Asia Pacific Education Review
2007, Vol. 8, No.2, 262-275.

Comparison of Three Growth Modeling Techniques in the Multilevel Analysis of Longitudinal Academic Achievement Scores: Latent Growth Modeling, Hierarchical Linear Modeling, and Longitudinal Profile Analysis via Multidimensional Scaling[1]

Tacksoo Shin
Seoul National University
Korea

This study introduces three growth modeling techniques: latent growth modeling (LGM), hierarchical linear modeling (HLM), and longitudinal profile analysis via multidimensional scaling (LPAMS). It compares the multilevel growth parameter estimates and potential predictor effects obtained using LGM, HLM, and LPAMS. The purpose of this multilevel growth analysis is to alert applied researchers to selected analytical issues that are required for consideration in decisions to apply one of these three approaches to longitudinal academic achievement studies. The results indicated that there were no significant distinctions on either mean growth parameter estimates or on the effects of potential predictors to growth factors at both the student and school levels. However, the study also produced equivocal findings on the statistical testing of variance and covariance growth parameter estimates. Other practical issues pertaining to the three growth modeling methods are also discussed.

Key words: longitudinal academic achievement study, multilevel growth analysis, latent growth modeling, hierarchical linear modeling, longitudinal profile analysis via multidimensional scaling

Longitudinal studies provide important sources of information when investigating how differences in various national and regional school policies, practices and compositional characteristics relate to differences in student achievement over a period of time. Therefore, it is not surprising that the use of growth modeling techniques in educational fields has rapidly increased. Recent years have produced a vast range of applications of structural equation modeling based (SEM) latent growth modeling (LGM) in applied longitudinal data analysis (Curran & Hussong, 2002; Meredith & Tisak, 1990; Muthén, 1991; Willett & Sayer, 1994). With the flexible usage of latent growth factors and measurement error structures, LGM is capable of testing the relative fit of various competing models. Within the domain of hierarchical structure, the hierarchical linear model (HLM) has created a powerful set of techniques for research on individual change (Bryk & Raudenbush, 1987; Foorman, Francis, Novy, & Liberman, 1991; Huttenlocher, Haight, Bryk, & Seltzer, 1991; Rogosa & Willet, 1985). When applied with valid measurements from a multiple-time-point design, this model affords an integrated approach for studying the structure and predictors of individual growth (Raudenbush & Byrk, 2002). Both techniques have the ability to test models that include multiple levels of hierarchical structured data and the capacity to embed assessments in more complex models that assess potential predictors, mediators, and consequences of change (Curran & Hussong, 2002; Raudenbush & Byrk, 2002). Additionally, such approaches are better able to assess

Tacksoo Shin, Department of Education, Seoul National University, Korea.
Correspondence concerning this article should be addressed to Tacksoo Shin, 10-1-601, BK21 Academic Leadership Institute for Competency-based Education Reform, Department of Education, Seoul National University, San 56-1, Sillim-dong, Gwanak-gu, 151-742, Seoul, Korea. e-mail:shin0143@snu.ac.kr

the multivariate patterning change across multiple measures, compared to other traditional methods (Willett & Sayer, 1996). The differences are that in LGM, each participant is required to have the same number and spacing of time points, and level-1 predictors with random effects must have the same distribution across all participants in each subpopulation, while HLM allows unequal numbers and spacing of time points as well as the different distributions of level-1 predictors (Raudenbush & Bryk, 2002).

However, the underlying assumptions therein would appear to limit their applicability. LGM and HLM are based on a likelihood-based approach in computing population parameters. This maximum likelihood (ML) technique[2] is developed under the assumption of multivariate normality and requires a large sample size (Bollen, 1989; Curran, 2003; Jöreskog, 1969). A large sample yields a covariance matrix that is able to produce a better estimate of the population covariance matrix, and should therefore be expected to offer enhanced opportunities to successfully identify the correct model (MacCallum, 1986). Moreover, under the assumption of a multivariate normal distribution of the observed variables, ML estimators have the desirable asymptotic properties of being unbiased, consistent, and efficient (Kmenta, 1971). Unfortunately, there were many situations in which these assumptions were violated in practice. For instance, Miccéri (1989), in his survey of empirical data sets, found that the distributional characteristics of 440 large-sample achievement and psychometric measures were all significantly nonnormal. Despite the development of robust approaches, Chou, Bentler, and Satorra (1991) and Gold, Bentler, and Kim (2003) argued that improvements in Yuan-Bentler ML and asymptotic distribution free method (ADF) are clearly needed, in particular under extreme nonnormal conditions and/or small sample size. Many researchers (Boosma, 1985; Anderson & Gerbing, 1984; Enders & Bandalos, 2001) have pointed out that sample sizes of 100 or less tended to result in high rates of non-convergence in SEM analysis. Therefore, the likelihood estimators may not have the desirable properties. In addition, a covariance structure approach has proven problematic when intercorrelations among some variables are high. Kline (1998) suggested that multicollinearity is the major reason why a sample covariance matrix may be non-positive, and that certain mathematical operations are either impossible, or the results are unstable, because some denominators are very close to zero. Thus, the covariance matrix cannot be inverted to compute the parameter estimates, and it may yield larger variances and covariances of parameter estimates which in turn can affect significance tests (Biesanz, Deeb-Sossa, Papadakis, Bollen, & Curran, 2004; Brekke, Long, Nesbitt, & Sobel, 1997; Huttenlocher, Haight, Bryk, & Seltzer, 1991; Smith, Landry, & Swank, 2000; Stoolmiller, 1995).

In recent years, some pioneers (Davison, Gasser, & Ding, 1996; Davison, Kang, & Kim, 1999; Ding, Davison, & Petersen, 2005; Kim, Frisby, & Davison, 2004) have demonstrated that concepts of individual growth modeling can be accommodated within the framework of multidimensional scaling named longitudinal profile analysis via multidimensional scaling (LPAMS). The strength of LPAMS is to allow for the simultaneous estimation of intra- and inter-individual growth processes using a relatively small number of statistical assumptions. Since LPAMS is similar to a hierarchical modeling in the sense that it allows one to estimate both overall (level-2) and individual (level-1) growth rates, further considerations in interpreting the test results are not required. Additionally, it does not require a large sample size or multivariate normality, but assumes only the homogeneity of variance across time. As with any exploratory approach, it is designed to identify patterns of growth underlying a set of data, not to test a priori hypotheses regarding patterns (Ding, 2003). Therefore, it is considered most useful in the early stages of a research program in which little is understood about the underlying growth patterns. Lastly, because LPAMS is not based on covariance structure but on proximity distance measure (e.g., squared Euclidean distance), multicollinearity does not cause any computational problems within LPAMS technique.

