Social and Personality Psychology Compass 3/6 (2009): 979–991, 10.1111/j.1751-9004.2009.00224.x

The ABC's of LGM: An Introductory Guide to Latent Variable Growth Curve Modeling

Terry E. Duncan* and Susan C. Duncan
Oregon Research Institute

Abstract

In recent years, we have witnessed an increase in the complexity of theoretical models that attempt to explain behavior from both contextual and developmental perspectives. This increase in the complexity of our theoretical propositions regarding behavior parallels recent methodological advances for the analysis of change. These new analysis techniques have fundamentally altered how we conceptualize and study change. Researchers have begun to identify larger frameworks to integrate our knowledge regarding the analysis of change. One such framework is latent growth modeling, perhaps the most important and influential statistical revolution to have recently occurred in the social and behavioral sciences. This study presents a basic introduction to a latent growth modeling approach for analyzing repeated measures data. Included is the specification and interpretation of the growth factors, primary extensions such as the analysis of growth in multiple populations, and structural models including both precursors of growth, and subsequent outcomes hypothesized to be influenced by the growth functions.

Over the past several decades, a majority of research on individual differences in social and behavioral development has focused on how the combination of personality, socio-cultural factors and broader societal factors contribute to social relationships throughout one's life. Understanding developmental processes in conjunction with opportunities, constraints, and demands of the salient contexts that individuals experience is important for the identification of exactly how such experiences contribute to, or impede, optimal developmental outcomes.

The representation and measurement of change is both an interest and a fundamental concern to almost all scientific disciplines. Researchers interested in studying change in behavior over time must use a longitudinal research design where the same individuals are repeatedly assessed. Such a design provides intra-individual as well as inter-individual variation, allowing the researcher to study change within individuals over time. Longitudinal designs, whether true, quasi-, or nonexperimental, pose several unique problems because they involve variables with correlated observations. No single statistical procedure exists for the analysis of longitudinal data because different research questions dictate different data structures, and thus different statistical methods and models. The search for the best methods to address complex issues in behavior change has been a persistent theme of recent developmental research (e.g. Collins & Horn, 1991; Collins & Sayer, 2001; Duncan, Duncan, & Strycker, 2006; Preacher, Wichman, MacCallum, & Briggs, 2008), and has prompted a shift in analytic strategies.

Traditional approaches to studying change have used ANOVA and multiple regression. These approaches analyze only mean changes and treat differences among individual subjects as error variance. However, this 'error variance' may contain valuable information about change. Rather than focusing on homogeneous populations and inter-individual variability, analysts are turning to new methods to explore both inter- and intra-individual variability and heterogeneity in growth trajectories of various human behaviors. Recently, a number of different analytical approaches to modeling hierarchical and longitudinal data have extended researchers' ability to describe individual differences and the nature of change over time (e.g. random-effects ANOVA, random coefficient modeling, multilevel modeling, and hierarchical linear modeling). In these approaches, individual differences in growth over time are captured by random coefficients. However, within these approaches, statistical modeling has been largely limited to a single response variable. As such, they do not fully accommodate the complexity and analytical needs of current developmental theories (Muthén & Curran, 1997).

The need to answer increasingly complex substantive questions inspires the development of new statistical methods. These new analysis techniques have fundamentally altered how we conceptualize and study change. Methodology for the study of change has matured sufficiently that researchers are beginning to identify larger frameworks in which to integrate knowledge. One such framework is latent growth modeling (LGM). The LGM makes available to a wide audience of researchers an analytical framework for a variety of analyses of growth and developmental processes (e.g. Burt, Mcque, Carter, & Iacono, 2007; Byrne, Lam, & Fielding, 2008; Neyer & Lehnart, 2007).

The latent growth modeling (LGM) framework allows more flexibility to fully examine questions often posited by developmental and behavioral researchers. Once the random coefficient model has been placed within the latent variable framework, many general forms of longitudinal analyses can be studied. Building upon traditional longitudinal models, methodologists have extended the latent variable framework to accommodate longitudinal models that include multivariate or higher-order specifications, multiple populations, the accelerated collection of longitudinal data, multilevel or hierarchical structures, and complex relations including mediation, moderation, and reciprocal causation.

A Developmental Framework

An appropriate developmental model is one that not only describes a single individual's developmental trajectory, but also captures individual differences in these trajectories over time. If, for example, trajectories produced a collection of straight lines for a sample of individuals, the developmental model should reflect individual differences in the slopes and intercepts of those lines. Another critical attribute of the developmental model is the ability to study predictors of individual differences to answer questions about which variables exert important effects on the rate of development. At the same time, the model should be able to capture the vital group statistics so that researchers can study development at the group level. One methodology with all these attributes is the latent growth model or LGM.

