LONGITUDINAL STUDIES OF ACHIEVEMENT GROWTH USING LATENT VARIABLE MODELING

BENGT O. MUTHÉN

UNIVERSITY OF CALIFORNIA, LOS ANGELES

SIEK-TOON KHOO

ARIZONA STATE UNIVERSITY

ABSTRACT: This article gives a pedagogical description of growth modeling of longitudinal data using latent variable methods. The growth modeling is described using an example of mathematics achievement developing over grades 7 to 10 in two cohorts of students. The article describes the basic idea behind growth modeling of individual differences in growth over time and applies it to mathematics achievement development as a function of background variables such as gender, mother's education, and home resources. The modeling ideas are described in words, diagrams, and formulas. The discussion covers modeling that assesses the form of the growth, the influence of background variables on the growth, multiple-cohort analysis, analysis with missing data, and multiple-group analysis of males and females. A corresponding set of analyses are performed on the mathematics data to illustrate the modeling ideas.

Educational studies are often used to answer questions about educational progress and obstacles to such progress. Longitudinal data with repeated measurements on a set of individuals provide information for answering such questions. Examples of large-scale, longitudinal data bases are NELS (the National Education Longitudinal Study), LSAY (the Longitudinal Study of American Youth), and NYS (the National Youth Survey). Often, questions are tangible but

Direct all correspondence to: Bengt O. Muthén, Graduate School of Education & Information Studies, University of California, Los Angeles, CA 90095. E-mail: bmuthen@ucla.edu

Learning and Individual Differences, Volume 10, Number 2, 1998, pages 73–101.

ISSN: 1041-6080

regard complex underlying processes that can only be observed with fallible indicators—e.g. what particular achievement disadvantages population subgroups have, what usefulness new test forms have, what gender differences develop over time in math and science achievement, what effects tracking has on mathematics skill development, and what developmental differences exist across adolescents with respect to problem behavior. Data are available in complex multivariate form, often involving different test forms, attrition, and students sampled hierarchically within schools.

This article focuses on new latent variable technology that is suitable for studies of the above kind. Methods are described for modeling individual differences in growth. In such modeling, however, data are usually not available in a form that allows standard analysis. Instead, it is frequently the case that some data are missing. Methods have recently been developed for modeling with data that are not necessarily missing completely at random. These new techniques are shown to fit into a conventional structural equation modeling framework, including mean structures and multiple-group techniques. Existing structural equation modeling software can be used. The article focuses on developments for continuous response variables, using conventional, normal-theory estimators such as maximum-likelihood. This makes it easier to cover the topics in one article, utilizing their common features. The central unifying feature is latent variable constructs.

The article is structured as follows. First, a motivating example will be presented. Using this example, the growth modeling will be introduced, starting from a relatively simple situation and making the modeling increasingly more realistic and complex. The method will be presented using four components. The methods ideas will be presented conceptually in words, statistically in terms of formulas, graphically in terms related to software specifications, and numerically in terms of analysis results for the example. The presentation is hereby made more generally accessible. For example, the statistical sections can be skipped by the statistically less sophisticated reader, and the reader familiar with the ideas of the analyses can go straight to the graphical representation to find out how to set up the software input.

A LONGITUDINAL STUDY OF GROWTH IN MATHEMATICS ACHIEVEMENT

Sponsored by the National Science Foundation, The Longitudinal Study of American Youth (LSAY) is a national study of performance in and attitudes towards science and mathematics. LSAY uses a national probability sample of about 100 public schools, testing an average of about 50 students per school. It was conducted as a longitudinal survey of two cohorts spanning grades seven to twelve. The first wave of data gathering was carried out in 1987 and each testing occasion took place in the Fall. Background information was collected from parents, teachers, and school principals. In this article, we will focus on mathematics achievement. LSAY uses mathematics items from the 1986 National Assessment of Educational Progress (NAEP). The test measures mathematics skills in a number of subtopics including algebra, probability and statistics, geometry, measurement, and arithmetic. To properly capture growth in achievement over grades, the test items that are administered differ in part across grades and across students within grades. An adaptive testing strategy was used in order to avoid floor and ceiling effects and to maximize the information obtained on the students' achievement level. Given the performance at a certain grade, an easy, medium, or hard test form was chosen for the next grade. The test forms also differed across grades within difficulty designation. The various test forms do, however, have many items in common so that achievement scores from the various forms can be equated. This equating was carried out via Item Response Theory.

Figure 1 shows the means of the mathematics achievement scores for grades 7, 8, 9, and 10 (the younger cohort, called Cohort 2 in LSAY) and for grades 10, 11, and 12 (the older cohort, called Cohort 1 in LSAY). Here, results are presented by gender (Table 1 gives further detail). It is seen that females start out ahead of males in the Fall of the 7th grade, but that males overtake females in higher grades. The differences are small, however. Considering for example grade 12, the standard deviation for males is about 13 and for females it is about 11 while the mean is only one unit higher for males than for females. In fact, the gender differences in means are not significant for any of the grades. The total increase in the mean over the six grades is about 14 for males, i.e. about one 12th grade standard deviation. For

FIGURE 1
Mathematics achievement in grade 7–12 in the longitudinal study of American youth.

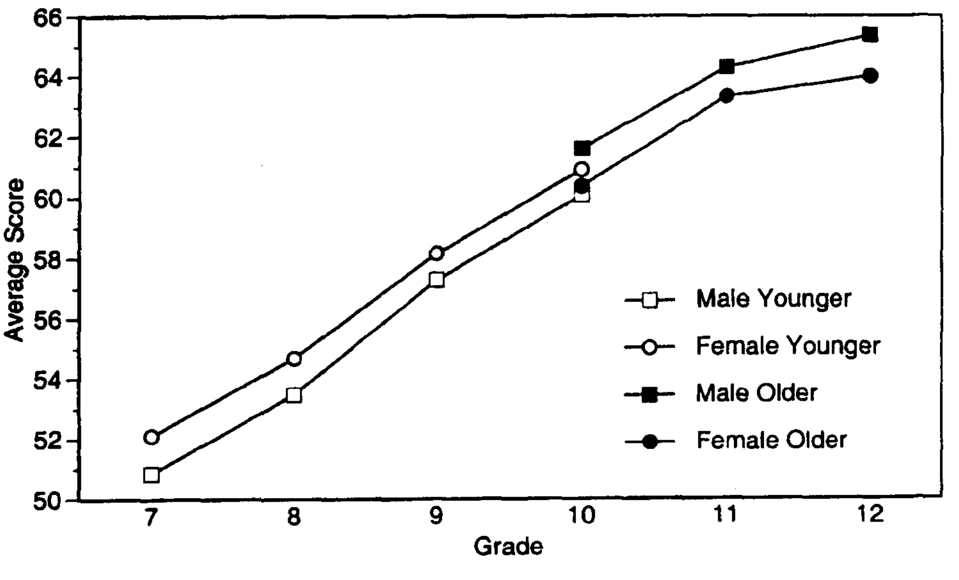


TABLE 1
Sample Statistics

Means and Standard Deviations

	Younger Cohort				*Older Cohort*			
	Male (N =1393)		*Female (N=1331)*		*Male (N=1111)*		*Female (N=1209)*	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Y7	50.87	10.31	52.12	9.31				
Y8	53.49	11.12	54.70	9.53				
Y9	57.30	12.34	58.17	11.05				
Y10	60.09	14.02	60.92	12.20	61.61	11.65	60.37	9.82
Y11					64.29	12.14	63.34	9.94
Y12					65.35	12.55	63.68	10.69
ME	2.42	1.07	2.31	1.01	2.36	1.01	2.35	1.01
HR	3.32	1.80	3.02	1.51	3.16	1.60	2.90	1.41