The purpose of this study is to investigate the differences and similarities and to discuss the strengths and weaknesses of LGM, HLM, and LPAMS. In so doing, this study compares multilevel growth parameter estimates, model-fit indices, and potential predictor effects in answering four central research questions: 1) What students' mathematical growth trajectories should be expected, 2) How academic growth rates differ regarding potential predictors (English language program learner; ELP, special educational program; SEP, and gender), 3) In multilevel analysis, what achievement patterns found at the school level should be expected, and 4) How growth trajectories are influenced by a school level predictor (school location: urban and suburban schools). Through four primary analytical approaches, this study aims to alert applied researchers to selected analytical issues that are required for consideration in terms of the decision to apply one of these approaches to academic growth analysis.

Growth Modeling Techniques

Latent Growth Modeling (LGM)

Latent growth modeling (LGM) pertains specifically to latent growth factors, rather than to observed repeated measures of a construct over time. LGM attempts to smooth over observed measures in order to estimate the continuous trajectory that gives rise to these time specific observed measures (Curran & Hussong, 2002). Therefore, we use the observed repeated measures to estimate latent growth factors, thereafter focusing on analyzing these latent growth factors. To estimate these growth factors, the conceptual models are written as the following formations of matrices. Suppose that with a linear growth structure one has four repeated measures.

$$\underbrace{y_{it}}_{\text{Four time points}} = \underbrace{\eta_{ik}}_{\text{Individual growth factors}} \bullet \underbrace{\Lambda_y}_{\text{Factor loadings}} + \underbrace{\varepsilon_{it}}_{\text{Measurement error}}$$

$$n \text{ subjects} \left\{ \begin{bmatrix} y_{11} & y_{12} & y_{13} & y_{14} \\ y_{21} & y_{22} & y_{23} & y_{24} \\ . & . & . & . \\ . & . & . & . \\ y_{n1} & y_{n2} & y_{n3} & y_{n4} \end{bmatrix} \right. = \begin{bmatrix} \eta_{11} & \eta_{12} \\ \eta_{21} & \eta_{22} \\ . & . \\ . & . \\ \eta_{n1} & \eta_{n2} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} & \varepsilon_{14} \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} & \varepsilon_{24} \\ . & . & . & . \\ . & . & . & . \\ \varepsilon_{n1} & \varepsilon_{n2} & \varepsilon_{n3} & \varepsilon_{n4} \end{bmatrix} \qquad (1)$$

where *i* indicates a specific individual and *n* notes total number of subjects.

$$\underbrace{\eta_{ik}}_{\text{Individual growth factors}} = \underbrace{\alpha_{0k}}_{\text{Mean growth factors}} + \underbrace{\zeta_{ik}}_{\text{Random error}}$$

$$\begin{bmatrix} \eta_{11} & \eta_{12} \\ \eta_{21} & \eta_{22} \\ . & . \\ . & . \\ \eta_{n1} & \eta_{n2} \end{bmatrix} = \begin{bmatrix} \mu_{01} & \mu_{02} \\ \mu_{01} & \mu_{02} \\ . & . \\ . & . \\ \mu_{01} & \mu_{02} \end{bmatrix} + \begin{bmatrix} \varsigma_{11} & \varsigma_{12} \\ \varsigma_{21} & \varsigma_{22} \\ . & . \\ . & . \\ \varsigma_{n1} & \varsigma_{n2} \end{bmatrix} \qquad (2)$$

The first row of the left side matrix in equation 1 indicates that the first subject (i = 1) is repeatedly measured four times. Thus, *n* subjects have been collected across four time points. In LGM formula, observed repeated measures can divide into two parts that are on the one hand true latent growth function and on the other error term. If a linear model is imposed, true latent growth is obtained through multiplying two latent growth factors (e.g., η_{i1}: the individual intercept latent factor, η_{i2}: the individual linear growth latent factor) and two factor loadings (a row of 1's for intercept factor and the coding of time: 0, 1, 2, $t - 1$ for linear factor). From matrix equation 2, the estimates of the mean intercept and growth factors (μ_{01} and μ_{02}) are computed. These are referred to as the 'unconditional latent growth curve model.' When an exogenous variable (*X*) is presented so as to predict observed variability, the equation will be changed. In this case, the measurement equation remains the same, but the structural formula is extended to include the effects of these predictor variables.

Regression coefficients
to intercept and linear growth factors

$$\begin{bmatrix} \eta_{11} & \eta_{12} \\ \eta_{21} & \eta_{22} \\ . & . \\ . & . \\ \eta_{n1} & \eta_{n2} \end{bmatrix} = \begin{bmatrix} \mu_{01} & \mu_{02} \\ \mu_{01} & \mu_{02} \\ . & . \\ . & . \\ \mu_{01} & \mu_{02} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \\ . & . \\ . & . \\ \gamma_{n1} & \gamma_{n2} \end{bmatrix} \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix} + \begin{bmatrix} \varsigma_{11} & \varsigma_{12} \\ \varsigma_{21} & \varsigma_{22} \\ . & . \\ . & . \\ \varsigma_{n1} & \varsigma_{n2} \end{bmatrix} \quad (3)$$

Potential predictor

This model is referred to as the 'conditional growth curve model.' For the purposes of this research, these conditional and unconditional LGM equations are used to estimate the growth factors, and test the relative fit of the competing models.

In education, studying growth patterns of standardized test scores has been explored in order to investigate the effects of improvement efforts on academic performance. These efforts are often evaluated in terms of the growth effects of within-level (e.g., individual and student level) growth and in terms of between-level (e.g., teacher, class, school, and school district level) growth[3]. Therefore, policy makers and school administrators are able to obtain valuable information through a multilevel analysis in academic growth studies. For instance, multilevel models are used to answer whether the growth rates of certain schools are faster than those of others. It also explores how aggregated level factors (e.g., teacher qualification, school budget, school size, etc) affect between-level growth rates. A number of researchers (Bentler & Liang, 2003; Linda, Lee, & Poon, 1993; Longford & Muthén, 1992; Muthén, 1989, 1994) suggested that this multilevel strategy allows for the disaggregation of the within (e.g., student level) and between (e.g., school level) covariance structures within a single partitioned covariance matrix. This is then used as the unit of analysis in the estimation. SEM software (e.g. EQS, LISREL, *M*-plus) is used to compute between-level and within-level covariance matrices and between-level means from a raw data file, as well as the current iteration estimates of model parameters.