Although strongly resembling a classic confirmatory factor analysis, the LGM's growth factors are actually interpreted as individual differences in the growth trajectories of attributes over time (McArdle, 1988). For example, Figure 1 presents growth trends of marital discord in a subset of 10 individuals over four time points. In this collection of individual growth trajectories are two potentially interesting attributes of growth: level of marital discord and the rate of change in marital discord over time. The figure shows differing levels and rates of change. For example, some individuals start high and decrease over time, some start low and increase over time, and others remain fairly stable across time points.

In simple, straight-line growth models' level and rate of change are interpreted as the slope and intercept, respectively. Repeated-measures (RM) polynomial analysis

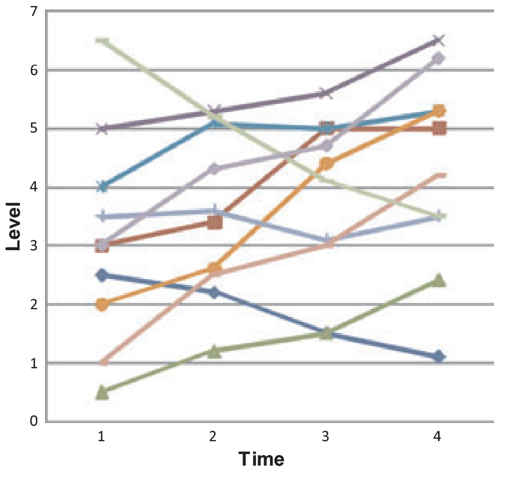

Figure 1 Representation of a collection of individual growth trajectories.

of variance (ANOVA) models are actually special cases of LGMs in which only the factor means (corresponding to group level information) are of interest (Meredith & Tisak, 1990). Variances are considered nuisance parameters. In contrast, a fully expanded latent growth analysis takes into account both factor means and variances, which correspond to individual differences. Heuristically, growth curve methodology can be thought of as consisting of two stages. In the first stage, a regression curve, not necessarily linear, is fit to the RM of each individual in the sample. In the second stage, the parameters for an individual's curve become the focus of the analysis rather than the original measures.

The modeling task, therefore, involves identifying an appropriate growth curve form which will accurately and parsimoniously describe individual development. If, for example, the trajectories were well described by a collection of straight lines for a sample of individuals, the developmental model should reflect individual differences in the slopes and intercepts of those lines. Beyond describing and summarizing growth at the group and individual level, however, the model can also be used to study predictors of individual differences to answer questions about which variables exert important effects on the rate of development. Researchers such as Meredith and Tisak (1990) and McArdle (1988) extended the basic model to permit the testing procedures and standards of estimation found in current structural equation modeling (SEM) programs. Thus, statistical tests of overall model fit and the significance of individual model parameters are available. Because LGM is carried out using SEM methodology, it shares many of the same strengths. These include an ability to test the adequacy of the hypothesized growth form, to incorporate both fixed and time-varying covariates, to correct for error in the observed indicators, to incorporate growth on several constructs simultaneously, and to develop from the data a common developmental trajectory, thus ruling out cohort effects.

Specification of the Latent Growth Model

The simplest LGM involves one variable (e.g. marital discord) measured the same way at two time points. Two time points are not ideal for studying development or for using LGM because the collection of individual trajectories are limited to a collection of straight lines. While two observations of marital discord provide information about the *amount* of change, they address other research questions quite poorly (Rogosa, Brandt, & Zimowski, 1982). For example, two temporally separated observations allow estimation of the *amount* of change, but it is impossible to study the *shape* of the developmental trajectory or the rate of individual *change* in marital discord. Two-wave panel designs preclude testing theories related to the shape of development and, as such, are appropriate only if the growth process is considered irrelevant or is known to be linear.

In general, developmental studies should include three or more assessment points. With multi-wave data (more than two observations), the validity of the straight-line growth trajectory model can be evaluated (e.g. tests for nonlinearity can be performed). In addition, the precision of parameter estimates will tend to increase along with the number of observations for each individual.

To introduce the LGM, a model with three time points (representing RM of marital discord) is presented in Figure 2.

Intercept

In Figure 2, the first latent factor (F_1) is labeled 'Intercept'. The intercept is a constant for any individual across time, hence the fixed values of 1 for factor loadings on the RM. The intercept in this model for a given individual has the same meaning as the intercept of a straight line on a two-dimensional coordinate system: it is the point where the line 'intercepts' the vertical axis. The intercept factor presents information in the sample about the mean (M_i) and variance (D_i) of the collection of intercepts that characterize each individual's marital discord growth curve.