Correlations (off-diagonals) and Standard Deviations (diagonals)

Younger Male						
Y7	10.32					
Y8	0.851	11.12				
Y9	0.801	0.857	12.34			
Y10	0.751	0.772	0.812	14.02		
ME	0.289	0.250	0.261	0.259	1.07	
HR	0.375	0.343	0.337	0.334	0.247	1.80
Younger Female						
Y7	9.31					
Y8	0.841	9.53				
Y9	0.812	0.849	11.05			
Y10	0.761	0.788	0.818	12.20		
ME	0.318	0.278	0.297	0.300	1.01	
HR	0.326	0.310	0.313	0.297	0.262	1.51
Older Male						
Y10	11.65					
Y11	0.890	12.14				
Y12	0.859	0.897	12.55			
ME	0.296	0.278	0.284	1.01		
HR	0.327	0.291	0.289	0.194	1.60	
Older Female						
Y10	9.82					
Y11	0.844	9.94				
Y12	0.802	0.853	10.69			
ME	0.234	0.254	0.242	1.01		
HR	0.331	0.322	0.319	0.212	1.14	

females the total mean increase is 12, also about one 12th grade standard deviation. The graph of Figure 1, however, only shows the average performance. It does not show the amount of individual variation in the achievement growth. It is therefore not clear how much overlap there is in male and female growth curves.

The figure also does not show how the growth curves for males and females vary as a function of their background characteristics. It is the aim of growth modeling to make these further investigations.

CONCEPTUAL DESCRIPTION OF GROWTH MODELING

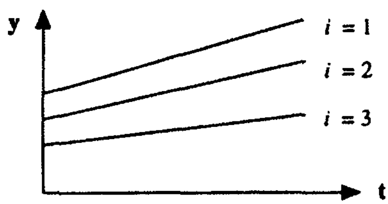
A basic idea behind growth modeling is that individuals differ in their growth over time. This notion is highly relevant to the LSAY example of mathematics achievement. Students are likely to show differences in growth as a function of differences in background characteristics such as gender and home environment. Furthermore, the mathematics curriculum is quite varied in the U.S. and students are likely to show differences in growth due to differences in course taking as well.

Two main parts of the description of individual differences in growth will be considered. First, individuals are likely to differ with respect to their performance at the first testing occasion, mainly due to experiences preceeding this testing occasion. This will be referred to as individual differences in initial sta-

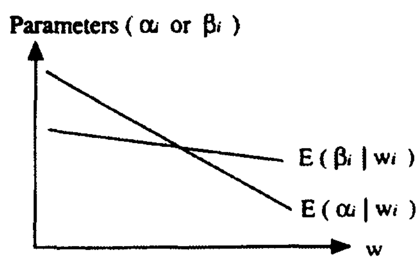
FIGURE 2
Growth modeling in terms of random coefficients and a multilevel model.

Individual i at time t:

(1) $y_{it} = \alpha_i + \beta_i t + \varepsilon_{it}$



(2) $\alpha_i = \alpha + \gamma_\alpha w_i + \delta_{\alpha i}$
(3) $\beta_i = \beta + \gamma_\beta w_i + \delta_{\beta i}$



tus. In the statistical section, this part of the description will use the statistical term of random intercepts. Second, individuals are likely to differ with respect to the growth in their performance across the testing occasions, referred to as individual differences in the rate of growth. Growth can be described as a linear trend and in this case the statistical term of random slopes will be used. Growth deviating from a straight line is, however, important and will also be considered here. Figure 2 illustrates the idea of individual growth trajectories. The top part of the figure shows the development over time for three individuals.

Key modeling results are estimates of the average initial status, the average growth rate, and estimates of the variation across individuals of initial status and of growth rate. For each testing occasion, time-specific factors also influence the performance so that a certain performance that is expected of an individual is in fact not realized. In statistical terms, these factors are described as residuals.

Growth curve analysis is particularly useful when an attempt is made to explain the individual variation in initial status and growth rate using background variables for the individual. These variables are viewed as causes of growth preceeding the testing occasion and do not vary across time. Such variables are of substantive interest in that they are predictors of the growth. The bottom part of Figure 2 shows how the initial status (α) and growth rate (β) are described as functions of a (time-invariant) background variable (w). Figure 2 is discussed in the statistical section below. More elaborate analysis may also attempt to account for the fact that the progression along an individual's growth curve may be hampered or enhanced by time-specific background variables.

Growth studies often use longitudinal data from more than one group of individuals as with the two cohorts of LSAY. These two cohorts both include measurements at grade 10. The older cohort has grade 10 data from the first year of the study, 1987, while the grade 10 data for the younger cohort was collected three years later, in 1990. A cohort design of this type is often used to be able to describe growth across all grades, 7–12, assuming that the two samples come from the same population. This means that any differences due to the two groups being three years apart are taken to be ignorable. For example, it assumes that grade 10 performance for the younger cohort in 1990 is statistically equivalent to grade 10 performance for the older cohort in 1987; the student composition is taken to be the same and the school environment is taken to be the same. This assumption can be the subject of a special investigation. The baseline assumption is that growth in performance is to be described over six grades, assuming two samples drawn from a single population. Given that the younger cohort provides data only from the four lower grades and the older cohort provides data only from the three higher grades, each of the two samples are viewed as having missing data from some of the six grades. The missingness is completely random and missing data techniques can be applied.

STATISTICAL DESCRIPTION OF GROWTH MODELING

The statistical developments to be drawn on are termed random coefficient models. They go beyond conventional structural equation modeling of longitudinal data and its focus on auto-regressive models (see e.g., Wheaton, Muthén, Alwin, & Summers 1977). This is in line with theory provided in biostatistics (see e.g., Laird & Ware 1982) and extended to latent variables in psychometrics (Meredith & Tisak 1984, 1990). The latent variable developments have been applied to psychology by McArdle and Epstein (1987), to education by Muthén (1993) and Willett and Sayer (1994), and to epidemiology by Muthén (1983, 1991). In structural modeling terms, we will see that the model involves both a mean and a covariance structure. The growth model is a multilevel model in that an individual's observations over time are correlated. Correspondingly, a second-level part to the model describes individual variation in growth parameters in terms of person-specific, time-invariant covariates.

