors; however, the second model allows intercepts to randomly vary among respondents whereas the first model does not. The ANOVA function is a generic function used to contrast alternative models and can be used to compare –2 log likelihood values (i.e., deviances) between `model.1` and `model.2`. The significance of the –2 log likelihood difference is based upon a chi-squared distribution using the *df* associated with the number of model differences between the contrasted models.

```
> anova(model.1, model.2)
          Model  df  logLik     Test    L.Ratio  p-value
model.1   1      3   -1880.773
model.2   2      4   -1726.617  1 vs 2  308.3123 <.0001
```

In this example, the likelihood ratio of 308.31 is significant on the one degree of freedom associated with the fixed versus free intercept. Thus, we conclude that the model that allows individuals to randomly vary in terms of their initial job satisfaction values fits the data better than does a model that fixes the intercept to be constant across individuals.

An examination of the parameter estimates and standard errors reveal that the parameter estimate for time in the random intercept model is nearly identical to that in the regression model (.0518 versus .0516, respectively). As expected from Bliese (2002b) and Kenny and Judd (1986), however, the standard error associated with time is smaller in the random intercept model than in the regression model (.022 versus .030). Due to the decreased magnitude of the standard error, time is significantly related to job satisfaction in the random intercept model ($p < .02$). Thus, accounting for between-person differences in job satisfaction (i.e., nonindependence due to respondent) increases the power to detect level-1 effects.

Step 3

The third step is to determine whether there is significant slope variation among respondents. A model with a random slope for time is estimated as follows:

```
> model.3<-lme(JSAT~TIME,random=~1+TIME|SUBNUM,
data=univbct,na.action=na.omit)
```

or, alternatively, as

```
> model.3<-update(model.2,random=~1+TIME|SUBNUM)
```

In Model 3, we have added `TIME` to the random component of the model. By so doing, we indicate that the slope associated with `TIME` can randomly vary among respondents.[4] In Bryk and Raudenbush's (1992) notation, the model being specified is the complete model presented in Equations (4) through (6).

The comparison of Model 2 and Model 3 returns a log likelihood value of 31.10. This value is significant on two degrees of freedom and indicates that the model with the random slope (Model 3) fits the data significantly better than does the model without the random slope. There are two degrees of freedom associated with the difference between Models 2 and 3 because Model 3 estimates the variation in the slopes among individuals and estimates the covariance between the intercept and the slope.

The `VarCorr` function provides estimates of the random effects—that is, estimates of the variances associated with the intercept and slope variance and an estimate of the correlation between the intercept and slope.

```
>VarCorr(model.3)
. . . .
                Variance       StdDev         Corr
(Intercept)     0.6701022      0.8185977      (Intr)
TIME            0.1005343      0.3170714      -0.578
Residual        0.3286164      0.5732508
```

The results from this command reveal that the intercept variance (τ_{00}) is .67; the variance of the time-job satisfaction slope across individuals (τ_{11}) is .10; the residual variance (σ^2) is .33, and the correlation between the slope and intercept is –.58. The negative correlation indicates that individuals with high initial job satisfaction values tend to have weak slopes, and individuals with low initial job satisfaction values tend to have strong slopes. Note that the correlation is rarely interpreted substantially because its magnitude is heavily influenced by arbitrary factors such as one's centering strategy. For instance, adding a constant to time so that it ranges from 1 to 3 instead of 0 to 2 changes the correlation to –.77. Hofmann and Gavin (1998) provided details about this phenomenon.

The model specified in Step 3 is clearly an improvement over the basic regression model initially encountered in Equation (1) as evidenced by the likelihood contrasts. Substantively, it provides a more appropriate description of the data than the regression model. RCM, however, allows us to further improve Model 3 by including estimates of autocorrelation and heteroscedasticity in the within-group errors.

Step 4

As previously noted, discussions of RCM in the organizational literature generally apply the methodology to multilevel modeling (e.g., Bliese, 2002b; Hofmann, Griffin, & Gavin, 2000). Importantly, however, in multilevel models there is no systematic order associated with level-1 variables. There is, for instance, generally nothing of analytic importance associated with being the first respondent versus the last respondent in a group of 10 respondents. Consequently, with multilevel data, it is logical to assume that the within-group errors (r_{ij}) are independent once group membership is controlled (Bryk & Raudenbush, 1992; Pinheiro & Bates, 2000).

In contrast, the level-1 variable of time in growth models has a logical ordering, and this logical ordering raises the possibility that within-person errors will be related to the level-1 variable. The ordering associated with time potentially affects the independence of within-person errors in two ways. First, it is possible that the within-person errors will display some degree of autocorrelation—responses close in time may be more strongly related to each other than responses far apart in time. Second, within-person errors may not be independent because responses may become more or less variable over time (Willett & Sayer, 1994).