Hierarchical Linear Modeling (HLM)

Raudenbush and Bryk (2002) noted that many individual change phenomena are able to be represented through a two-level hierarchical level. At level-1, each person's development is represented by an individual growth trajectory that is dependent on a unique set of parameters. They added that these individual growth parameters become the outcome variables in a level-2 model. Thus, under a linear model at level-1 and level-2, equations 4 and 5 simplify to

$$Y_{it} = \pi_{i1} + a_{t-1}\pi_{i2} + e_{it} \quad (4)$$

$$\pi_{i1} = \beta_{01} + r_{i1} \quad (5)$$
$$\pi_{i2} = \beta_{02} + r_{i2}$$

In this regression equation, y_{it} is an observed test score for person *i* at time *t* and a_{t-1} is the coding of time variable (e.g., 0, 1, 2, …, $t - 1$ for a linear growth structure). π_{i1} and π_{i2} are respectively the individual intercepts and the regression coefficients with e_{it} representing the residual error term. When one predictor variable (*X*) is presented,

$$\pi_{i1} = \beta_{01} + \lambda_{i1}X + r_{i1}$$

$$\pi_{i2} = \beta_{02} + \lambda_{i2}X + r_{i2} \quad (6)$$

The level-3 model represents the variability in an aggregated variable (*k*). We can view the level-2 means, β_{01k} and β_{02k}, as varying randomly around a grand mean:

$$\beta_{01k} = \gamma_{011} + u_{01k},\ \beta_{02k} = \gamma_{022} + u_{02k} \quad (7)$$

This equation can also be extended for the level-3 conditional model, as per equation 6.

The method for formulating the model in HLM is similar to that in the LGM approach. In equation 1, the individual growth factors (η_{i1} and η_{i2}) are equivalent to a matrix formation of π_{i1} and π_{i2} in equation 4. β_{01} and β_{02} indicate respectively the average of intercept and slope growth factors the same as μ_{01} and μ_{02} of LGM.

However, the estimation methods of LGM and HLM are fundamentally different. Curran (2003) explained that the assumption of independence of observations is highlighted in the standard estimation of the SEM in that the discrepancy function is based on a single aggregate sample covariance matrix that allows for the covariance structure within any other level of nesting that is assumed to be null. In contrast, the estimation of the HLM outcome incorporates complex data structures among lower and higher levels of data hierarchy in that the nesting in the data is explicitly modeled. Thus, Curran (2003) added that whereas nested data structures pose a significant problem to standard ML estimation in SEM, the estimation of the HLM explicitly allows for these dependent structures. Moreover, time scores are treated as data in HLM whereas time scores of LGM are treated as parameters.

Longitudinal Profile Analysis via Multidimensional Scaling (LPAMS)

The longitudinal profile analysis via multidimensional scaling (LPAMS) starts its analysis from the following equation:

$$y_{it} = c_i + \sum w_i x_t + \varepsilon_{it} \qquad (8)$$

where y_{it} is an observed test score for person *i* at time *t*, c_i is a $i \times 1$ vector of constant intercept terms, w_i is a profile match index characterizing person *i*, and x_t is a growth scale value reflecting the location of a repeated variable at time *t* (Kim, Frisby, & Davison, 2004). If *k* ($\geq$ 2) growth dimensions are required (i.e., multidimensional growth cases), there will be *k* numbers of w_i ($w_{i1}, w_{i2}, \ldots, w_{ik}$) and x_t ($x_{t1}, x_{t2}, \ldots, x_{tk}$). LPAMS consists of three steps (Ding, 2003). The analysis begins with a matrix containing a proximity measure defined over all possible pairs of variables. In this study, the variables are time points, and the proximity measured for each possible pair of time points, a squared Euclidean distance measure[4], is computed from the raw data. When proximity measures are submitted to an appropriate multidimensional scaling algorithm, the analysis is expected to yield one dimension for each growth curve. Since the objective of LPAMS model analysis is to determine whether there are particular trend shapes found within the data, and whether those shapes are linear curves, nonlinear curves, or time-series periodic curves, a multidimensional scaling (MDS) estimation method would be capable of identifying one or more such trend shapes where such curves exist in the data.

In the second step, the zero point of the scale values are reset so that the scale values indicate growth rates for each time-point, and in order that the intercept estimate can be interpreted as the initial growth level. Ding, Davison, and Petersen (2005) and Kim, Frisby, and Davison (2004) have explained that if the zero point is set to correspond with the growth scale value as at the first time-period, then that will result in x_1^* = 0 for all profile *k*. Therefore, the growth scale value of the first time-point is reset to zero on each profile in such a way that $x_t^* = x_t - x_1$. The rescaled growth scale values estimates would have a range of zero to some positive or negative numbers, depending on the shape of the curve. These growth scale values would represent the growth rate for each time-point. The authors added that c_i becomes the expected score under the model for person *i* at initial time *t* = 1, reflecting the initial growth level. In the third step, the intercept (c_i) and profile match index parameters (w_i) can be estimated for each individual by regressing the observed scores (y_{it}) onto the rescaled value estimates (x_t^*).

These LPAMS mathematical terms can be interpreted as per these of LGM. The averages of c_i and w_i are respectively equivalent to μ_{01} and μ_{02} indicating the means of intercept and growth factors in LGM. Also, x_t^* can be viewed as a special case of factor loading (Λ_y) in LGM (see equation 1). The difference is that LPAMS directly calculates x_t^* from the data (random effect), whereas LGM considers a fixed effect (e.g. 0, 1, 2,…, *t* – 1 for the linear growth model). Table 1 presents an analogy of the elements and notations of three growth modeling structures.