In terms of the LSAY math achievement example, consider an achievement score y_{it} for individual i at time point t. Here, t corresponds to the different grades. It is convenient to set t=0 for the lowest grade. The growth model is specified as

$$y_{it} = \alpha_i + \beta_i t + \zeta_{it} \tag{1}$$

Here, α_i and β_i are individual-specific parameters describing initial level of achievement and rate of achievement growth, respectively, and ζ_{it} represent time-varying residuals. The regression intercept and slopes are random parameters that vary over individuals. The top part of Figure 2 illustrates this equation for three individuals.

In any given application, it is not necessary that both the intercept and the slope are random, but one of the two can be fixed. For example, repeated measurement, random effects ANOVA customarily uses a random intercept only model to describe the correlations among the observations over time for a given individual.

The specification of linear growth is not necessary. As will be seen below, the degree of non-linearity in the growth curve can be estimated whenever there are data on a sufficient number of time points. For the models we will propose, at least four time points are desirable.

Extending this model to include time-invariant covariates w_i, the individual variation in these parameters is specified as

$$\alpha_i = \alpha + \gamma_\alpha w_i + \delta_{\alpha i}, \tag{2}$$

$$\beta_i = \beta + \gamma_\beta w_i + \delta_{\beta i}. \tag{3}$$

If the mean of w is zero, α and β represent average intercept and slope parameter values, respectively, whereas more generally they are intercepts in the regressions on w. The γs are regression slope parameters, and δ's represent residuals. The residuals are allowed to be correlated. The bottom part of Figure 2 illustrates a case where w_i has negative effects (negative γs) on the intercepts and slopes. The top part of the figure shows a positive correlation between the intercepts and slopes, perhaps largely due to both being influenced by w.

The model may be further expanded by adding a time-varying covariate v_{it} to the growth curve of (1), introducing time-specific deviations from the growth curve,

$$y_{it} = \alpha_i + \beta_i t + \gamma_t v_{it} + \zeta_{it}, \tag{4}$$

Assuming for simplicity that there are no time-varying covariates v, the model can be seen to imply growth in means and variances as a function of t,

$$E(y_{it} \mid w_i) = \alpha + \gamma_\alpha w_i + (\beta + \gamma_\beta w_i)t \tag{5}$$

$$V(y_{it} \mid w_i) = \sigma_\alpha^2 + 2t\sigma_{\alpha\beta} + t^2\,\sigma_\beta^2 + \sigma_\zeta^2 \tag{6}$$

In general, the model imposes a structure on the means. For example, in a linear growth model without covariates, only two parameters (α and β) are used to explain the progression of observed means.

The growth model can be viewed as a structural equation model with latent variables. The α_i and β_i can be viewed as latent variables instead of random parameters in the sense that both are unobserved i.i.d (independently and identically distributed) variables varying across individuals. In the type of application considered here, t does not vary over individuals because all students are in the same grade at a given testing occasion. In this way, t in (4) can be considered a fixed regression parameter (or factor loading) for the variable β_i. This parameter can in fact be estimated when fixing the first two t values, thereby capturing nonlinear growth.

GENERAL MODEL FRAMEWORK

This latent variable model fits into the following conventional structural equation modeling framework, letting the observed variables y_{it}, w_i, and v_{it}, be stacked in the vector y and the latent variables of α_i and β_i be stacked in η,

$$y = v + \Lambda\eta + \varepsilon, \tag{7}$$

$$\eta = \alpha + B\eta + \zeta, \tag{8}$$

where v and Λ contain measurement intercept and loading (slope) parameters, respectively, and ε denotes a vector of measurement errors. The α and B contain structural regression intercepts and slopes, respectively, and ζ denotes a vector of residuals. With $E(\eta) = \alpha$, $V(\varepsilon) = \Theta$, $V(\zeta) = \Psi$, usual assumptions give the mean and covariance structure for the y vector as

$$\mu = v + \Lambda(I - B)^{-1}\alpha \tag{9}$$

$$\Sigma = \Lambda(I - B)^{-1}\Psi(I - B)^{-1'}\Lambda' + \Theta \tag{10}$$

Assuming multivariate normality for the vector y, maximum-likelihood (ML) estimation is carried out by minimizing the conventional ML fitting function

$$\sum_{p=1}^{P} \left\{ N_p[\ln|\Sigma_p| + tr(\Sigma_p^{-1}T_p) - \ln|S_p| - r] \right\} N^{-1}, \tag{11}$$

where

$$T_p = S_p + (\bar{y}_p - \mu_p)(\bar{y}_p - \mu_p)' \tag{12}$$

In maximum-likelihood (ML) estimation of conventional structural equation models with latent variables, this is the fitting function corresponding to independent random samples from P populations with sample sizes N_p and total sample size N. Here, an r-dimensional vector y, say, is observed with sample covariance matrix S_p, sample mean vector $\bar{y}_p$, population covariance matrix Σ_p, and population mean vector μ_p. The items containing $ln\,|S_p| - r$ are offsets so that a perfectly fitting model has the function value of zero. The sample covariance matrices S_p are the ML estimates of the unrestricted Σ_p matrices and are therefore divided by N_p, not $N_p - 1$. Multiplying the minimum value for any model by $2 \times N$ then gives the value of the likelihood-ratio chi-square test of the H_0 model against the H_1 model of unrestricted mean vectors μ_p and covariance matrices Σ_p. Many models do not impose any restrictions on μ_p in which case the second term on the right-hand-side of (12) vanishes and only covariance matrices are involved in the estimation. The simultaneous analysis of several populations is considered when the populations have parameters in common, so that equality constraints of parameters across populations are invoked.

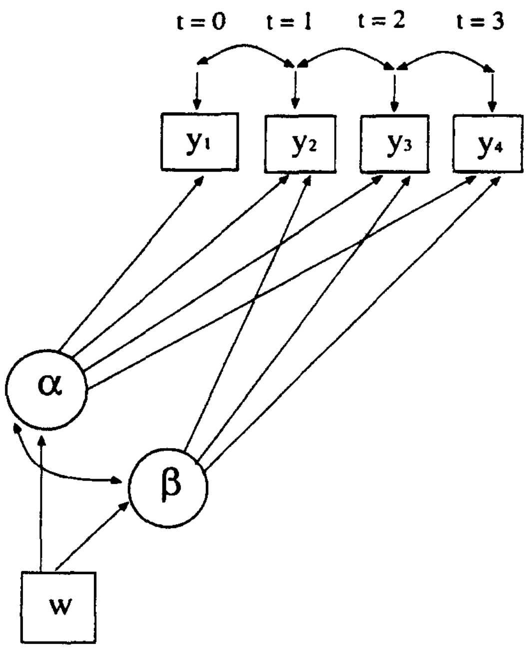
In a simple growth model, only one population is involved and $P = 1$. In an analysis of multiple cohorts assumed to be sampled from a single population the $P > 1$ feature is used (see below). In this case, a single population is still assumed and the use of $P > 1$ is merely a technical solution to carrying out the analysis. With multiple subgroups such as gender, $P > 1$ even in a single cohort. In this case, the fact that $P > 1$ reflects a genuine multiple-population analysis. The growth model imposes a structure not only on the covariance matrix but also the mean vector so that both terms of (12) are involved in the estimation. It is clear from (7) that this modeling framework also encompasses multiple indicators of latent variable constructs.