In general, failing to account for autocorrelation leads to underestimation of the standard errors and inflated *t* values. Failing to account for heteroscedasticity is more complicated. In some cases, it can lead to inflated standard errors; and in other cases, it can lead to underestimated standard errors; nonetheless, it certainly leads to incorrect estimates of standard errors and therefore to possibly incorrect inferences. In modeling, the way to control for these effects is to (a) examine the residuals and determine whether they show evidence of autocorrelation or heteroscedasticity, and (b) include terms to account for the nature of the autocorrelation and heteroscedasticity in the final model if one finds evidence of these effects in the residuals. The procedure is analogous to including known covariates in a regression model prior to interpreting effects of interest to avoid bias. As illustrated below, we can determine whether the residuals show evidence of heteroscedasticity and autocorrelation by contrasting competing models (see also Pinheiro & Bates, 2000).

In the lme function, autocorrelation is included in the model by using the correlation option in lme. For instance, to include a first-order autoregressive structure in our ongoing example, we use the `corAR1()` option.

```
> model.4a<-update(model.3,correlation=corAR1())
```

The log likelihood difference between Model 3 and Model 4a is 6.36 with 1 degree of freedom. This difference is significant, indicating that Model 4a is a better fit than Model 3. A summary of Model 4a indicates that the lag 1 correlation estimate (phi) is .37. The first-order structure specifies that the lag 2 correlation is estimated to be $.37^2$ or .14 and the lag 3 correlation is estimate to be $.37^3$ or .05. This suggests a fairly high degree of autocorrelation despite only three data points.

We can impose still further restrictions on the error variance-covariance matrix by examining whether the errors associated with job satisfaction are homoscesdastic across time. A preliminary analysis suggests that variance in job satisfaction may be decreasing over time. In the lme function, heteroscedasticity is modeled using the weights option and various variance specifications such as `varPower` and `varExp`. The `varPower` option models heteroscedasticity by increasing or decreasing the within-group residual variance by delta (a single estimated power function describing the nature of the variance change). For instance, conceptually, if delta was estimated to be 1.4 and the residual variance at Time 1 was 3, then the estimated residual variance at Time 2 would be $3^{1.4}$ or 4.66, and at Time 3 it would be $4.66^{1.4}$ or 8.62. Pinheiro and Bates (2000) note that `varPower` will not work if the covariate (time) includes a zero (as it does in our example); thus, we use `varExp`. The `varExp` option models heteroscedasticity as an exponent of the covariate instead of as a power function of the covariate, but it is conceptually similar to `varPower`. Model 4b provides the code to model heteroscedasticity in our example.

```
> model.4b<-update(model.4a,weights=varExp(form=~TIME))
```

The log likelihood difference between Model 4a and Model 4b is 3.32. The *p* value of this log likelihood on 1 degree of freedom is $p = .066$. A summary shows that the delta estimate associated with the variance function is –.08. Given the nonsignificant *p* value and small delta estimate, we assume homoscedasticity in subsequent analyses.

Table 3
Growth Model Parameter Estimates

Variable	*Estimate*	SE	t *Value*	p
Intercept	3.219	0.044	72.429	.000
TIME	0.053	0.024	2.214	.027

The final estimates of the fixed effects associated with the growth model are presented in Table 3. Notice that the estimates in Table 3 are similar to those from Table 2 (the simple regression results). In the final model, the slope associated with time is significant, as it was in the random intercept model, whereas time was not significant in the regression model. These results highlight how correctly analyzing the data can lead to more accurate substantive interpretations; indeed, the support for a theory may depend on this information. This is certainly true in our example. If we closely adhere to the $p < .05$ significant value, the regression approach suggests that job satisfaction did not change, whereas the RCM indicates that job satisfaction increased over time.

Level 2: Modeling Intercept and Slope Variation

Broadly speaking, the steps conducted in level-1 analyses were designed to determine the relationship between job satisfaction and time. These steps led us to conclude that (a) there is a small but significant linear increase in job satisfaction over time, (b) individuals differ in terms of their initial job satisfaction levels, and (c) the linear individual growth pattern varies among individuals. Furthermore, we can make these three inferences knowing that we have accounted for a fairly high degree of autocorrelation in the error variance-covariance matrix.

Clearly, the basic growth model that we have described is a significant advance over the regression model that we first introduced; nonetheless, the real strength associated with growth modeling is that it allows us to examine two additional aspects of our longitudinal data. First, it allows us to examine why individuals vary in terms of intercept values (i.e., initial job satisfaction). Second, it allows us to examine why individuals have different slopes for the relationship between time and job satisfaction.

In the following two sections, we explore the idea that age is related to both intercept variation and slope variation. Specifically, we propose that older respondents will have higher initial job satisfaction and tend to have small changes in job satisfaction over time. In contrast, we expect younger respondents to have low initial job satisfaction levels and to show more change over time.

Intercept Variation

We begin level-2 analyses by examining how age is related to initial levels of job satisfaction (i.e., intercept variation). To model this relationship, the individual-level characteristic, age, must be included in the univariate data set. The value for age will be repeated across all observations for each respondent as it is in Table 4.

The model that we test is represented in Equations (8) through (10).