Slope

The second factor (F_2), labeled 'Slope', represents the slope of an individual's marital discord trajectory. In this case, it is the slope of the straight line determined by the three RM. The slope factor has a mean (M_s) and variance (D_s) across the whole sample which, like the intercept mean and variance, can be estimated from the data. The two factors, Slope and Intercept, are allowed to covary, Ris, which is represented by the double-headed arrow between the factors. The error variance terms are labeled as E_1, E_2, and E_3. The error variances affect the interpretation of the model parameters by correcting the measured variances for random error. For example, the variance of

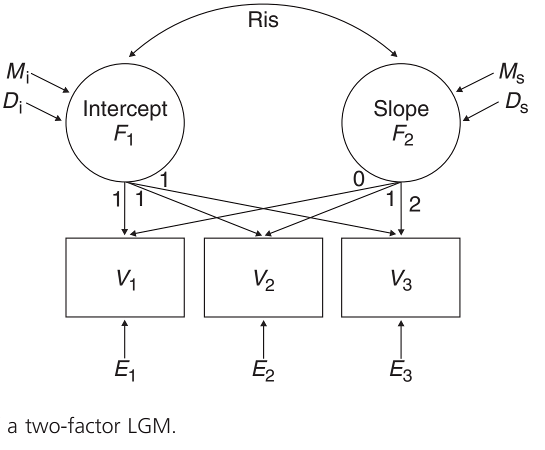

Figure 2 Representation of a two-factor LGM.

the slope factor, F_2, would be the variance of change in marital discord corrected for random error, and the variance of the intercept factor (F_1) is the true score variance of the intercept of marital discord. By including the error variance terms, the model parameters retain their same basic interpretations but are now corrected for random error.

To identify this model, two slope loadings must be fixed at two different values. Although the choice of loadings is somewhat arbitrary, the intercept factor is bound to the time scale. Shifting the loadings on the slope factor alters the scale of time, which affects the interpretation of the intercept factor mean and variance.

The slope factor's mean and variance differ from the intercept factor's mean and variance in that changing the fixed loadings, and thereby changing the time scale, rescales the slope factor's mean and variance, in this case by constants. Rescaling by constants does not change the fundamental meaning or affect significance tests of the parameters. It also does not affect the correlations between the slope factor and other predictors in the model.

In estimating the LGM, researchers must ensure that the model is identified. If the model has fewer parameters to estimate than pieces of information in the data, the model is identified, meaning the model provides a unique fit to the data. If the model has more parameters to estimate than pieces of information in the data it cannot be uniquely estimated, and is therefore not 'identified'. Having an unidentified model implies that it is impossible to compute a reasonable estimate for one or more of the model's parameters and that the model cannot generally be relied on.

In the model depicted in Figure 2, there are eight parameters to be estimated (the Intercept's mean and variance, the Slope's mean and variance, the covariance between the Intercept and Slope, and three error variances) and there are nine pieces of known information (3 variances, 3 means, and 3 covariances) from which to estimate the model. Note that the factor loadings for the Intercept and Slope are not included as estimated parameters in the model as they have been fixed to user-specified values. Given that the model is identified, the difference between the number of pieces of information in the data and the number of parameters to be estimated in the model should equal the number of degrees of freedom in the model test. If the number of pieces of information in the data and the number of parameters to be estimated in the model are the same, then the model is saturated or 'just identified'. Unfortunately, there is no way to test or confirm the plausibility of a saturated or 'just identified' model because saturated models will always provide a perfect fit to the data. However, despite the lack of an overall test of the hypothesized model in this situation, tests of individual parameters can provide the researcher with some secondary evidence of the adequacy of the proposed model describing relationships in the data.

Interpreting the Growth Factors

As indicated previously, the choice of loadings can affect the interpretation of both the intercept and slope factors. For example, in the model depicted in Figure 2, growth in marital discord might be assessed over three approximately equally spaced occasions (2 units of time) where there is a 1-unit increase in mean levels of marital discord (e.g. $M_{T3} - M_{T1} = 2$). The factor loadings, representing values of the time metric, are fixed at values that represent polynomial contrasts used to identify the scale of the F variables. To further explain the nature of the contrasts used, consider the equation for a linear growth curve for a single individual with three data points:

$$y_i = b_1 + t_i b_2 + e_i, \tag{1}$$

where b_1 is the intercept, b_2 is the slope (amount of vertical increase per unit of horizontal run of the growth curve), t_i is the ith value of time, e_i represents the time-specific errors of prediction (set at zero in this example), and i is the value of time. For this hypothetical individual, therefore, the set of equations is:

$$y_1 = b_1 + 0b_2 + e_1 \tag{2}$$

$$y_2 = b_1 + 1b_2 + e_2, \text{and} \tag{3}$$

$$y_3 = b_1 + 2b_2 + e_3. \tag{4}$$

Relating these equations back to Figure 2, for any given individual, b_1 corresponds to the intercept factor score (F_1), b_2 corresponds to the slope factor score (F_2), and e_i corresponds to the time-specific errors of prediction. If $F_1 = 4$, $F_2 = 1$ and $E_t = 0$, $t = 1$, 2, and 3, then the model for this individual implies the following trajectory:

$$y_1 = 4 + 0(1) + 0 = 4 \tag{5}$$

$$y_2 = 4 + 1(1) + 0 = 5, \text{and} \tag{6}$$

$$y_3 = 4 + 2(1) + 0 = 6, \tag{7}$$

where $t_1 = 0$ at time 1 simply starts the curve at this point by rescaling the intercept factor to represent initial status, $t_2 = 1$ at time 2 indicates that from time 1 to time 2 there is one unit of change, and $t_3 = 2$ at time 3 indicates that from time 2 to time 3 there is one unit of change. Thus, t_i describes a linear relation of change in terms of linear differences from initial status at time 1.

Because the factor loadings represent an increasing trend, the slope factor is interpreted as positive growth where higher scores on the factor (i.e. factor scores) represent more positive, or greater increases, in marital discord. A positive correlation between the intercept and slope factors would indicate that those individuals with greater values at T_1 tended to have higher slope scores or more positive growth in marital discord over time. A negative correlation between the intercept and slope factors would suggest that individuals with greater values at T_1 tended to have lower slope scores or less positive growth in marital discord over time.

Suppose that the model in Figure 2 depicted factor loadings set at values of −2, −1, and 0. With these factor loadings, the mean for the intercept factor would no longer represent initial status at T_1, but would now be interpreted as status at T_3. (Note: the variable mean, associated with the time point where the factor loading on the slope factor is fixed at 0, defines the intercept factor mean). The loadings still represent an increasing trend, thus the slope factor mean is again positive, and higher scores on the factor represent greater increases in marital discord. Because the interpretation of the intercept factor has changed (from T_1 to T_3), the correlation between the intercept and slope factors has also changed as would the correlation between the intercept and any other covariate.

Consider now the same model depicted in Figure 2 with factor loadings fixed at values of 0, −1, and −2. Note that the zero loading for the first variable of the slope factor allows for the intercept factor to once again be interpreted as initial status at T_1. However, because the loadings on the slope factor represent a negative trend from T_1, the slope mean is negative ($M_s = -1$) and higher scores on the slope factor now represent more negative or greater decreases in marital discord. Compared to the model with loadings of 0, 1, and 2, the correlation between the intercept and slope would be of the same magnitude, but of opposite sign. The LGM approach offers the researcher a wide range of modeling possibilities. The choice of factor loadings would also allow the researcher to express the intercept factor as representing a point prior to or following the time frame in which the data were collected. However, extrapolation beyond the range of scores used in the growth trajectory estimation is generally not advisable as it assumes that the trajectory continues to be linear (in our example depicted in Figure 2) when in fact it may be curvilinear.

The ability to center growth around various time points provides great latitude when specifying conditional LGMs that involve various predictors of growth. For example, in a longitudinal study of marital discord over three time points, the researcher can alternately specify predictors of initial status (time 1), average status at time 2, or the ending or terminal status (time 3), in addition to predictors of the growth trend from T_1 to T_3.

Representing the Growth form Over Time

Thus far in our examples, the factor loadings have only carried information regarding the *amount* (e.g. linear change) of change over time. With three or more time points, however, the factor loadings also provide information about the *shape* of growth over time. Three or more time points provide an opportunity to test for nonlinear trajectories. Perhaps the most familiar approach to nonlinear trajectories is the use of polynomials. The inclusion of quadratic or cubic effects is easily accomplished by including an additional factor or two. The factor loadings can then be fixed to represent a quadratic or cubic function of the observed time metric. When polynomial contrasts are used to identify the scale of the *F* variables (e.g. 1,1,1; 0,1,2; 0,1,4), the selected contrasts rescale the observed variables to represent intercept or initial status, linear, and quadratic factors (Figure 3). Polynomials are only one possibility for modeling a three-factor LGM. The use of orthogonal polynomials with loadings of 1,1,1; −1,0,1; and 1,−2,1 for the intercept, linear, and quadratic factors, respectively, would rescale the intercept mean to represent a constant or the average level over time (e.g. the mean at T_2). This is the usual interpretation of the initial growth factor in RM ANOVA. Therefore, the linear and quadratic functions would not be centered around the initial time point, but around the constant or average level over time, influencing their variances and covariances with other variables in the model. RM ANOVAs are, in fact, special cases of latent growth curve models (Meredith & Tisak, 1990), and the different methods share many similarities.