Consider now the case of analyzing two cohorts as in the LSAY data. We view this as a missing data situation; for theory on missing data analysis, see e.g. Little and Rubin (1987). We consider the data from the two cohorts as two i.i.d samples from the same population, where the younger cohort has missing data in the later grades and the older cohort has missing data in the earlier grades. The data is missing completely at random due to the sampling design. The total sample consists of n_y observations from the younger cohort and n_o observations from the older cohort. Assume for simplicity that there are no covariates. Let the vector of y variables for the two cohorts be denoted y_y and y_o, respectively. The vector y_y has length four (for grades 7–10) and the vector y_o has length three (for grades 10–12). This means that the log likelihood for the total sample can be written as

$$\log L = \sum_{i=1}^{n_y} \log\phi(y_{yi}; \mu, \Sigma) + \sum_{j=1}^{n_o} \log\phi(y_{oj}; \mu, \Sigma) \tag{13}$$

where ϕ denotes a multivariate normal density. Note that the population mean vector μ and the population covariance matrix Σ are the same for the two cohorts with the exception that different elements in these two arrays are operative in the two densities (corresponding to the first four variables in the younger cohort and the last three variables in the older cohort). Allison (1987) and Muthén, Kaplan, and Hollis (1987) showed that from the point of view of structural equation modeling software, the two terms of the log likelihood in (13) can be incorporated in a two-population analysis using the fitting function of (11) with $P = 2$.

FIGURE 3
Graphical representation of a growth model for four time points.



GRAPHICAL DESCRIPTION OF GROWTH MODELING IN LATENT VARIABLE FORM, MODEL STRUCTURE, AND ANALYSIS SPECIFICATIONS

It is convenient to view initial status and growth rate as latent variables. In this way, conventional path diagrams of structural modeling can be used to show the growth model graphically (squares represent observed variables and circles represent latent variables). This graphical representation can be directly translated into conventional structural equation modeling software. Furthermore, it can be easily generalized in many ways, for example analysis with multiple indicators at a given time point and analysis with categorical indicators.

Figure 3 gives an example with test scores y from four time points (e.g. four grades) with one time-invariant covariate w. This is the same model as described above in the statistics section; see also Figure 2. The figure does not show the growth curves, but the sources of variation that influence the growth curves. The latent variable of initial status is denoted α (intercept) and the latent variable of growth rate is denoted β (slope).

MODEL STRUCTURE

The growth model is not only a covariance structure model, but also a mean structure model. It is therefore useful to consider the specification of this model in two parts: terms contributing to the means of the observed variables and terms contributing to variances and covariances among these variables.

Consider first the mean structure. The intercepts in the regressions of the y's on the two latent variables are parameters which should be held equal across time to reflect the fact that the same variable is measured at all time points (e.g a test score in the same metric). The mean of the latent variable α (initial status) is fixed at zero (when more than one population is analyzed, mean differences across the populations can be estimated). With a single indicator as in Figure 3, the y intercept parameter may alternatively be viewed as the mean of the initial status factor, but with multiple indicators it is not convenient to estimate this parameter as the mean of the initial status factor. The growth in observed variable means over time is captured by the latent variable β having a free mean to be estimated. This will be referred to as the growth rate in results tables. The model adds this mean to the expected value of the y variable at a certain time point using the factor t, that is multiplying the β latent variable mean by either 0, 1, 2, or 3 for the four time points. When the model includes a w variable (as in Figure 3), the free mean of the β variable is achieved by estimating a free intercept in the regression of the β variable on w. The mean of the covariate w is free (the covariate part of the model is unrestricted and can be fixed at observed sample values).

Consider next the variance-covariance structure. The growth model specifies that the regression coefficients of the y's on these two latent variables are not all free to be estimated, but some have prespecified values. The growth model specifies that the coefficient for the regression of a certain y on the α variable is fixed at one for all time points. For the regression of each y on the β variable the values of the coefficients are also fixed with values taken as the convenient scaling 0, 1, 2,...for the variation in t over time.

The regressions of the two latent variables on w are to be estimated. The residual variances and covariance for α and β are to be estimated. The variances of the time-specific residuals for the y's are also to be estimated. Usually, some form of across-time covariance for these residuals is needed to represent the data, e.g. allowing different correlations among adjacent time points. The w covariate variance (and covariances if there are more than one covariate) is free to be estimated (the covariate part of the model is unrestricted and can be fixed at observed sample values).

When no covariates are present (no w's or v's), the above linear growth modeling for four time points gives a model with two restrictions and thefore a chi-square test with two degrees of freedom. It can be seen that these two restrictions come from the mean structure of the model in that the four means are explained in terms of only two parameters. With four time points the covariance structure part is just-identified (not restricted) in this approach to growth modeling. This implies that a misfitting model suggests that linear growth is not realistic for the data.

Linear growth is established by fixing the coefficients (the scores for the time variable t) for the latent variable β (the slope) at 0, 1, 2, 3. Using these growth scores means that the growth increments from one occasion to the next are $\beta \times 1$.

FIGURE 4
Graphical representation of a growth model for four time points with covariates.

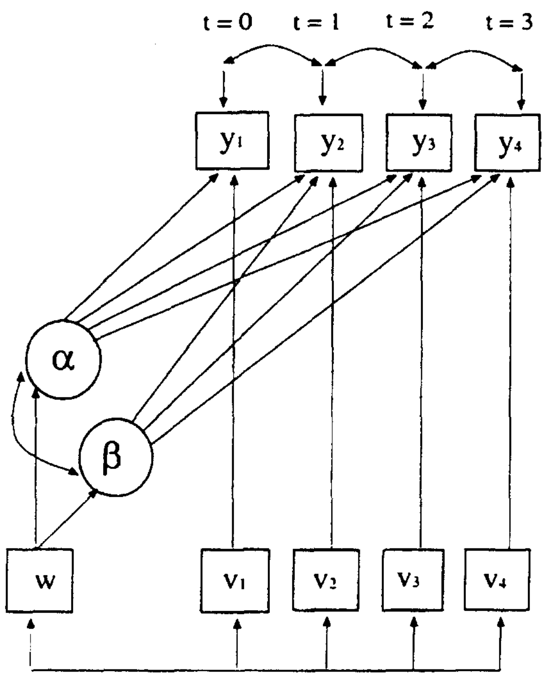
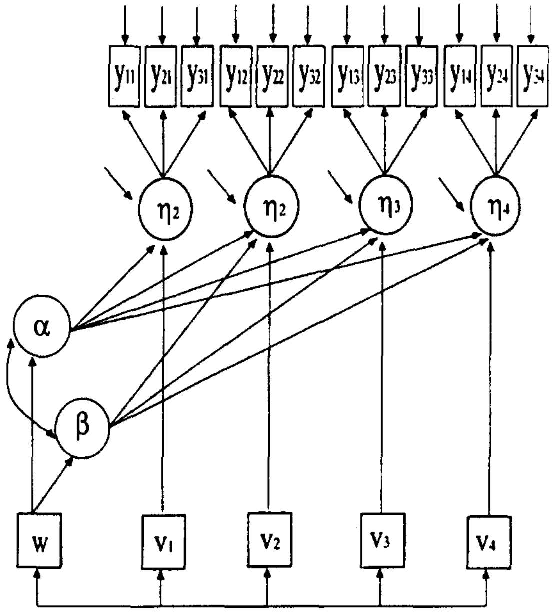