Table 4
Data in Univariate Form With an Individual-Level Attribute (Age)

Subnum	*Time*	*Jsat*	*Age*
1	0	3.4	20
1	1	3.6	20
1	2	3.3	20
2	0	2.5	24
2	1	3.5	24
.			
.			
1,485	2	3.4	19

$$Y_{ij} = \pi_{0j} + \pi_{1j}(\text{Time}_{ij}) + r_{ij} \tag{8}$$

$$\pi_{0j} = \beta_{00} + \beta_{01}(\text{Age}_j) + u_{0j} \tag{9}$$

$$\pi_{1j} = \beta_{10} + u_{1j}. \tag{10}$$

Equations (8) through (10) state that respondent *j*'s initial job satisfaction (π_{0j}) can be modeled as a function of two things. One is the mean level of job satisfaction (β_{00}) for all respondents. The second is a coefficient associated with the individual's age (β_{01}). Note that the error term for the intercept (u_{0j}) now represents the difference between an individual's intercept and the overall intercept after accounting for age. The model is specified as

```
> model.5<-lme(JSAT~TIME+AGE,random=~TIME|SUBNUM,
correlation=corAR1(),na.action=na.omit,data=univbct)
```

Model 5 differs only from Model 4a in that Model 5 includes a new fixed effect, `AGE`. The results are presented in Table 5. Notice that age is positively related to initial levels of job satisfaction. Also notice that there are fewer degrees of freedom for age than for time as age is an individual (level-2) attribute.[5]

In interpreting Table 5, we conclude that in cases where age is 0 and where time is 0, the expected level of job satisfaction is 2.347. In some ways, this interpretation is strange because age will never actually be 0 in this population. Consequently, it may be useful to reparameterize age by grand-mean centering the variable (see Singer, 1998). Grand-mean centering involves subtracting the overall mean from each observation. A model using a grand-mean centered version of age (`AGE2`) is presented in Table 6.

In Table 6, the intercept estimate of 3.216 now represents the initial job satisfaction value for a respondent of average age (25.7 years old). Notice that the *t* values for time and age did not change from those listed in Table 5.

Table 5
Relationship Between Age and Job Satisfaction Intercept

Variable	*Estimate*	SE	df	t *Value*	p
Intercept	2.347	0.146	897	16.086	.000
TIME	0.053	0.024	897	2.205	.028
AGE	0.034	0.005	486	6.241	.000

Table 6
Relationship Between Grand-Mean Centered Age and Job Satisfaction Intercept

Variable	*Estimate*	SE	df	t *Value*	p
Intercept	3.216	0.043	897	74.564	.000
TIME	0.053	0.024	897	2.205	.028
AGE2	0.034	0.005	486	6.241	.000

Modeling Slope Variability

The final aspect of growth modeling involves attempting to determine attributes of individual respondents that are related to slope variability. In this section, we attempt to determine whether age is related to the time-job satisfaction slope variation.

The model that we estimate is identical to that estimated in Equations (8) through (10) except that Equation (10) is now specified as $\pi_{1j} = \beta_{10} + \beta_{11}(\text{Age}_j) + u_{1j}$. This change indicates that the slope between time and job satisfaction for an individual (π_{1j}) is a function of an overall slope estimate (β_{10}), a coefficient associated with individual age (β_{11}), and an error term (u_{1j}). In lme, the model is specified as

```
> model.6<-lme(JSAT~TIME*AGE2,random=~TIME|SUBNUM,
correlation=corAR1(),na.action=na.omit,data=univbct)
```

The only change between Model 5 and Model 6 is that the predictors are modeled as `TIME*AGE2` in Model 6 rather than `TIME+AGE2` as in Model 5. The use of the notation `TIME*AGE2` indicates that job satisfaction should be regressed on time, age, and the Time × Age interaction. Table 7 reveals that the Time × Age interaction is significant (note that we used the grand-mean centered version of `AGE`).

A plot of the interaction is presented in Figure 2. Notice that older respondents tend to report higher levels of job satisfaction at the first time interval, and over the course of the year, their job satisfaction increased only slightly. In contrast, younger respondents tended to report relatively low job satisfaction initially, but job satisfaction had a more pronounced increase over the year.[6]

Table 7
Relationship Between Age and Job Satisfaction Intercept and Slope

Variable	*Estimate*	SE	df	t *Value*	p
Intercept	3.215	0.043	896	74.667	.000
TIME	0.054	0.024	896	2.243	.025
AGE2	0.043	0.007	486	6.180	.000
TIME × AGE2	–0.008	0.004	896	–2.153	.032