Since LPAMS offers parameter estimates and a fit index of individual growth, it appears to be useful for applied researchers interested in predicting each subject's initial score and growth rate. On the other hands, LGM provides average growth information and an overall fit index. One other difference is that the LPAMS model estimates within individual changes with respect to the latent change curves using the MDS method and estimates between-individual variations via a conventional analysis of variance (ANOVA) approach, while HLM represents a clustering of individuals within group, and variables are measured at all available levels. This model, then, combines variables from different levels in one statistical model (Ding, 2003). In addition, the LPAMS model is based on a distance model (Borg & Groenen, 1997; Davison, 1983) rather than a linear model. Thus, it can be used to model data that, by nature, are nonlinear. For conditional analysis, LPAMS is easily capable of testing the effects of potential predictors to growth profiles

Table 1

Analogy of elements and notations of three growth modeling structures

	LGM	HLM	LPAMS
Factor loading	Λ_y	a_{t-1}	x_t^*
	Fixed growth effect coding $\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ . & . \\ . & . \\ 1 & t-1 \end{bmatrix}$	Fixed growth effect coding same as LGM factor loading	Sample driven random growth effect coding $\begin{bmatrix} 1 & x_1^* = 0 \\ 1 & x_2^* \\ 1 & x_3^* \\ 1 & x_4^* \\ . & . \\ . & . \\ 1 & x_t^* \end{bmatrix}$
Growth parameters			
Intercept	η_{i1}	π_{i1}	c_i
Slope	η_{i2}	π_{i1}	w_i
Mean intercept	μ_{01}	β_{01}	$\bar{c}$
Mean slope	μ_{02}	β_{02}	$\bar{w}$

Note. This comparison is based on a linear growth structure with *t* time points

Note. The different way of fixed growth effect time coding (e.g., centered time coding) in LGM and HLM can be implemented (Biesanz, et al., 2004; Brekke, et al., 1997; Huttenlocher, et al., 1991; Smith, Landry, & Swank, 2000; Stoolmiller, 1995).

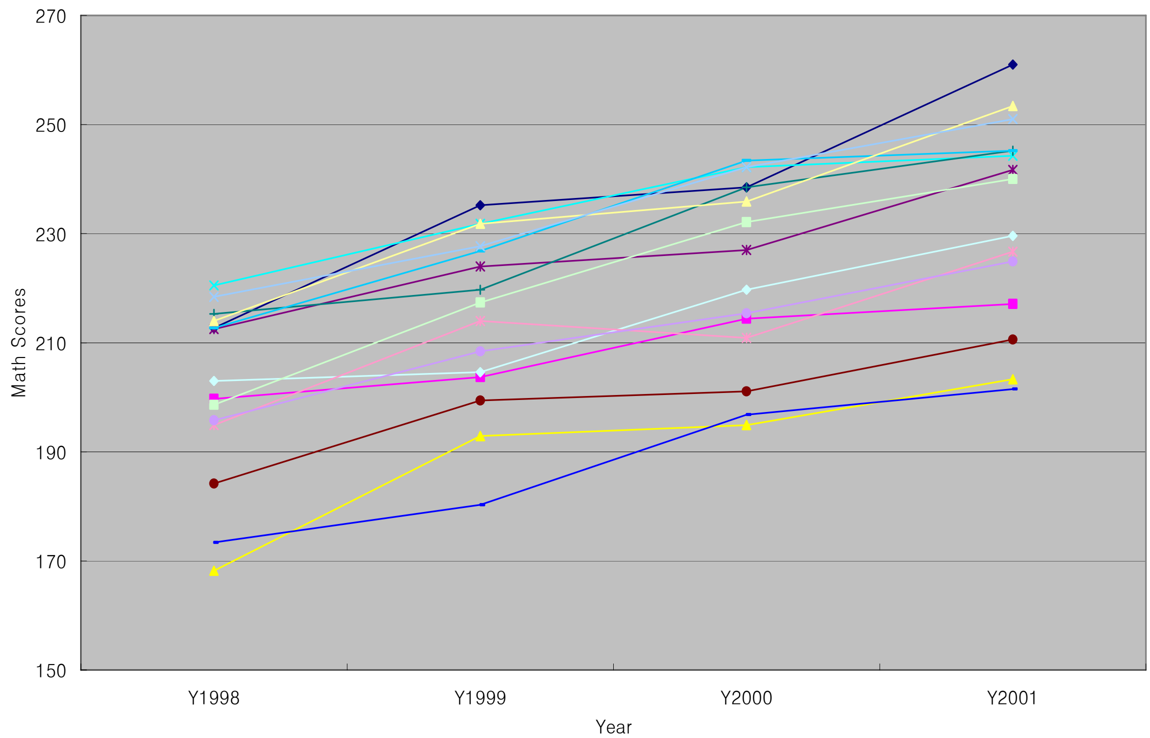

Figure 1. 15 Randomly Selected Individual Growth Curves for Mathematics Achievement Scores over Four Time Points

by using the ANOVA approach. LPAMS performs aggregation of the individual test results into high level variable (aggregated variable) for multilevel analysis.

Methodology

Sample Description

The data set was obtained from a sample of 1244 students from a large school district. These students were in the 2nd grade at the first instance of measurement, and were then followed for four years (from 1998 to 2001). The students took a mathematics achievement test, (the Northwest Achievement Level Math Test), with vertically equated scale scores each year from the 2nd to 5th grades. Both LGM and LPAMS models require time-structured data so that all sample members were observed on the same set of occasions. Figure 1 displays the math growth trajectories of 15 randomly chosen students. From the graph of their individual curves, we might infer that most individual math scores tended to increase over time.

Three covariates measured for student level analysis at the first time-point were considered. These potential predictors of change were coded: 0 = students were not in special educational program (SEP), and 1 = students were in SEP; 0 = female, and 1 = male; 0 = non-English language leaner (ELL), and 1 = ELL. 53% of the students were male. 12% were in special educational programs and 24% had limited English proficiency levels. The students were enrolled in a total of 28 schools. The ranges of the number of students nested within one school were 30 through 68. 7 schools were situated in suburban area, with this school location considered an exogenous variable for school level analysis.

Results

Table 2 summarizes the unconditional test results of student and school levels for LGM, HLM, and LPAMS models. The quadratic growth models of LGM and HLM yielded the best fit in terms of student and school level. A LGM's standard decrement-to-chi-squared test between linear and nonlinear models (e.g., student level: change in x^2 = 13.370 on *df* = 3 with the critical value x^2 = 7.815 at α = .05) and HLM's deviance test between two models (e.g., student level: change in deviance = 40.862 on *df* = 5 with the critical value x^2 = 11.070 at α = .05) also revealed that the quadratic model fit the data better than the linear model. However, the means and variances of quadratic factors for both student and school levels were not significant. Additionally, the linear growth model with heterogeneous error variance structure presented reasonable model-fit statistics. Thus, this study interprets the change and variability of math achievement with a linear growth function. As previously mentioned, where there is more than one growth trend (e.g., nonlinear curve) in the data, the LPAMS model can simultaneously identify these curves with single growth factor. Therefore, the model-fit testing between linear and nonlinear models is not necessary under LPAMS analysis conditions.