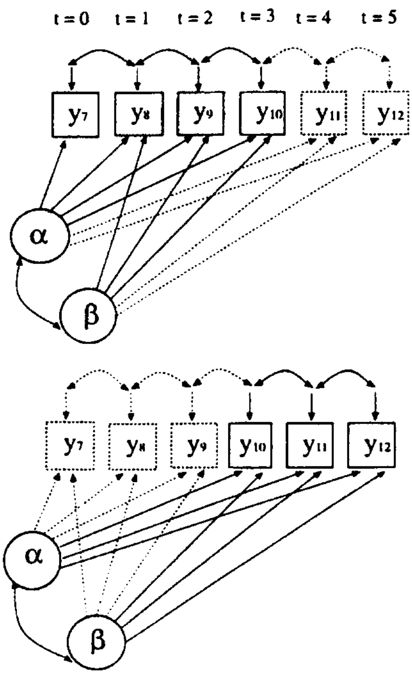
FIGURE 5
Growth model for four time points with multiple indicators and covariates.



Non-linear growth can be estimated by instead using scores 0, 1, t_{12}, t_{23}, where t_{12} and t_{23} are growth score parameters to be estimated. If for example t_{12} is greater than 2, growth is larger between time points 2 and 3 than between time points 1 and 2. Non-linear growth can also be accomplished by using scores generated by a non-linear function such as a logistic growth curve or an exponential decline curve.

It is useful to also consider the model structure when only three time points are available. Consider for simplicity the case with no covariates present (no w's or v's). With linear growth, there is one over-identifying restriction on the three y means because there are only two parameters specific to the mean structure, the common intercept in the regressions of the y's and the mean of the growth rate. The covariance part has six elements. When the y residuals are allowed to be correlated as above, these six elements are expressed in terms of eight parameters. The covariance part is therefore not identified but there are two indeterminacies among the eight parameters. As an example of this indeterminacy, we note that both residual covariances among the y's and the covariance between the initial status and growth rate factors contribute to the covariances between the y's. Adding w covariates will not alter this. Two restrictions need to be applied to the eight parameters. For example, the residual covariances between the y's may be fixed to zero, but if this is not correct, the six remaining parameters will be misestimated.

FIGURE 6
Graphical representation of a two-cohort growth model with missing data.



This dilemma illustrates the usefulness of having at least four time points in growth modeling.

Time-varying covariates v can be incorporated as exemplified in Figure 4. The v's in this figure may for example correspond to the amount and type of course work the student has experienced in the time period preceeding the testing occasion. The influence on the means of the y's that such covariates have may explain deviations from an otherwise linear growth function.

The observed variables y can be replaced by latent variable constructs η, each of which has multiple indicators y. This is a useful approach to avoiding biasing effects of measurement error in y. Such a model is shown in Figure 5.

SEVERAL COHORTS; MISSING DATA

Analysis of both cohorts simultaneously can be carried out using missing data specifications as follows. Figure 6 describes the situation graphically in the case where there are no covariates. There are six time points totally (t = 0, 1, 2, 3, 4, 5). The top part of the figure shows the younger cohort with achievement scores y for grades 7, 8, 9, and 10. Scores for grades 11 and 12 are not observed (missing variables). The bottom part of the figure shows the older cohort with achievement scores y for grades 10, 11, 12. Scores for grades 7, 8, and 9 are not observed (missing variables).

In the missing data two-group analysis, it must be recognized that the two samples are assumed to be drawn from a single population. For the younger cohort the parameters are the variances and covariance for the latent variables α and β, the residual variances and covariances for the residuals of the y's, the common intercept in the regression of the y's on the two latent variables, and the mean of the latent variable β. In total, this is 12 parameters. There are a total of four means and ten variances/covariances for the four observed y variables. The younger cohort therefore has two more observed pieces of information than the number of parameters and would give a two-degree of freedom chi-square test if analyzed alone. Including the older cohort in the joint analysis of the two cohorts introduces only four new parameters: the two residual variances for grades 11 and 12 and the two residual covariances for the last three grades. The other parameters are equal to those of the younger cohort: the variances and covariance for the latent variables α and β, the residual variance for grade 10, the common intercept in the regression of the y's on the two latent variables, and the mean of the latent variable β. In total then, six equality constraints need to be applied in the two-group analysis. As explained in Muthén et al. (1987), the missing variables in the two cohorts can be handled by using dummy entries in the sample mean vectors and covariance matrices and by fixing parameters specific to these parts. Two analyses need to be performed in order to get the correct chi-square test of the growth model. The H_0 model is the one discussed above. The H_1 model is a model where no underlying latent variable structure is imposed, but where equality constraints are imposed for mean vector and covariance matrix elements that the two cohorts have in common (these correspond to the grade 10 y variable in Figure 6). Note that in this two-cohort example, the resulting degrees of freedom is not the same as would have been obtained had there been a single cohort observed across all six time points (no missing data). This is because none of the two cohorts have observed covariances between the y's of the three lowest grades (7, 8, 9) and the y's of the two highest grades (11, 12). The number of degrees of freedom is reduced by the number of missing covariances which would otherwise have added to the number of H_1 parameters.

The missing data approach to the analysis of the two cohorts assumes that the two groups are random samples from the same population. There may, however, be substantively meaningful exceptions from this assumption. For example, main parameters related to the growth curve may be assumed to be the same for both groups while less central parameters related to the marginal distribution of the covariates or the residuals of the y's may differ. In such cases, both the H_0 and H_1 models need to relax the corresponding equality constraints.

SEVERAL POPULATIONS

It may also be of interest to simultaneously analyze growth in several populations. For example, males and females may be seen as representing different growth curves. With two cohorts, this would lead to a four-group analysis allowing for various hypotheses of invariance across gender.

To test the degree of gender invariance of the growth model formally, the following series of analysis steps are useful. In a first analysis, full invariance across gender of the growth model parameters is imposed. This gender invariance model can be tested against the model of no gender invariance by subtracting the sum of the degrees of fredom and chi-square values obtained in the two gender-specific analyses. If non-invariance is found, the sources of non-invariance may be investigated in the following three steps. First, we may relax invariance for the marginal part of the model consisting of the covariates because the growth model does not concern itself with this part. Second, we may relax invariance of the growth model residual variances (variances remaining when conditioning on the covariates), namely the intial status and growth rate residuals and the variances of the achievement score residuals. Third, we may relax invariance of the growth model's conditional means given the covariates, namely the growth rate intercept and the y intercept.

SOFTWARE IMPLEMENTATION

The growth modeling can be carried out using latent variable structural equation modeling computer software. The senior author's program Mplus (Muthén & Muthén, 1998) makes this implementation particularly simple. Once a model has been layed out in a path diagram such as the one in Figure 3, the analysis specifications can be given without referring to matrices or equations.