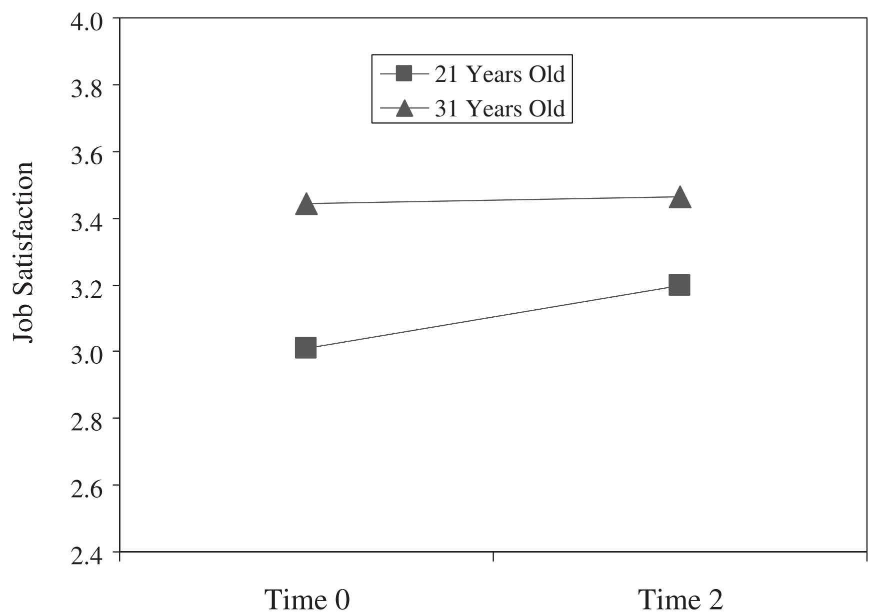

Figure 2: Interaction Between Age and Job Satisfaction Slope

Building and Testing Growth Models: Practical Advice

The first portion of this article illustrated the steps involved in building a growth model and highlighted the value of using log likelihood tests to contrast competing models when establishing the level-1 model. We contend that the model comparison approach is a key strength associated with conducting growth modeling in RCM for several reasons. First, the model contrasting approach arguably requires a greater understanding of the growth process than does running and interpreting parameter estimates from a single model. Second, the approach can be very useful in identifying the source of problems when models fail to converge during estimation. That is, one tends to know exactly why a model fails to converge if it is done in a series of steps.

Finally, the model contrasting approach avoids the statistically questionable practice of interpreting the significance tests associated with the variance component parameters that are provided in many analytic packages. Singer (1998) and others have noted that standard error estimates of random effects (and associated *t* tests of the significance) have two key problems. First, they rely on large sample approximations;

and second, they are skewed because they cannot take on negative values. The inability to take on negative values means that variance terms are not really able to randomly vary around zero even when there is no effect. Because of these problems, the model testing approach is generally considered a better way of determining whether intercepts and slopes randomly vary across level-2 units.

In growth modeling, properly specifying the level-1 model is critical because the validity of the growth model depends heavily on one's ability to build a level-1 model that correctly specifies how time is related to the outcome (Raudenbush, 2001). Currently, however, there is a lack of guidance on how one should build the level-1 model in longitudinal research. In contrast to hierarchical regression (e.g., Cohen & Cohen, 1983), structural equation modeling (Vandenberg & Lance, 2000; Willet & Sayer, 1994), or even multilevel RCM (e.g., Bliese, 2002b; Bryk & Raudenbush, 1992; Hofmann et al., 2000), where a sequence and order for testing these models is available, we find no clear guidance for the practicing random coefficient modeler interested in exploring the complexity of longitudinal data. Therefore, we propose a sequence of tests designed to specify both the level-1 and level-2 models.

Our approach is based on the existing literature on random coefficient models (Pinheiro & Bates, 2000), regression (Cohen & Cohen, 1983), growth modeling (Chan, 1998; Ployhart & Hakel, 1998; Willett & Sayer, 1994), and our own experience in working with these models. In our approach, we start with the most basic model and subsequently add complexity. Where appropriate, we contrast the deviance of the simpler model with that from the more complex model using a log likelihood test (Pinheiro & Bates, 2000). If the more complex model fits the data significantly better than the simpler model, we will retain the more complex model.

By focusing on deviances and the log likelihood tests for random effects, we avoid problems associated with parameter estimates of random effects (Singer, 1998) while still being able to estimate a model that fits the data well. We hope that one contribution of our model testing sequence is a movement away from relying on *t* tests of significance for intercept and slope variability. The sequence of steps that we suggest for readers interested in conducting growth modeling is similar to the steps followed in the preceding illustration. There are, however, two major differences. First, we propose specific shortcuts; and second, we consider cases where the growth trajectory follows a nonlinear pattern. Table 8 provides an outline for the sequence of model comparison tests.

Level-1, Step 1

In our preceding example, we used a regression model as our baseline model. This regression model assumed that ratings of job satisfaction would be independent even though the observations came from the same individuals over time. In practice, however, it is very unlikely that longitudinal data will be independent; consequently, one typically assumes that data collected from individuals over time will display significant nonindependence. Thus, one generally begins with a random intercept model. In the random intercept model, it is often valuable to estimate the intraclass correlation coefficient (ICC) to determine the strength of the nonindependence (Bryk & Raudenbush, 1992). Details on estimating ICC values can be found in Bliese (2000; 2002b). In R, the ICC is estimated from a null lme model and the `VarCorr` command.