The LGM presented earlier can be easily expressed in terms of the general linear modeling ANOVA techniques typically used in RM analyses. In a simple RM analysis, all dependent variables represent different measurements of the same variable for different values or levels of a within-subjects factor. The within-subjects factor distinguishes measurements made on the same individual, rather than between different individuals. In models capturing growth or development over time, the within-subject factor is the time of measurement. Between-subjects factors and covariates can also be included in the model, just as in models not involving RM data.

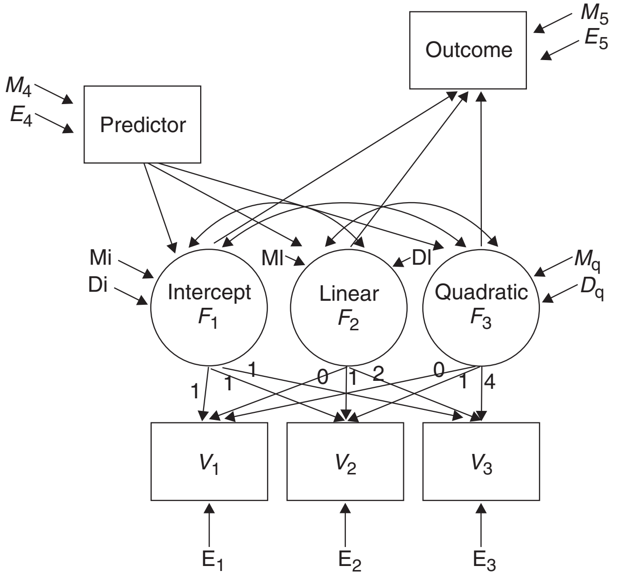

Figure 3 Representation of the conditional LGM.

This fixed effects framework of the ANOVA model can be extended to include random effects. Within the random-effects approach, each of the levels in the data structure (e.g. repeated observations within individuals) is represented by its own submodel, which represents the structural relations and variability occurring at that level. Unlike fixed-effects models, such as the RM ANOVA, the random-effects ANOVA makes use of the within-cluster differences in parameter estimates, treating these differences as a meaningful source of variance rather than as within-group error or a nuisance (Kreft, 1994). The simplest random-effects ANOVA model is fully unconditional (i.e. no predictor variables are specified) and allocates variation in an outcome measure across the different levels. Conditional models include predictors and a general structural model at each level.

However, the LGM provides greater flexibility in the measurement of change compared to the more traditional ANOVA, the most notable being LGM's ability to approximate random changes in measurement error. Another is LGM's ability to use variables simultaneously as both independent and dependent variables in the model, allowing for complex representations of growth and correlates of change, an approach that both the fixed and random-effects ANOVA are unable to accommodate.

The characteristics of the developmental trajectories that comprise the sample determine not only the magnitude of the estimated model parameters, but also the number of factors adequate to describe the data. Although it is possible to add factors until a satisfactory data fit is obtained, LGM is most powerful when a small number of factors describe the data. Questions about the number of factors needed for a given growth form, or how well a small number of factors approximates a particular nonlinear trajectory, are covered in detail in Tucker (1958), Tisak and Meredith (1990), and Burchinal and Appelbaum (1991).

Including Covariates of Change: the Conditional LGM

Once the shape of growth is determined in the LGM, the parameters for an individual's curve can be used to study predictors of individual differences and answer questions about which variables affect the rate of development and how development influences subsequent behaviors. Analysis of covariance allows for tests of continuous predictors of change and change as a predictor, but not for the *simultaneous* inclusion of change as both an independent and dependent variable. Compared to more traditional approaches, a major advantage of the LGM is the ability to use variables simultaneously as independent and dependent variables in the same model, allowing for complex representations of growth and correlates of change. Figure 3 represents a growth model that includes both exogenous predictors of growth, and growth forms impacting subsequent outcomes. Here, the growth factors are specified as antecedents and sequelae in the same growth model.

Modeling Between-Subjects Effects: Multiple Sample Growth Models

In addition to modeling growth for a single population, LGMs can also analyze behavior change among multiple groups (e.g. treatment and control conditions, age, gender). Just as RM ANOVA models can be considered special cases of the general LGM, so too can between-subjects RM ANOVAs be considered a special case of the multiple-sample LGM approach.

In the typical LGM, data are assumed to represent a random sample of observations from a single population. However, in practice, this assumption is not always reasonable. In an experimental design, it is often appropriate and necessary to test for the existence of multiple populations (e.g. treatment and control groups) and multiple developmental pathways across condition rather than a single underlying trajectory for all. Many studies involving multiple populations have either examined separate models for each group and compared the results, or have collapsed across different populations (e.g. gender, ethnicity), which may mask potential group differences important to the study of change. Unfortunately, such procedures do not allow a test of whether there are multiple developmental pathways or a common developmental trend across groups.