EXAMPLES OF GROWTH MODELING USING LSAY MATHEMATICS ACHIEVEMENT DATA

Mathematics data from LSAY will now be analyzed by the growth model using the latent variable formulation and maximum-likelihood estimation in conventional structural equation modeling software. A series of increasingly more complex models will be estimated. A theme in the analyses is the study of gender differences. Because of this, males and females will be analyzed separately. To begin with, the two cohorts will also be analyzed separately, leading to a final, joint analysis.

The sample statistics for the four groups defined by gender and cohort are given in Table 1. This table also includes sample statistics for the covariates that will be used. As time-invariant (w) covariates we will use the student's mother's education and a measure of home resources of an academic nature. For simplicity in the exposition, no time-varying covariates (v) will be used.

ANALYSIS OF THE YOUNGER COHORT: NO COVARIATES

The younger cohort was tested in grades 7, 8, 9, and 10. In the first analysis, we will examine a simple growth model with only achievement scores (the y's) and no covariates. The sample sizes are 1,393 for males and 1,331 for females.

TABLE 2
Non-Linear Growth for the Younger Cohort: No Covariates

	Male			Female		
n	*1393*			*1331*		
df	*0*			*0*		
	Estimate	*SE*	*T-value*	*Estimate*	*SE*	*T-value*
y Intercept	50.87	0.10	500.61	52.12	0.09	560.05
(initial status mean)						
Growth Scores/Steps						
Grade 8/7–8	1.00*/1.00*	-/-	-/-	1.00*/1.00*	-/-	-/-
Grade 9/8–9	2.46/1.46	0.12/0.12	20.08/11.97	2.35/1.35	0.11/0.11	21.26/12.27
Grade 10/9–10	3.53/1.07	0.20/0.11	18.01/9.82	3.41/1.06	0.17/0.10	20.02/11.16
Growth Rate Mean	2.62	0.16	16.68	2.58	0.14	18.19
Variance of						
Initial Status	87.21	7.12	12.26	77.5	5.65	13.72
Growth Rate	1.61	0.94	1.71	0.67	0.77	0.87
Covariance of						
Initial Status, Growth Rate	6.05	2.23	2.72	2.66	1.72	1.55
Residual Variance of *y*						
Grade 7	19.25	6.34	3.04	9.07	4.80	1.89
Grade 8	22.72	3.24	7.00	7.00	2.57	2.73
Grade 9	25.70	4.15	6.12	29.97	3.21	8.71
Grade 10	46.42	7.71	6.02	45.76	5.75	7.96
Residual covariance of *y*						
Grade 7, Grade 8	4.35	4.07	1.07	-5.79	3.15	-1.84
Grade 8, Grade 9	5.00	1.31	4.29	1.01	0.98	1.03
Grade 9, Grade 10	3.05	4.74	0.64	12.07	3.64	3.31

Note: *Fixed.

TABLE 3
Non-Linear Growth for the Younger Cohort: Two Time-Invariant Covariates

	Male			Female		
n	*1393*			*1331*		
df	*4*			*4*		
chi-sq	*9.32*			*16.81*		
p-value	*0.053*			*0.002*		
	Estimate	*SE*	*T-value*	*Estimate*	*SE*	*T-value*
y Intercept	50.88	0.10	498.94	52.12	0.09	557.85
(initial status mean)						
Growth Scores/Steps						
Grade 8/7–8	1.00*/1.00*	-/-	-/-	1.00*/1.00*	-/-	-/-
Grade 9/8–9	2.48/1.48	0.13/0.13	19.82/11.84	2.39/1.39	0.12/0.12	20.76/12.07
Grade 10/9–10	3.58/1.10	0.20/0.11	17.77/9.82	3.48/1.09	0.18/0.10	19.55/11.12
Growth Rate Mean	2.58			2.53		
(from growth rate equation)						
Initial Status Regression						
Mother's ED	1.93	0.24	8.10	2.09	0.23	9.00
Home Resources	1.83	0.14	12.91	1.57	0.16	10.13
Growth Rate Regression						
Intercept	2.58	0.16	16.41	2.53	0.14	17.68
Mother's ED	0.12	0.06	1.96	0.17	0.06	2.87
Home Resources	0.09	0.04	2.52	0.10	0.04	2.42
Estimated Variance of						
Initial Status	91.33			78.05		
Growth Rate	2.58			1.52		

Residual Variance of						
Initial Status	72.84	6.20	11.75	65.40	5.02	13.04
Growth Rate	2.53	0.84	3.02	1.45	0.69	2.12
Residual Covariance of						
Initial Status, Growth Rate	3.41	1.91	1.78	1.08	1.52	0.71
R^2						
Initial Status	0.20			0.16		
Growth Rate	0.02			0.04		
Residual Variance of *y*						
Grade 7	14.27	5.58	2.56	7.02	4.34	1.62
Grade 8	23.34	2.94	7.97	9.29	2.38	3.90
Grade 9	24.78	3.91	6.34	25.51	3.03	8.33
Grade 10	38.19	7.33	5.21	37.28	5.45	6.85
Residual covariance of *y*						
Grade 7, Grade 8	2.61	3.61	0.72	-5.37	2.88	-1.87
Grade 8, Grade 9	6.68	1.26	5.28	1.75	0.95	1.84
Grade 9, Grade 10	-0.79	4.50	-0.18	7.03	3.46	2.03

Note: *Fixed.

The linear growth model does not fit well, especially not for males (the chi-square values with two degrees of freedom are 21.5, p=0.000, and 13.5, p=0.001, respectively for males and females). As previously discussed, the source of this misfit should not be sought in the covariance structure, but in the mean structure. The misfit suggests that linear growth is not realistic. We allow for non-linear growth by allowing growth scores to be estimated for grades 9 and 10 instead of fixed at the values 2 and 3. While keeping the growth step from grade 7 to 8 fixed at one, this means that the steps from grade 8 to 9 and from grade 9 to 10 are allowed to be different from one. The model is then just-identified (has zero degrees of freedom) and a chi-square test is not available (the model fits the data trivially). In this model, the means for the y variables as estimated from the model are in fact the same as the sample means (in practice, they are the same only to two digits). The estimates from this model are presented in Table 2.

It is interesting to compare the estimated growth scores to the t (time) score sequence for linear growth in grades 7-10: 0, 1, 2, 3. For both males and females this shows that there is an acceleration in growth after the eighth grade testing occasion. The estimated growth steps given in Table 2 are computed as the difference between growth scores for adjacent grades. This shows that for both males and females there is a larger growth between the grade 8 and grade 9 testing occasions than between grades 7 and 8 and grades 9 and 10.

The growth rate has only a slightly higher point estimate for males than for females, 2.62 versus 2.58. To better understand the magnitude of the estimated growth rate it is useful to look at its impact on the achievement mean increase from one grade to the next as estimated by the model. This is obtained as the product of the growth rate and the growth step. Consider for example the increase from grade 8 to 9. The product gives the estimated achievement mean increase 3.83 for males and 3.48 for females. These values can be compared to the grade 9 achievement standard deviation of 12.34 for males and 11.05 for females. In grade 9 standard deviation units, the increases are 31% and 32%, respectively, for males and females.