Table 8
Growth Model Building Sequence

Model Building Steps	*What to Interpret*
Build the Level 1 Model	
I. Estimate the intraclass correlation coefficient for the outcome	Intraclass correlation estimates
II. Determine the fixed functions for time	Significance tests of parameters
A. Linear-only model	
B. Quadratic model	
C. Higher order models	
III. Determine variability in the growth parameters	Differences in likelihood ratios
A. Variance in linear parameter	
B. Variance in quadratic parameter	
C. Variance in higher order parameters	
IV. Determine the error structure	Differences in likelihood ratios
A. Errors uncorrelated and homogeneous	
B. Errors correlated (autoregressive)	
C. Errors correlated and heterogeneous	
D. Different structures (e.g., toeplitz, band-diagonal)	
Build the Level 2 Model	
V. Add individual difference predictors of intercept and slope variability in order of theoretical importance	Significance tests of parameters

```
> null.model<-lme(JSAT~1,random=~1|SUBNUM,data=univbct,
na.action=na.omit)
> VarCorr(null.model)
. . . .
               Variance      StdDev
(Intercept)    0.4337729     0.6586144
Residual       0.4319055     0.6571952
```

The intercept or between-group variance (τ_{00}) is .434 and the residual within-group variance (σ^2) is .432. ICC is calculated as $\tau_{00}/(\tau_{00} + \sigma^2) = .50$, indicating a nontrivial degree of nonindependence. Recall that the likelihood ratio contrasting Model 1 and Model 2 was very large (308.31, to be specific), indicating a highly significant degree of nonindependence. In cases where ICC values are marginal or small (<.10), one may resort to contrasting log likelihood values between fixed versus random intercept models as we did in Steps 1 and 2 of the illustration to establish whether the degree of nonindependence is significant. In most cases, however, the ICC will be of sufficient magnitude that one can reasonably assume nonindependence and begin with a random intercept model.

Level-1, Step 2

In our illustration, we only considered the possibility that time would have a linear relationship with job satisfaction. In many cases, however, time may have a more complex relationship with the outcome (Raudenbush, 2001). For instance, it is possible that job satisfaction could have a quadratic relationship with time—job satisfaction might start high, decrease at the second time period, and then rebound at the third time period.

How one models time is important in growth modeling. If one models time such that Time 1 = 0, Time 2 = 1, Time 3 = 2, the intercept refers to the first time period or what is often called "initial status." Alternatively, if one specified Time 3 = 0, then the intercept would refer to final status. Other variations of modeling time, such as the use of orthogonal polynomials, can have a dramatic effect on the estimate of the time regression coefficient and its variance. For example, Ployhart et al. (2002) show how simply changing the coding of time can result in different significance tests for the time parameters.

Based on trend analysis in regression (e.g., Cohen & Cohen, 1983), we recommend that one start with the most parsimonious model, which in growth modeling is the linear model. This model suggests that change follows a linear trajectory for all respondents. Next, one would test whether a quadratic function captures the trajectory better than a linear trajectory. If this occurs, one would next compare the quadratic model to a cubic model, and so on, until the effects are no longer significant. From our experience, few measures have the reliability to permit an assessment of models greater than a cubic model.

In terms of tests of significance, this is a case where model contrasting is not appropriate because one has different fixed effects in the two models. That is, it would not be appropriate to estimate a baseline model with a linear time effect and a second model with both linear and quadratic time effects and then compare the log likelihood ratios between the two models (Pinheiro & Bates, 2000). Rather, one should determine the significance of each specific growth term by examining the parameter estimate, standard error, and *t* value associated with the specific term. Note that examining *t* values is always the preferred means of significance testing when one is focusing on new fixed effects (i.e., new predictors) as we are doing in this step.

Level-1, Step 3

An important characteristic of the model to this point is that it assumes that there is no significant variance in the growth trajectories among respondents (although it does allow for random intercepts). That is, the model assumes that growth for all participants follows the same trajectory. Therefore, after determining the structure of the fixed effects for time, it is important to determine whether there are individual differences in the fixed effects. This is performed by comparing the growth model with no variance in the linear, quadratic, and so on, growth parameters to models that allow variance in these parameters. Comparisons should be done in a step-by-step process. The baseline model will generally be the random intercept model, the second model will be a model where the linear slope is allowed to randomly vary, the third model will allow random variation in both the linear and quadratic terms, and so on. Likelihood contrasts between competing models will help select the most appropriate model.

We contend that it is important to include lower level growth parameters in the random effects prior to testing for higher level growth parameters. For instance, in testing for quadratic random effects, one will also want to include the linear parameter. We base this recommendation on the fact that it is necessary to control for lower level parameters in tests of higher level parameters in basic regression analyses (Cohen & Cohen, 1983).