Developmental hypotheses involving multiple populations can be evaluated simultaneously provided that data on the same variables over the same developmental period are available in the multiple samples. Multiple-sample LGM has the potential to test for similarities and differences in developmental processes across different populations, including differences in levels of behaviors, developmental trajectories, rates of change, and effects of predictors and outcomes. Thus, when data from multiple populations are available, a multiple-sample LGM is likely to be most appropriate and advantageous.

Conventional longitudinal multiple population LGMs specify a common growth model in multiple groups, testing for equality of parameters across the different populations. However, an alternative approach (Muthén & Curran, 1997), now widely used in the analysis of treatment and prevention data, introduces an additional growth factor, the 'added growth' factor, for one population (Figure 4). Whereas the first two factors (i.e. intercept and slope) are the same in both groups, the 'added growth' factor, specified in one group, represents the incremental/decremental growth that is specific to that group.

In Figure 4, the added growth factor is specified to capture linear differences between the two groups. In this case, the linear slope factor captures normative growth that is common to both groups. For all factors except initial status, one may specify an added growth factor. For example, one group may have a linear and quadratic growth factor beyond the intercept or initial status factor, and the remaining group may have added factors for both the linear and quadratic trajectories. The multiple-sample LGM framework affords a powerful design for detecting differences in program effectiveness by attributes such as subgroup membership. Examples can be found in Duncan, Duncan, and Alpert

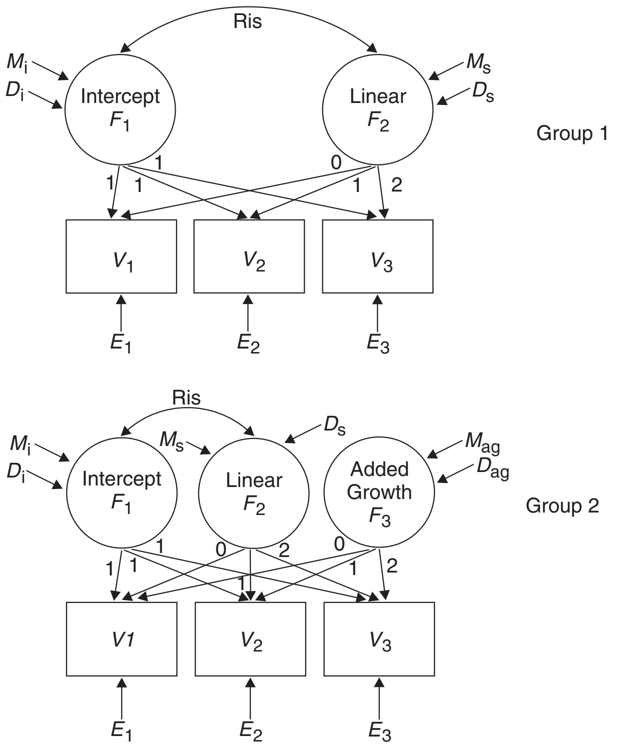

Figure 4 Representation of the added growth LGM.

(1997), Jo and Muthén (2001), McArdle, Hamagami, Elias, and Robbins (1991), Muthén and Curran (1997), and Tisak and Tisak (1996).

Discussion

LGM methodology provides advantages to researchers interested in studying change and development in behaviors over time. The LGM describes a single individual's developmental trajectory and captures individual differences in these trajectories over time. It is able to study predictors of individual differences to answer questions about which variables exert important effects on the rate of development. At the same time, the LGM is able to capture important group statistics in a way that allows the researcher to study development at the group level.

Given more than two assessment points, LGMs are able to test both linear and nonlinear growth functions. With the judicious choice of factor loadings, the general two-factor LGM can be parameterized so that the intercept and slope factors have straightforward interpretations as initial status and change, respectively. Using this parameterization, investigators can study predictors of change separately from correlates of initial status. When appropriate, the LGM can also include more than two factors to capture developmental trends through the use of specified growth functions and additional factors (e.g. quadratic, cubic growth).

Compared to more traditional approaches, LGM is able to use variables simultaneously as both independent and dependent variables. Both static and time-varying variables can be included as predictors and outcomes of growth functions, thus allowing researchers to address questions related to the antecedents and consequences of development.

In addition to the basic LGM approaches presented in this study, there are a number of additional noteworthy extensions. These include the accelerated or cohort sequential design, multivariate extensions, interrupted time series (ITS) models, and hierarchically nested designs.