Table 2 shows a significant amount of individual variation in initial status for both males and females, indicating that seventh graders are a heterogeneous group in terms of math performance. The estimated variation is larger for males than for females, but not significantly so. The individual variation in the growth rate is not significant for males or for females implying that individuals of a certain gender grow at the same rate. For both males and females initial status and growth rate have a positive correlation, 0.51 for males and 0.37 for females (assuming that females do vary in their growth rate), so that students who start off high grow faster, as expected.

ANALYSIS OF THE YOUNGER COHORT: TWO TIME-INVARIANT COVARIATES

The next analysis step attempts to explain the individual variation in the initial status and growth rate by means of two time-invariant covariates (w's): mother's education and home resources. The non-linear growth specification found in the first analysis step is maintained. Including the two covariates, the model is no longer just-identified but can be tested by chi square. The model now has four degrees of freedom, where the four restrictions imposed by the model arise because the influence from the two covariates to the four math achievement scores is mediated by the two latent variables of initial status and growth rate so that eight covariances are expressed by four parameters. This model achieves a marginally acceptable fit for males (chi square of 9.32, d.f. = 4, p=0.053) but not for females (chi square 16.81, d.f. = 4, p=0.007). It may be noted that replacing mother's education with an SES composite including father's education and income, fits the data better (chi square for males is 6.28, p=0.179; chi square for females is 6.57, p=0.160). It appears that mother's education has influence on performance beyond what is explained by the growth curve, particularly for females. Allowing one direct effect from mother's education to the math achievement score at one time point decreases chi square signficantly for females, particularly if the direct effect is for achievement at grade 7. The main parameters, however, obtain about the same estimates. This elaboration of the model will not used here, but results from the more parsimonious model will be reported. The estimates from this model are given in Table 3.

An interesting finding here is that for both males and females, mother's education and home resources both have positive influence on the student's initial status as well as the student's growth rate. In terms of the growth rate of males, it may be noted that replacing mother's education with SES did not achieve significant influence on the 5% level (z value of 1.39).

The two covariates explain 20% of the variance in the initital status for males and 16% for females. For the growth rate, however, only 2% of the variance is explained for males and only 4% for females, showing that many more well-chosen covariates are needed to explain the growth rate variation across students. A somewhat surprising finding is that this model, as opposed to the previous one without covariates, does describe the male and female growth rate as varying across individuals.

ANALYSIS OF THE OLDER COHORT: TWO TIME-INVARIANT COVARIATES

The older cohort was tested in grades 10, 11, 12. The same two w covariates as for the younger cohort are considered here. The sample sizes are 1,111 for males and 1,209 for females.

Due to the fact that only three time points are available for this cohort, the linear growth model including correlated residuals among the achievement scores is not identified as previously discussed. Two restrictions need to be applied to the parameters to make the model identified. For example, one may restrict the two residual covariances to zero. This will not be pursued here for two reasons. First, the analysis of the younger cohort does not support the assumption of zero residual covariances. Second, this assumption is not necessary given that the older cohort can be analyzed jointly with the younger cohort in which case this identification problem does not arise.

TABLE 4
Two-Cohort Simultaneous Analysis: Two Time-Invariant Covariates

	Male			Female		
n	*1393 and 1111*			*1331 and 1209*		
H_0	*chisq=57.72 (df=57, p=0.448)*			*chisq=54.38 (df=57, p=0.574)*		
H_1	*chisq=45.85 (df=48, p=0.561)*			*chisq=22.57 (df=48, p=0.999)*		
difference test	*chisq=11.87 (df=9, p=0.221)*			*chisq=31.81 (df=9, p=0.0002)*		
	Estimate	*SE*	*T-value*	*Estimate*	*SE*	*T-value*
y Intercept	51.24	0.08	658.81	52.00	0.07	717.30
(initial status mean)						
Growth Scores/Steps						
Grade 8/7–8	1.00*/1.00*	-/-	-/-	1.00*/1.00*	-/-	-/-
Grade 9/8–9	2.61/1.61	0.13/0.13	20.66/12.78	2.56/1.56	0.13/0.13	19.86/1209
Grade 10/9–10	3.90/1.29	0.20/0.11	19.07/11.32	3.75/1.19	0.20/0.11	19.06/11.33
Grade 11/10–11	5.00/1.10	0.27/0.09	18.38/11.83	5.02/1.27	0.28/0.10	18.18/12.33
Grade 12/11–12	5.43/0.43	0.30/0.07	18.28/5.97	5.31/0.29	0.30/0.07	18.00/4.02
Growth Rate Mean	2.49			2.30		
(from growth rate equation)						
Initial Status Regression						
Mother's ED	2.04	0.21	9.87	1.78	0.19	9.32
Home Resources	1.80	0.12	14.63	1.63	0.13	12.57
Growth Rate Regression						
Intercept	2.49	0.14	17.61	2.30	0.13	17.36
Mother's ED	0.13	0.05	2.85	0.10	0.04	2.48
Home Resources	0.05	0.03	1.86	0.09	0.03	2.96
Estimated Variance of						
Initial Status	91.76 (89.83)†			76.57 (75.93)†		
Growth Rate	3.74 (0.47)†			3.00 (0.19)†		

Residual Variance of						
Initial Status	73.94	5.63	13.14	65.31	4.86	13.45
Growth Rate	3.71 (0.44)†	0.57 (0.54)	6.48 (0.82)	2.97 (0.16)†	0.49 (0.45)	6.10 (0.35)
Residual Covariance of						
Initial Status, Growth Rate	2.44	1.49	1.64	0.27	1.30	0.21
R^2						
Initial Status	0.19(0.18)†			0.15(0.14)†		
Growth Rate	0.01(0.07)†			0.01(0.17)†		
Residual Variance of y						
Grade 7	11.62	4.73	2.46	4.63	4.01	1.16
Grade 8	26.76	2.82	9.50	13.30	2.38	5.59
Grade 9	18.76	3.57	5.26	17.35	2.83	6.13
Grade 10	17.44	2.41	7.23	16.44	1.93	8.52
Grade 11	19.42	3.45	5.63	14.27	2.95	4.84
Grade 12	23.90	3.43	6.97	25.97	2.80	9.27
Residual covariance of y						
Grade 7, Grade 8	3.05	3.35	0.91	-4.32	2.88	-1.50
Grade 8, Grade 9	7.53	1.24	6.07	2.33	0.96	2.44
Grade 9, Grade 10	-12.03	2.78	-4.43	-5.62	2.20	-2.56
Grade 10, Grade 11	2.90	1.80	1.61	0.11	1.45	0.08
Grade 11, Grade 12	5.86	2.80	2.09	4.24	2.45	1.73

Note: *Fixed; †value for the older cohort in parentheses.