Level-1, Step 4

The fourth step in developing the level-1 model involves assessing the error structure of the model. It is incumbent on growth modeling researchers to carefully scrutinize the error structure present in their data because significance tests may be dramatically affected. We recommend that the baseline model used to test different error structures is the model one has selected at the end of Step 3. The goal of Step 4 is to determine whether one's model fit improves by incorporating a (a) correlated and/or (b) heterogeneous error structure. One should test for different error structures by contrasting models as demonstrated in the illustrative example.

There are a variety of different error structures that include such forms as compound symmetric, autoregressive, toeplitz, and unstructured. Several researchers, including DeShon et al. (1998) and Littell, Milliken, Stroup, and Wolfinger (1996) have discussed these various structures in detail; we do not have the space to give them service here. However, it is important that researchers move beyond the simple uncorrelated and homogeneous error structure because such a structure is rarely tenable in longitudinal research. Furthermore, different error structures have different substantive interpretations and may be used to test different theoretical questions. For example, if errors are uncorrelated, the cognitions/attitudes/behaviors of individuals at two or more time periods are not related, indicating no degree of underlying stability (beyond the hypothesized trend or modeling of time). Alternatively, the presence of correlated errors suggests that the cognitions/attitudes/behaviors of individuals "carry over" to later time periods above and beyond those found with the modeling of time. For example, one might expect such carryover effects for emotions but not for more transient moods.

Note that once error structures have been determined, one should go back and recheck the estimates for the growth parameters. Substantial differences (e.g., growth parameters that are no longer significant) may suggest that the growth model should be reassessed using the newly determined error structure.

Level-2

The most difficult part of growth modeling is developing the level-1 model. In contrast, building the level-2 model is fairly straightforward. For each of the growth parameters that contain meaningful variability (e.g., intercept, linear), the appropriate individual difference predictors may be added to the model. Additional predictors may be added as theory suggests. One important note here is that the researcher should conduct this analysis one growth parameter at a time; that is, first determine the significant predictors for the intercept, then for the linear term, and so on. Furthermore, different growth parameters may be predicted by different variables. For example, the intercept may be predicted by age and gender whereas the linear term may be predicted by age.

Model Building Summary

Although these steps may seem to require a lot of analyses to build a growth model, they cannot be ignored when the desire is to build a parsimonious model that fits the data well. For example, one may argue that a simple linear growth model will be examined, yet if the data truly represent a quadratic model, the model will be misspecified. It may be parsimonious, but it will be wrong. Similarly, allowing all growth parameters to be heterogeneous when they are not adds complexity to the model and increases the possibility that one will encounter model convergence problems. Thus, we believe the model building strategy that we outlined is a valuable way to develop parsimonious models that fit the data.

Conclusion

The purpose of this article has been to provide an introduction to growth modeling aimed at readers with a background in multiple regression. In an attempt to further broaden researchers' accessibility to growth modeling, we have also emphasized the actual estimation of growth models using open source software. Finally, we have proposed a sequence of model building steps that we hope gives practical guidance to growth modeling researchers. Our goal is to make these methods more widely accessible to organizational researchers to spur the collection and reporting of data from longitudinal studies.

Although our proposed model building strategy is consistent with practices among growth modeling researchers from a variety of domains, it remains possible that refinements are possible. For instance, we tested for autocorrelation prior to heteroscedasticity; however, it is possible that this procedure should be reversed. Our order choice on this matter was entirely arbitrary. Thus, we encourage future researchers to examine model building strategies in different ways, and determine the consequences of following the wrong strategy, to better understand how RCM can provide accurate and complete descriptions of longitudinal data. The validity of our theories and the knowledge we generate rests on the quality of this information.

We conclude by noting that even researchers who are comfortable with growth modeling in SEM may want to consider adding growth modeling in RCM to their repertoire of analytic skills for several reasons. The first reason is that RCM is generally considered to be more robust in cases where missing data are present (see Schnabel, Little, & Baumert, 2000). RCM models estimate growth parameters on the available data and do not require complete data from all respondents. This can be a benefit in longitudinal field studies where respondents miss data collection points for numerous reasons. Consider, for instance, that Lance, Vandenberg, and Self (2000) used SEM to conduct a growth model with data collected at three time points. Unfortunately, however, only 104 of the 281 respondents provided the complete data needed for the SEM approach. RCM would have used the data from all 281 respondents even though individuals missed one or two reporting periods.

Second, RCM models are more flexible when modeling time. Specifically, MRC models do not require "time-structured data" (Willett & Sayer, 1994). In fact, Willett and Sayer (1994) noted that in RCM models, "each individual in the data set can possess an empirical growth record containing different numbers of waves of data with randomly assigned temporal spacing" (p. 379). Thus, there may be cases where growth modeling in RCM allows one to do analyses that would not be possible in SEM. For instance, suppose that one was interested in the relationship between number of months at work and job satisfaction. Furthermore, suppose that the three data collection times were roughly 6 months apart but that it took 3 to 4 months to complete each collection. In this case, Time 1 would have considerable variability because of differences in respondent tenure. In addition, over time, respondents would vary in data collection intervals—every 6 months for some, 4 months and 8 months for others, and so on. Although these data would be problematic to analyze in SEM, analysis would not be difficult in RCM.