The accelerated design consists of linking adjacent segments of temporally overlapping repeated measurements of independent age cohorts to determine the existence of a common developmental trend or growth curve (Duncan, Duncan, & Hops, 1996; Raudenbush & Chan, 1992). This technique allows the researcher to determine whether trends observed in the RM are corroborated within short time periods for each age cohort. The cohort-sequential LGM design approximates a long-term longitudinal study by conducting several short-term longitudinal studies of different age cohorts simultaneously.

While development in a single behavior is often of interest, it is sometimes important to examine a number of behaviors simultaneously (e.g. various family behaviors) to determine the extent to which their development is interrelated. The univariate longitudinal model is actually a special case of the general multivariate growth curve model. With multivariate LGMs, it is possible to determine whether development in one behavior covaries with other behaviors. Multivariate LGMs can include associative models or *factor-of-curves* and *curve-of-factors* models (McArdle, 1988).

Piecewise LGMs subdivide a series of repeated measurements into meaningful segments, summarize growth in each segment, and provide a means of examining (i) whether rates of change differ as a function of growth period, (ii) whether individual variability in rates of change differ between periods, and (iii) important predictors of change unique to a particular developmental period. This segmenting approach has its corollary with the time series prevention and treatment literature in which the analysis of interest is usually an ITS and the goal of the analysis is to evaluate the intervention's effect. Within LGM, ITS LGM analyses can evaluate naturally occurring interruptions (e.g. middle- to high-school transition). ITS LGM allows examination of various developmental transitions within each stage and adjustments between each transition (Duncan & Duncan, 2004). ITS designs allow for assessments of the onset (i.e. abrupt, gradual) and duration (i.e. permanent, temporary) of change in response to the transition.

How individuals and social factors operate independently and interactively to shape development can only be adequately studied in the context of longitudinal and hierarchical research. In many studies, intact groups (e.g. communities, families, marital dyads) rather than individuals are assessed. Within these intact groups, it is assumed that the responses of individuals in these groups will be similar by virtue of the experiences they share in those settings (Raudenbush, 1995). Hierarchical/multilevel LGM techniques are available for the analysis of nested or hierarchical data. Just as ANOVA and multiple regression techniques can be considered special cases of the general SEM (Hoyle, 1995), so too can hierarchical linear models be viewed as special cases of the general multilevel covariance structure model. The multilevel covariance analysis LGM approach is available for multilevel growth modeling and is particularly beneficial because it can estimate and test relationships among latent as well as observed variables.

The search for the best methods to address complex issues in studies of behavioral change has been a major focus of recent developmental research (e.g. Collins & Sayer, 2001; Duncan, Duncan, & Strycker, 2006). It has prompted a movement away from analyses of homogeneous populations and inter-individual variability to a focus on both inter- and intra-individual variability and heterogeneity in growth trajectories over time.

The ultimate purpose of these new statistical methods is to provide a means for drawing conclusions from the data. At their best, these statistical methods enjoy a symbiotic relationship with substantive research. The need to answer substantive questions inspires the development of new statistical methods, and subsequently, these methods prompt substantive researchers to view their data in new ways and to posit new substantive questions. Just as there are a plethora of substantive questions posed by researchers, so too are there a broad and varied assortment of newly developed statistical methods available to answer these questions. Although these techniques have fundamentally changed how researchers think of and study change, the application of these methods in practice for many fields has been implemented slowly. The LGM approach is extremely beneficial and flexible. Perhaps more than any other current statistical model in use by social and behavioral researchers, it provides a highly comprehensive and varied approach to the analysis of growth and behavioral processes.

Acknowledgments

This research was supported by Grant AA11510 from the National Institute on Alcohol Abuse and Alcoholism. Partial support in preparing this manuscript was provided by Grant DA018760 from the National Institute on Drug Abuse. The authors would like to thank two anonymous reviewers for their insightful comments on an earlier draft of this manuscript.

Short Biography

Terry E. Duncan is currently a Senior Research Scientist at Oregon Research Institute in Eugene, Oregon. He earned a PhD in physical education in 1989 from the University of Oregon. He is an active researcher and has been the Principal or Co-Investigator on 16 NIH-funded studies. His expertise is in statistical methods for longitudinal and multilevel designs, structural equation modeling, the etiology of substance use and development, youth problem behavior, and exercise and health behavior.

Susan C. Duncan received her PhD in physical education (sport and exercise psychology) in 1992 from the University of Oregon. She is currently a Senior Research Scientist at Oregon Research Institute in Eugene, Oregon. Duncan has been the Principal or Co-Investigator on 13 NIH-funded studies. She is a productive scientist and author, known for her research focusing on substantive, statistical, and methodological issues related to youth health risk (e.g. substance use) and health promoting (e.g. physical activity) behaviors.