Unconditional Student and School Level Analyses

For the unconditional student level analysis, all methods yielded similar mean growth parameter estimates. The average initial and growth scores of LGM, HLM, and LPAMS were respectively 194.580, 194.532, and 194.286, and 8.800, 8.975, and 9.824. However, the predicted mean scores for each year should be calculated differently. In this study, while the growth factor loading is fixed, such as at 0, 1, 2, 3 for a linear model in LGM and HLM, the LPAMS model uses a sample-driven factor loading, such as at 0, .992, 1.757, 2.723. Again, the factor loading (scale value) has a random effect in LPAMS, whereas time is entered as the values of the factor loadings relating the repeated measures to the underlying latent growth factors in LGM and HLM. The growth levels of 15 randomly selected individual students under LPAMS are presented in Table 3.

The variance components of initial and growth parameter estimates for all three approaches were statistically significant. This suggests that there was meaningful evidence of inter-individual heterogeneity at the first time-point and over time. However, LGM and HLM tended to produce the smaller intercept and slope variance estimates (168.503 and 4.744 for LGM; 169.303 and 4.446 for HLM) than LPAMS (194.652 and 12.728). Significantly, the major difference was found in covariance estimate testing. The covariance parameter estimate found under LGM and HLM (respectively 10.512 and 10.571) revealed a significant positive relationship between intercept and growth factors (correlation = .372 for LGM and .385 for HLM). This estimate suggests that growth rates of students reporting higher math scores in the first year would be relatively faster than those of students reporting lower achievement scores in the same year. In contrast, the covariance of LPAMS (.083) was not found to be significant. This result indicates that students' initial scores would not be related to their academic growth rate levels.

Table 2
Results of Unconditional Multilevel Analyses

	Student Level			School Level			
	LGM Heterogeneity	HLM Heterogeneity	LPAMS	LGM Heterogeneity	LGM Homogeneity	HLM Homogeneity	LPAMS
Average Intercept	194.580***	194.532**	194.286***	194.138***	194.195***	194.195***	193.893***
Average Slope	8.800***	8.795***	9.824***	8.794***	8.707***	8.701***	9.725***
Variance of Intercept	168.530***	169.303***	194.652***	24.445***	27.354***	29.305***	34.917***
Covariance of Intercept and Slope	10.512 (.372)**	10.571 (.385)**	.083 (.002)	2.805 (.625)*	2.235 (.528)	1.435 (.245)	1.602 (.114)
Variance of Slope	4.744**	4.446**	12.728**	.825	.957	1.168*	1.753*
Model Fit	χ^2 = 52.126 with *df*=5 *p*-value = .000, RMSEA=.064	Deviance = 35449.761	STRESS < .001, Average R-square = .865	χ^2 = 19.466 with *df*=8 *p*-value = .013, RMSEA=.045	χ^2 = 24.921 With *df*=11 *p*-value=.009, RMSEA=.048	Deviance = 35264.986	STRESS < .001, Average R-square = .857

*** $p < .001$; ** $p < .01$; * $p < .05$.

Note. The parenthesis ( ) indicates the correlation values.

Note. Heterogeneity refers to a heterogeneous error structure being imposed. Homogeneity indicates that the homogeneous error variance is implemented.

Table 3
Growth Information of 15 Randomly Selected Students by LPAMS

Student ID	Intercept (c_i)	Growth Profile (w_i)	Model-Fit (r-square)
464480	199.87	6.02	0.81
701170	174.96	7.41	0.86
513985	169.18	8.75	0.96
546458	188.64	5.73	0.93
637986	185.19	12.76	0.97
748844	178.41	6.75	0.92
546325	182.21	11.75	0.94
837838	175.55	11.35	0.98
337442	174.16	10.3	0.88
917907	186.28	7.33	0.91
925041	185.07	13.93	0.95
338119	195.95	13.54	0.94
688223	176.05	11.09	0.88
294498	172.36	14.21	0.89
836006	187.79	9.84	0.95

The different model-fit indices were considered for the three modeling techniques. In LGM, most applied researchers are interested in the chi-square fit test and the root mean square error of approximation (RMSEA). Various fit indices (e.g., GFI, AGFI, NFI, CFI, etc.) are also supplied by SEM-based software. HLM yields the deviance values and finds the most plausible model through the deviance-difference test. For LPAMS, the adequacy of one-dimensional MDS solution is verified by the MDS fit index, STRESS formula 1 (Kruskal, 1964). Furthermore, an r-square value is imposed on the model so as to examine the goodness of fit. In this study, the linear LGM exhibited a reasonable fit. The chi-square fit statistic was 52.126 with $df = 5$, $p < .001$. With over 1000 people in the sample, this chi-square statistic has a large degree of power to detect even small deviations from the model. The GFI (goodness of fit index) equaled .995. The RMSEA equaled .064. According to the guidelines of Browne and Cudeck (1993), an RMSEA of .05 or less signifies a close fit, with that at or below .08 indicates a reasonably fitting model. While there were statistically significant deviations from the model, these deviations were small, and the covariance structure showed a reasonably good level of fit. In HLM, the linear model with a heterogeneous error structure model fit the data better than that with the homogeneous error variance model (deviance of difference $x^2 = 28.465$ on $df = 3$ with a critical value $x^2 = 7.815$ at $\alpha = .05$). Lastly, the STRESS formula 1 value of LPAMS was less than .001, indicating the rank ordering of four proximity data points could be perfectly reproduced by the one-dimensional solution obtained. The mean r-square value of the LPAMS approach was .865. This indicated that the average 86.5 % of total variability was accounted for by the LPAMS growth equations.

All three modeling techniques again yielded similar mean growth parameter estimates at the school level. These results indicated that schools were reporting significant initial math scores of approximately 194 points, and linear rates of increase around 9 points per time-point. However, dissimilar results of statistical testing on variance and covariance parameter estimates were discovered. The variance estimates of LPAMS were larger than those of LGM and HLM. Moreover, the variance of LGM growth rates (.825) was not significant, while that of HLM and LPAMS growth profiles (respectively 1.168 and 1.753) expressed meaningful evidence of school variability over time. In the case of the covariance parameter estimate, LGM found a significant positive relationship between initial and true growth factors (covariance = 2.805, correlation = .625). This estimate suggests that the schools that reported higher math scores at the beginning time-point tended also to report faster rates of increase than schools reporting lower math scores in the first year. On the other hands, HLM and LPAMS revealed no significant relationship between intercept and growth factors.