Third, RCM models can easily incorporate multiple levels of nesting. Technically, this is possible in SEM as well; however, it is arguably easier to specify the model in RCM. For instance, individuals in our study provided multiple measures, so we modeled two levels of nesting—repeated measures within individuals. In fact, however, the individuals were also members of army companies nested within army battalions. A growth model that included random intercepts for both company and battalion could be specified in lme as follows:

```
>FINAL.MODEL<-lme(JSAT~TIME*AGE2,random=list
~1|BTN, ~1|COMPANY, ~TIME|SUBNUM), correlation=corAR1(),
na.action=na.omit, data=univbct)
```

The results of the `VarCorr` function applied to the final model (not shown) indicate that there is a trivial amount of battalion-level variability in job satisfaction but a nontrivial amount of variance in job satisfaction at the company level. Even syntactically highly efficient SEM programs such as Mplus (Muthen & Muthen, 1998) would require considerably more code to specify a four-level multilevel growth model of this nature. For these reasons, even researchers who have expertise in SEM may want to consider RCM models in cases where multiple levels of nesting are involved.

Of course, this is not to say that SEM does not provide some benefits over RCM when analyzing longitudinal data. First, SEM is able to account for measurement error in the manifest measures. Second, SEM allows users to test whether the factor structure (and thus interpretation of the factor) remains invariant across time (see Lance, Vandenberg, & Self, 2000). Finally, SEM is flexible in that partial invariance may be imposed over time when responses to some of the manifest measures change over time. Such concerns are important when the meaning of the construct may change across time, such as when knowledge tests are dependent on experience or performance tests are affected by training.

Given the challenges associated with longitudinal analyses, researchers are likely to benefit from being familiar with both RCM and SEM; nonetheless, in terms of application, we believe that RCM and the modeling building strategy provide an excellent and broadly accessible approach for conducting longitudinal analyses. We hope that a broader use of growth modeling in organizational research will substantially improve our understanding of organizational processes.

Notes

1. Of course, many researchers have examined correlations or mean differences between two points in time. However, most have not collected data with enough time points to allow an examination of growth over time. For our purposes, we focus on growth modeling where the research question is directed toward understanding within-person (or within-group, within-organization) change over time.

2. For instance, ordinary least squares algorithms return residual sums of squares as estimates of model deviance, whereas maximum likelihood estimation algorithms provide log likelihood values as estimates of model deviance. Thus, one could not contrast a model based on an

ordinary least squares algorithm with a model based on a maximum likelihood algorithm. To contrast two competing models, one needs to ensure that both models provide the same type of deviance term.

3. When restricted maximum likelihood estimation is used (as it is here), likelihood tests contrasting different models are appropriate only when models differ only in terms of random effects. The tests, however, are slightly conservative. Technically, likelihood tests can also be conducted when fixed effects differ (i.e., models have different predictors) if full maximum likelihood estimation is used; however, likelihood tests involving different fixed effects tend to be too liberal and are not recommended. Further details can be found in Pinheiro and Bates (2000).

4. In specifying Model 3 in R, it is technically not necessary to include the `1` in the random statement. That is, the random portion of the model could be specified as `random= ~TIME|SUBNUM` because the program infers that a random intercept is included if a random slope is specified.

5. The linear mixed-effects (lme) program "knows" that `AGE` is a level-2 attribute because the values of `AGE` are constant across individuals. In our experience, we occasionally see cases where an individual-level attribute such as `AGE` has degrees of freedom associated with the number of observations, suggesting that it is a level-1 attribute. In all cases, this has been because data manipulation errors have resulted in one or more cases where the level-2 variable has within-individual variability. For example, if age was accidentally coded as 20 (Time 1), 200 (Time 2), and 20 (Time 3) for just one individual, it would now have within-individual variability and be treated as a level-1 variable.

6. This effect could also be explained by ceiling effects for the older respondents.

References

Abelson, R. P. (1995). *Statistics as a principled argument*. Hillsdale, NJ: Lawrence Erlbaum.

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations* (pp. 349-381). San Francisco: Jossey-Bass.

Bliese, P. D. (2002a). *Multilevel modeling in R: A brief introduction to R, the multilevel package, and the NLME package*. Washington, DC: Walter Reed Army Institute of Research.

Bliese, P. D. (2002b). Multilevel random coefficient modeling in organizational research: Examples using SAS and S-PLUS. In F. Drasgow & N. Schmitt (Eds.), *Measuring and analyzing behavior in organizations: Advances in measurement and data analysis* (pp. 401-445). San Francisco: Jossey-Bass.

Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, *101*, 147-158.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models*. Newbury Park, CA: Sage.

Bryk, A. S., Raudenbush, S. W., & Congdon, R. T. (1994). *Hierarchical linear modeling with the HLM/2L and HLM/3L programs*. Chicago: Scientific Software, International.