Endnote

* Correspondence address: Terry E. Duncan, PhD, Oregon Research Institute, 1715 Franklin Blvd., Eugene, OR 97403, USA. Email: terryd@ori.org

References

Burchinal, M., & Appelbaum, M. I. (1991). Estimating individual developmental functions: Methods and their assumptions. *Child Development*, **62**, 23–43.

Burt, S. A., Mcque, M., Carter, L. A., & Iacono, W. G. (2007). The different origins of stability and change in antisocial personality disorder symptoms. *Psychological Medicine*, **37**(1), 27–38.

Byrne, B. M., Lam, W. W. T., & Fielding, R. (2008). Measuring patterns of change in personality assessments: An annotated application of latent growth curve modeling. *Journal of Personality Assessment*, **90**(6), 536–546.

Collins, L., & Horn, J. L. (1991). *Best Methods for the Analysis of Change*. Washington, DC: APA Press.
Collins, L., & Sayer, A. (2001). *New Methods for the Analysis of Change*. Washington, DC: American Psychological Association.
Duncan, T. E., & Duncan, S. C. (2004). A latent growth curve modeling approach to pooled interrupted time series analyses. *Journal of Psychopathology and Behavioral Assessment*, **26**, 271–278.
Duncan, T. E., Duncan, S. C., & Alpert, A. (1997). Multilevel covariance structure analysis of family substance use across samples and ethnicities. *Journal of Gender, Culture, and Health*, **2**(4), 271–286.
Duncan, S. C., Duncan, T. E., & Hops, H. (1996). Analysis of longitudinal alcohol use data within accelerated longitudinal designs. *Psychological Measurement*, **1**, 236–248.
Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). *An Introduction to Latent Variable Growth Curve Modeling: Concepts, Issues, and Applications* (2nd edn). Mahwah, NJ: Lawrence Erlbaum Associates.
Hoyle, R. (1995). The structural equation modeling approach: Basic concepts and fundamental issues. In R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, Issues and Applications* (pp. 1–15). Thousand Oaks, CA: Sage.
Jo, B., & Muthén, B. (2001). Modeling of intervention effects with noncompliance: A latent variable approach for randomized trials. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New Developments and Techniques in Structural Equation Modeling* (pp. 57–87). Mahwah, NJ: Lawrence Erlbaum Associates.
Kreft, I. G. (1994). Multilevel models for hierarchically nested data: Potential applications in substance abuse prevention research. In L. Collins & L. Seitz (Eds.), *Advances in data analysis for prevention intervention research* (NIDA Research Monograph, No. 142; pp. 140–183). Washington DC: Government Printing Office.
McArdle, J. J. (1988). Dynamic but structural equation modeling of repeated measures data. In R. B. Cattell & J. Nesselroade (Eds.), *Handbook of Multivariate Experimental Psychology* (2nd edn, pp. 561–614). New York: Plenum Press.
McArdle, J. J., Hamagami, F., Elias, M. F., & Robbins, M. A. (1991). Structural modeling of mixed longitudinal and cross-sectional data. *Experimental Aging Research*, **17**, 29–52.
Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, **55**, 107–122.
Muthén, B. O., & Curran, P. J. (1997). General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods*, **2**, 371–402.
Neyer, F. J., & Lehnart, J. (2007). Relationships matter in personality development: Evidence from an 8-year Longitudinal study across young adulthood. *Journal of Personality*, **75**(3), 535–568.
Preacher, K. J., Wichman, A. L., MacCallum, R. C., & Briggs, N. E. (2008). Latent growth curve modeling. *Quantitative Applications in the Social Sciences*, **157**, Thousand Oakes, CA: Sage Publications, Inc.
Raudenbush, S. W. (1995). Statistical models for studying the effects of social context on individual development. In J. Gottman (Ed.), *The Analysis of Change* (pp. 165–201). Hillsdale, NJ: Lawrence Erlbaum.
Raudenbush, S. W., & Chan, W. (1992). Growth curve analysis in accelerated longitudinal designs. *Journal of Research in Crime and Delinquency*, **29**, 387–411.
Rogosa, D. R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measure of change. *Psychological Bulletin*, **92**, 726–748.
Tisak, J., & Meredith, W. (1990). Descriptive and associative developmental models. In A. von Eye (Ed.), *Statistical Methods in Developmental Research* (Vol. 2, pp. 387–406). San Diego, CA: Academic Press.
Tisak, J., & Tisak, M. S. (1996). Longitudinal models of reliability and validity: A latent curve approach. *Applied Psychological Measurement*, **20**, 275–288.
Tucker, L. R. (1958). Determination of parameters of a functional relation by factor analysis. *Psychometrika*, **23**, 19–23.