Conditional Student and School Level Analyses

The test results of the three different approaches suggest

Table 4
Results of Conditional Multilevel Analyses

	Student Level				School Level		
	LGM Heterogeneity	HLM Heterogeneity	LPAMS		LGM Heterogeneity	HLM Homogeneity	LPAMS
Intercept and LEP	-2.129 (-.377)**	-11.614**	-11.633**	Intercept and LOCATION	1.619 (.538)**	8.845**	9.197**
Intercept and SEP	-2.503 (-.272)**	-10.822*	-10.789**		--	--	--
Intercept and GENDER	1.621 (.198)*	1.514*	1.506*		--	--	--
Slope and LEP	-.027 (-.029)	-.070	-.006	Slope and LOCATION	.126 (.134)	.614	.448
Slope and SEP	-1.005 (-.204)*	-1.873*	-2.078*		--	--	--
Slope and GENDER	.069 (.063)	.258	.289		--	--	--

*** $p < .001$; ** $p < .01$; * $p < .05$.

Note. The parenthesis ( ) indicates the correlation values.

that a moderate level of collinearity exists between the covariates and the growth factors of math scores. LGM found that with a negative relationship between English language learner (ELL) status and initial factor (covariance estimate = - 2.129; correlation estimate = - .377), we are able to conclude that students who were in ELL reported lower math scores than students not in said program at the beginning time-point. With respect to the special education program (SEP), effects were found on both the intercept and growth terms. The negative relationship between SEP and the initial factor (covariance = - 2.503; correlation = -.272) indicated that SEP students reported lower math scores in the 2nd grade. Additionally, the growth rate of SEP students (covariance = - 1.005; correlation = - .204) was relatively lower than that found among non-SEP students. Finally, male students reported higher math scores in the first year (covariance =1.621; correlation = .198), while there was no significant gender-related impact on the growth rates. HLM and LPAMS also showed that students who took the English language program reported lower math achievement levels, to the extent of an average of 11.614 (11.633 for LPAMS) points in the first year. SEP students reported lower initial and growth scores of respectively 10.822 and 1.873 (10.789 and 2.078 for LPAMS) points than non-SEP students. Finally, male students tended to obtain slightly higher average math scores at the beginning time-point (i.e., as much as 1.514 points for HLM, and 1.506 pints for LPAMS).

There was one potential predictor for the conditional analysis at the school level, which was school location (7 suburban and 21 urban schools). By LGM, the results of testing the school location variable suggested that school location had a significant effect on the intercept factor, whereas growth rate was not influenced by the same variable. This result indicates that schools located in suburban areas reported higher math scores than urban-located schools at the beginning time-point (covariance = 1.619, correlation = .538). HLM also revealed that suburban schools reported higher math score, by as much as 8.845 (9.197 for LPAMS) points in the first year, with no subsequent relationship between school location and growth rate. Table 4 illustrates the test results of conditional student and school levels.

Discussion

LGM, HLM, and LPAMS methods of analysis resulted in no significant differences in average growth parameter estimates in the student and school level analyses, regardless of the imposition of different growth factor loadings. In the case of the effects of potential predictors to growth factors, all modeling techniques yielded the same test results. The math achievements of non-SEP, non-ELL, and male students tended to be higher than those of other students at the beginning time-point. The growth rate of non-SEP students was also faster than that of SEP students. For the purposes of the multilevel analysis, although urban schools had a tendency to report lower academic achievements, there were no significant differences in the growth rates depending on school location.

However, this study presented some dissimilar test statistics among LGM, HLM, and LPAMS methods. Firstly, the student and school levels' variance parameter estimates of LGM and HLM tended to be smaller than those of LPAMS. In the case of school level, HLM and LPAMS produced significant school level differences with regard to growth rates, whereas LGM did not. Importantly, the major discordance among growth techniques was the significance test results on covariance parameter estimates. LGM persistently showed that initial status and true gains were positively related in student and school level analyses. This indicates that these achievement gaps increased, as students and schools with lower scores at the beginning also reported lower growth rates. The patterns of variances for student and school levels over time fully support these interpretations. From 1998 to 2001, since the variances of student and school levels tended to increase over time (respectively, 214.564 > 209.987 < 259.662 < 294.077; 40.368 > 34.078 < 51.294 < 57.035), we should therefore conclude that the achievement gaps among students and schools widen. HLM showed a significant relationship between intercept and slope in the student level, while that relationship disappeared in the school level analysis. Additionally, none of the covariance estimates in student and school levels were deemed significant under the LPAMS approach. This suggests that the achievement gaps among students and schools were parallel across years.

There are a number of potential reasons behind dissimilarities in test results among LGM, HLM, and LPAMS methods. Firstly, LGM and HLM use the ML technique to estimate the population parameter estimates, whereas the ordinary least squares estimator is imposed to obtain LPAMS parameters. When observations are independent of one another and are normally distributed with a constant variance, both ML and least square methods intersect (Myung, 2003). However, most research situations do not meet these conditions. Especially under longitudinal study, since it tracks the same persons and involves observations of repeated measures over time meaning that auto-correlated errors are an important consideration, ML estimates are more likely to differ from ordiarny least square estimates. Secondly, the assumption of homogeneity of variance would affect the test statistics. Although LPAMS assumes a homogeneous error variance only, an equal variance assumption is more likely to violate this in actual longitudinal analysis. In this study, given that the homogeneity of the variance test was rejected, we should be careful when interpreting the test results produced by LPAMS. The level-3 (between-level) growth information would also be influenced by this assumption. Raudenbush and Bryk (2002) argued that when unidentified slope heterogeneity at lower-level appears as the heterogeneity of lower-level error variance, such slope heterogeneity might be expected to bias estimates of the higher-level coefficients. Related to this issue, the different measurement error structures may be sufficient to cause different statistical testing on variance and covariance estimates. LGM can impose both the heterogeneous and homogeneous error variance models on the within- (level-2 in HLM) and between-level (level-3 in HLM) testing, while HLM cannot implement the heterogeneous error variance model on the analysis of level-3. Therefore, where the homogeneous error structure is implemented within an LGM context at the school level, the variance of the slope factor tended to be larger, and covariance between the initial status and true gain became insignificant, as illustrated in Table 3.