Chan, D. (1998). The conceptualization and analysis of change over time: An integrative approach incorporating longitudinal means and covariance structures analysis (LMACS) and multiple indicator latent growth modeling (MLGM). *Organizational Research Methods*, *1*, 421-483.

Chan, D. (2002). Latent growth modeling. In F. Drasgow & N. Schmitt (Eds.), *Measuring and analyzing behavior in organizations: Advances in measurement and data analysis* (pp. 302-349). San Francisco: Jossey-Bass.

Chan, D., Ramey, S., Ramey, C., & Schmitt, N. (2000). Modeling intraindividual changes in children's social skills at home and at school: A multivariate latent growth approach to understanding between-settings differences in children's social skills development. *Multivariate Behavioral Research*, *35*, 365-396.

Chan, D., & Schmitt, N. (2000). Interindividual differences in intraindividual changes in proactivty during organizational entry: A latent growth modeling approach to understanding newcomer adaptation. *Journal of Applied Psychology*, *85*, 190-210.

Chou, C., 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.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Deadrick, D. L., Bennett, N., & Russell, C. J. (1997). Using hierarchical linear modeling to examine dynamic performance criteria over time. *Journal of Management*, *23*, 745-757.

DeShon, R. P., Ployhart, R. E., & Sacco, J. M. (1998). The estimation of reliability in longitudinal models. *International Journal of Behavior and Development*, *22*, 493-515.

Garst, H., Frese, M., & Molenaar, P. C. M. (2000). The temporal factor of change in stressor-strain relationships: A growth curve model on a longitudinal study in East Germany. *Journal of Applied Psychology*, *85*, 417-438.

Gonzales, R., & Griffin, D. (2001). Testing parameters in structural equation modeling: Every “one” matters. *Psychological Methods*, *6*, 258-269.

Gottman, J. M. (1995). *The analysis of change*. Mahwah, NJ: Lawrence Erlbaum.

Hand, D., & Crowder, M. (1996). *Practical longitudinal data analysis*. London: Chapman & Hall.

Hofmann, D. A. (1997). An introduction to the logic and rationale of hierarchical linear models. *Journal of Management*, *23*, 723-744.

Hofmann, D. A., & Gavin, M. (1998). Centering decisions in hierarchical linear models: Theoretical and methodological implications for research in organizations. *Journal of Management*, *24*, 623-641.

Hofmann, D. A., Griffin, M. A., & Gavin, M. B. (2000). The application of hierarchical linear modeling to organizational research. In K. J. Klein & S. W. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations* (pp. 467-511). San Francisco: Jossey-Bass.

Hofmann, D. A., Jacobs, R., & Baratta, J. E. (1993). Dynamic criteria and the measurement of change. *Journal of Applied Psychology*, *78*, 194-204.

Kenny, D. A., & Judd, C. M. (1986). Consequences of violating the independence assumption in analysis of variance. *Psychological Bulletin*, *99*, 422-431.

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

Lance, C. E., Meade, A. W., & Williamson, G. M. (2000). We should measure change—and here’s how. In G. M. Williamson & D. R. Shaffer (Eds.), *Physical illness and depression in older adults: A handbook of theory, research, and practice* (pp. 201-235). New York: Kluwer Academic/Plenum.

Lance, C. E., Vandenberg, R. J., & Self, R. M. (2000). Latent growth models of individual change: The case of newcomer adjustment. *Organizational Behavior & Human Decision Processes*, *83*, 107-140.

Littell, R. C., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (1996). *SAS system for mixed models*. Cary, NC: SAS Institute.

Little, R. J. A., & Schlenker, N. (1995). Missing data. In G. Arminger, C. C. Clogg, & M. E. Sobel (Eds.), *Handbook of statistical modeling for the social and behavioral sciences* (pp. 39-75). New York: Plenum.

Muthen, L. K., & Muthen, B. O. (1998). *Mplus user’s guide*. Los Angeles: Muthen & Muthen.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer-Verlag.

Ployhart, R. E., & Hakel, M. D. (1998). The substantive nature of performance variability: Predicting interindividual differences in intraindividual performance. *Personnel Psychology*, *51*, 859-901.

Ployhart, R. E., Holtz, B. C., & Bliese, P. D. (2002). Longitudinal data analysis: Applications of random coefficient modeling to leadership research. *Leadership Quarterly, 13,* 455-486.

Raudenbush, S. W. (2001). Comparing personal trajectories and drawing causal inferences from longitudinal data. *Annual Review of Psychology*, *52*, 501-525.

Schnabel, K. U., Little, T. D., & Baumert, J. (2000). Modeling longitudinal and multilevel data. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* (pp. 9-13). Mahwah, NJ: Lawrence Erlbaum.

Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, *24*, 323-355.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *2*, 4-69.

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

Paul D. Bliese is a research psychologist at the Walter Reed Army Institute of Research. His interests include multilevel and longitudinal data analysis techniques, occupational stress research, and leadership research.

Robert E. Ployhart is an assistant professor at George Mason University. His research interests include personnel selection, modeling latent job performance, longitudinal statistical methods, and applied measurement issues.