Another possible explanation of the differences in the test results may be due to sample size. LGM and HLM procedures, which are maximum likelihood based approaches, would not ordinarily produce stable results if the sample size at any level is too small. Additionally, given a small country-level sample size (27 countries), Cheung and Au (2005) reported that results at the individual level were quite stable even when using such small individual-level sample sizes, whereas the group-level parameter estimates and their standard errors were unsystematically affected by varying individual-level sample sizes. In this study, the total number of schools was 28 and the average number of students within the school level was approximately 41. The major concern here is the lack of a clear cut point determining how many samples need to be collected at any given level, and how many lower-level data are required within higher level studies. Although LPAMS has no serious limitation in relation to sample size, regression and ANOVA approaches under LPAMS analysis also require a large enough sample size to achieve the desired power level of significance testing. If not, there is no way to guarantee the stable result. Therefore, one may not be certain as to whether the estimator is estimating a meaningful value. Lastly, high correlations between measured variables may yield problematic test results on parameter estimates of LGM and HLM methods. This multicollinearity would cause inflated standard error estimates leading to inflating the type II error rate, especially for covariance structure analysis. Vasu and Elmore (1975) indicate that violation of the assumption of normality coupled with the condition of multicollinearity results in large standard errors in the sampling distributions of the standardized coefficients. According to the results of Grewal, Cote, and Baumgartner (2004), when multicollinearity is extreme, type II error rates can be unacceptably high, being greater than 80% in cases where the correlations are greater than .80. In this study, the ranges of correlations in student and school levels were respectively .82 to .91 and .88 to .97. Under these data abnormalities, LGM and HLM may be differentially sensitive. Fortunately, it may be anticipated that the assumption of multivariate normality did not greatly affect the test results. The ranges of univariate skewness and kurtosis of variables were respectively .043 to .122 and -.045 to -.207, with the value of Mardia-based Kappa (mulitivariate normality test) tending to zero (.102)

From a statistical point of view, these results make it difficult to select a single most appropriate modeling technique. Although applicability of LGM would be limited under several statistical conditions (e.g., multicollinearity, small sample size, multivariate nonnormality), it has a more flexible array of possible covariance structures for modeling random effects and residuals (Rovine & Molenaar, 1998). Additionally, LGM is able to model and comparatively evaluate a broader array of growth functions (du Toit & Cudeck, 2001). HLM is less likely to be associated with estimation problems (i.e., difficulty in obtaining convergence of the estimation procedure) and efficiently handles the unbalanced hierarchical data structure. However, this hierarchical modeling technique also shares most of the statistical limitations of LGM. LPAMS is an integrated technique for exploring developmental growth trends (e.g., systematic and directional growth) as well as exploring change patterns (e.g., oscillation between ups and downs) that are not growth in nature (Ding, Davison, & Petersen, 2005). In particular, if a study includes many time points (e.g., more than 7 time points) and more than two growth dimensions require change detection, LPAMS must be a useful technique. Thus, it is widely applicable to various longitudinal studies including academic achievement growth and psychological change (e.g., mood change per day). However, as discussed above, the homogeneous error variance assumption may be easily violated in real circumstance and the violation of this assumption threatens the validity of statistical inferences.

Lastly, several other research conditions could limit the findings of this study being able to be more widely generalized. First, the results of this analysis may not be applicable when growth trend shows a nonlinear change. Second, though the repeated measured variables tend to be highly correlated, the cross-time correlation pattern in this study is extraordinary high (.80 to .90) for much educational and psychological data. Unlike LPAMS, LGM and HLM function oddly at this level of abnormality of data. Thus, if lower correlations between variables are obtained, test results may be different. Related to this issue, further longitudinal study should examine the unique and combined effects of varying sample sizes at within (e.g., student level) and between (e.g., school level) measures and modeling diverse growth structures, as well as a changing degrees of nonnormality and of correlations between observed variables for each growth method.

Conclusion

This study suggests that applied researchers should select appropriate growth modeling techniques, depending on their research questions and circumstances. If one is interested in each individual growth rate, LPAMS is appropriate. In addition, since LPAMS does require a small number of statistical assumptions, it would be able to provide reasonable test results under a diverse range of adverse research conditions (e.g., multicollinearity, multivariate nonnormality, etc). Being flexible and efficient in terms of technical and structural aspects, LGM is deemed preferable for large sample based multilevel studies (e.g., national level studies) and for longitudinal research involving heterogeneity of error variance across time points. HLM would be superior to other growth techniques, in particular in unbalanced time-spacing data structure and where dependency among observations is observed. Significantly, when the modeling techniques yield equivocal results, especially in terms of statistical testing of variance and covariance estimates due to the imposition of differing statistical estimation techniques and assumptions, researchers should be careful when interpreting the inter- and intra-variability levels and the relationship between initial scores and true gains. The recommendation herein is to examine whether the variances of measurements change over time (e.g. increase and/or decrease), or remain parallel. If the trends of variances decrease over time, this may indicate that individual variability has decreased, and that the initial scores tended to be negatively related to growth rates. On the other hand, where variances increase over time, the individual differences may appear to have increased, and the relationship between initial and growth factors have exhibited the tendency to be positively related. While researchers remain in need of performing statistical tests of covariance parameter estimate, this descriptive information is necessary to understand the general trends in the data, as well as offering helpful advice in determining the most appropriate growth parameter estimates.

Notes

1. This work was supported by BK21 Academic Leadership Institute for Competency-based Education Reform, Department of Education at Seoul National University.
2. Simply, maximum likelihood technique is to find a parameter (θ) that maximizes the sample likelihood (i.e., make the observed data most likely). Suppose that we collected data (measurements *y*) from the population. Probability of *y* given θ is normalized probability density ($\int p\{y \mid \theta\} dy = 1$) of *y* as a function of parameter θ, then the parameter θ can be estimated by joint probability density for the *n* measurements *y:*

$$L(y \mid \theta) = \prod_{i=1}^{n} p(y_i \mid \theta)$$

This $L(y \mid \theta) = L(y_1, y_2, \ldots\ldots, y_n \mid \theta)$ is called likelihood function. Maximizing *L* by varying θ amounts to interpreting *L* as function of θ, given the measurements *y*.
3. Compared with the hierarchical Linear Modeling (HLM), the within-level analysis is referred to as level-1 and level-2. The between-level analysis is indicated as level-3.
4. Euclidean distance is to measure the ordinary distance between two points. In longitudinal study, these points become time points, and the proximity measure for each possible pair of time points (t, t') is a squared proximity distance measure, $\delta^2_{tt'}$, computed from the raw data as follows:

$$\delta^2_{tt'} = (1/i)\sum_i (y_{i(t)} - y_{i(t')})^2$$

Thus, in the matrix of squared Euclidean distance, all main diagonal values should be equal to zero and off diagonal values indicates the distances among time points. In the case of covariance matrix, the main diagonal and off diagonal values respectively notes variance and covariance.

References

Anderson, J. C., & Gerbing, D. W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*, *49*, 155-173.

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

Biesanz, J. C., Deeb-Sossa, N., Papadakis, A., Bollen, K. A., & Curran, P. J. (2004). The role of coding time in estimation and interpreting growth curve models. *Psychological Methods*, *9*, 30-52.

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

Boosma, A. (1985). Nonconvergence, improper solutions, and starting values in LISREL maximum likelihood estimation, *Psychometrika*, *50*, 229-242.

Borg, I., & Groenen, P. (1997). *Modern multidimensional scaling: Theory and applications*, New York: Springer.

Brekke, J. S., Long, J. D., Nesbitt, N., & Sobel, E. (1997). The impact of service characteristics on functional outcomes from community support programs for persons with schizophrenia: a growth curve analysis. *Journal of Consulting and Clinical Psychology*, *65*, 464-475.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 136-162). Beverly Hills, CA: Sage.

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

Chou, C. P., Bentler, P. M., & Satorra, A. (1991). Scaled test statistics and robust standard for non-normal data in covariance structure analysis: A Monte Carlo study. *British Journal of Mathematical and Statistical Psychology*, *44*, 359 – 368.

Cheung, M. W.-L., & Au, K. (2005). Applications of multilevel structural equation modeling to cross-culture research. *Structural Equation Modeling*, *12*, 598 – 619.

Curran, P.J. (2003). Have multilevel models been structural equation models all along? *Multivariate Behavioral Research, 38*, 529-569.

Curran, P. J., & Hussong, A. M. (2002). Structural Equation Modeling of Repeated Measures data: Latent Curve Analysis. In D. S. Moskowitz & Scott L. Hershberger (Eds.), *Modeling Intraindividual Variability with Repeated Measures Data: Methods and Application.* New Jersey: Lawrence Erlbaum Association.

Davison, M. L. (1983). *Multidimensional scaling*. New York: Wiley.

Davison, M. L., Gasser, M., & Ding, S. (1996). Identifying major profile patterns in population: An exploratory study of WAIS and GATB patterns. *Psychological Assessment*, *8*, 26 – 31.

Davison, M. L., Kang, H., & Kim, S. (1999). The structure of ability profile pattern: A multidimensional scaling perspective on the structure of intellect. In P. L. Ackerman, P. C. Kyllonen, & R. D. Roberts (Eds.), *Learning and individual differences: Process, trait, and content determinant* (pp. 187 – 204). Washington, D.C.: APA Books.

Ding, C. (2003). Exploratory longitudinal profile analysis via multidimensional scaling. *Practical Assessment, Research & Evaluation*, On-line serial *8*.

Ding, C. S., Davison, M. L., & Petersen, A. C. (2005). Multidimensional scaling analysis of growth and change. *Journal of Educational Measurement, 42*, 171-191.

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

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling*, *8*, 430 – 457.

Foorman, B., Francis, D., Novy, D., & Liberman, D. (1991). How letter-sound instruction mediates progress in first-grade reading and spelling. *Journal of Educational Psychology*, *83*, 456-469.

Gold, M, S., Bentler, P. M., & Kim, K. H. (2003). A comparison of maximum-likelihood and asymptotically distribution-free methods of treating incomplete nonnormal data. *Structural Equation Modeling*, *10*, 47 – 79.

Grewal, R., Cote, J. A., & Baumgartner, H. (2004). Multicollinearity and Measurement Error in Structural Equation Models: Implications for Theory Testing, *Marketing Science*, 23, 519-529.

Huttenlocher, J. E., Haight, W., Bryk, A. S., & Seltzer, M.

(1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*, 236 – 249.

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, *34*, 183-202.

Kim, S., Frisby, C. L., & Davison, M. L. (2004). Estimating cognitive profiles using profile analysis via multidimensional scaling (PAMS). M*ultivariate Behavioral Research, 39*(4), 595 – 624.

Kline, R. B. (1998). *Principles and Practice of structural equation modeling*. New York: The Guliford press.

Kmenta, J. (1971). *Elements of econometrics*. New York: MacMillan.

Kruskal, J. B. (1964). Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 29, 115 – 129.

Linda, N. Y., Lee, S.-Y., & Poon, W.-Y. (1993). Covariance structure analysis with three level data. *Computational Statistics & Data Analysis*, *15*, 159–178.

Longford, N. T., & Muth´en, B. O. (1992). Factor analysis for clustered observations. *Psychometrika*, *57*, 581–597.

MacCallum, R. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin*, *100*, 107-120.

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

Miccéri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, *105*, 156 – 166.

Muthén, B. O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika, 54,* 557–585.

Muthén, B. O. (1991). Analysis of longitudinal data using latent variable models with varying parameters. In L. M. Collins, & J. L. Horn (Eds.), *Best methods for the analysis of change: Recent advances, unanswered questions, future directions.* Washington, D.C.:APA.

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

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data*
analysis methods. California: Saga Publications, Inc.

Rogosa, D. R., & Willett, J. B. (1985). Understanding correlates of change by modeling individual differences in growth. *Psychometrika*, *50*, 203-228

Rovine, M. J. & Molenaar P. M. (1998). A LISREL model for the analysis of repeated measures with a patterned covariance matrix. *Structural Equation Modeling*, *5*, 318 – 343.

Smith, K. E., Landry, S. H., & Swank, P. R. (2000). Does the content of mothers' verbal stimulation explain differences in the children's development of verbal and nonverbal cognitive skills? *Journal of School Psychology*, *38*, 27-49

Stoolmiller, M. (1995). Using latent growth models to sutyd developmental processes. In J. M. Gottman (Ed*.), The analysis of change* (pp. 103-138). Mahwah, NJ: Erlbaum.

Vasu, E. S., & Elmore, P. B. (1975). *The effect of multicollinearity and the violation of the assumption of normality on the testing of hypotheses in regression analysis.* Paper presented at the Annual Meeting of the American Educational Research Association. Washington, D.C.

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

Received March 9, 2007
Revision received April 10, 2007
Accepted June 26, 2007