ANALYSIS OF THE YOUNGER AND OLDER COHORT: SIMULTANEOUS ANALYSIS WITH TWO TIME-INVARIANT COVARIATES

In the final analysis step, a single, non-linear growth model is applied to both cohorts jointly. It is assumed that the two cohorts are random samples from the same population, with growth determined by the same parameters. As discussed, the chi-square test for this model is obtained as the difference between the H_0 growth model and the H_1 model with no structure on the means and (co-)variances except across-cohort invariance. It was found, however, that already the H_1 model was rejected, implying that the two cohorts cannot be seen as random samples from a single population (with 50 d.f.'s the chi square and p value were 97.82 and 0.000 for males and 82.96 and 0.001 for females). The H_1 model rejection appeared to mainly be due to two causes. First, the variance of grade 10 performance is larger for the younger cohort than the older cohort as seen in Table 1. This may because more 10th grade students in 1990 (the younger cohort at the fourth measurement occasion) than in 1987 (the older cohort at the first measurement occasion) have access to more advanced courses. Second, the variance of the covariate Home Resources is larger for the younger cohort than for the older cohort. Perhaps this is due to a larger amount of measurement error in the reporting of this variable by the younger group of students. Relaxing these two equality restrictions resulted in a good fit of H_1 (with 48 d.f.'s the chi square value was 45.85 for males and 22.56 for females). Given this H_1 testing outcome, the H_0 model was modified accordingly in two ways. First, the two cohorts were allowed to have different variances for the Home Resources variable. Furthermore, the cohort differences in grade 10 variance was seen to continue through grades 11 and 12 (see Table 1) indicating a cohort effect on the variances for all three achievement scores. Because variance development over time is modelled by the growth rate variance, the second modification was therefore to allow the two cohorts to have different residual variance for growth rate. The corresponding H_0 and H_1 difference test of model fit indicated good fit for males (with nine degrees of freedom the chi square difference value was 11.87, p= 0.221) but not for females (with nine degrees of freedom the chi square difference value was 31.82, p= 0.000). The significant value of the difference test for females may be merely due to the low chi square value of the H_1 model, perhaps indicating that the H_1 model has unnecessarily many parameters for females. The H_0 model test value for females is of the same magnitude as for males (57.72 for males and 54.38 for females) indicating that the model may be equally adequate in the two groups. The results for this H_0 model are presented in Table 4.

The results of Table 4 appear rather similar for males and females. The non-linear growth model again shows that growth is accelerated between Fall of eighth grade and Fall of ninth grade, while slowed down when going into tenth grade. The information added by the older cohort shows that the growth from grade ten to eleven is about equal to that of grade nine to ten and that the growth from grade eleven to twelve is less than half of that. As with the analysis of the younger cohort, for both males and females, mother's education and home resources both have positive influence on the student's initial status as well as the student's growth rate. Home Resources does not, however, achieve significance influence on the 5% level for the growth rate of males (the z value is 1.86).

Given that Table 4 presents rather similar results for males and females, it is of interest to test gender invariance of the growth model formally. This can be done in a simultaneous analysis of the four groups defined by gender and cohort. The analysis steps previously suggested are used. The Table 4 model with no gender invariance fits well and serves as a baseline model (we may use the sum of the two H_0 chi-square tests, 112.10 with 114 d.f.'s, as baseline comparison values). In a first analysis of the degree of gender invariance, full invariance across gender of the Table 4 growth model parameters was imposed. Testing this model against the model with no invariance leads to rejection of full gender invariance of the growth model (chi square value is 406.95 with 145 d.f.'s; chi square difference test value comparing with the no invariance model is 297.85 with 31 d.f.'s, p=0.000). To investigate the sources of gender non-invariance, the following three steps were taken. First, we relaxed invariance for the marginal part of the model consisting of the covariates. This improved model fit, but the model still fit significantly worse than the model with no invariance (chi square with 139 d.f.'s is 288.57; chi-square difference comparing to the no invariance model is 176.47 with 25 d.f.'s). Second, we also relaxed invariance of the growth model residual variances (variances remaining when conditioning on the covariates), namely the initial status and growth rate residuals and the variances of the achievement score residuals. This improved model fit, but the model still fit significantly worse than the model with no invariance (chi square with 124 d.f.'s is 136.48; chi-square difference comparing with the no invariance model is 24.38 with 10 d.f.'s, p=0.007). Gender differences in both types of residual (co-)variances were indicated. Third, we also relaxed invariance of the growth model's conditional means given the covariates, namely the growth rate intercept and the y intercept (or, equivalently in this model, the initial status mean). While a y intercept gender difference was not evident, a gender difference in the growth rate intercept was. The model also allowing for gender differences in the growth rate intercept (but not the y intercept) fit well and did not fit significantly worse than the no invariance model (chi square with 123 d.f.'s is 127.42; chi-square difference comparing to the no invariance model is 15.32 with 9 d.f's, p=0.083). This is the final model shown in Table 5. This invariance testing sequence shows that the important gender differences in growth lie in the variances of the initial status and growth rate factors and in the mean of the growth rate factor. It is particularly important to note the finding that the y intercepts show gender invariance while the growth rate mean, by virtue of the growth rate intercept, is significantly higher for males. This says that males and females do not differ significantly in their achievement level when going in to seventh grade, but significantly stronger male achievement growth develops over time. As seen from the estimated growth rates, however, the difference is not large. In terms of the estimated mean increase from grade 8 to 9, the increase is 4.01 for males and 3.59 for females with grade 9 achievement standard deviations of 12.34 and 11.05, respectively.

TABLE 5
Four-Group Simultaneous Analysis: Two Time-Invariant Covariates

H_0	*chisq=127.42 (df=123. p=0.374)*					
H_1	*chisq=112.10 (df=114, p=0.533)*					
difference test	*chisq=15.32 (df=9, p=0.083)*					
n	*Male 1393 and 1111*			*Female 1331 and 1209*		
	Estimate	*SE*	*T-value*	*Estimate*	*SE*	*T-value*
y Intercept	51.663	0.05	974.09	invariant		
(initial status mean)						
Growth Scores/Steps						
Grade 8/7–8	1.00*/1.00*	-/-	-/-	1.00*/1.00*	-/-	-/-
Grade 9/8–9	2.57/1.57	0.09/0.09	28.73/17.44	invariant		
Grade 10/9–10	3.81/1.23	0.14/0.08	27.01/16.04	invariant		
Grade 11/10–11	5.00/1.19	0.19/0.07	25.91/17.35	invariant		
Grade 12/11–12	5.35/0.35	0.21/0.05	25.71/6.93	invariant		
Growth Rate Mean	2.49			2.32		
(from growth rate equation)						
Initial Status Regression						
Mother's ED	1.89	0.14	13.48	invariant		
Home Resources	1.72	0.09	19.32	invariant		
Growth Rate Regression						
Intercept	2.49	0.11	23.59	2.32	0.10	23.89
Mother's ED	0.12	0.03	3.72	invariant		
Home Resources	0.07	0.02	3.36	invariant		
Estimated Variance of						
Initial Status	89.11 (87.35)†			78.18 (77.45)†		
Growth Rate	3.80 (0.49)			2.94 (0.14)†		

Residual Variance of						
Initial Status	73.08	5.70	12.83	65.55	4.78	13.70
Growth Rate	3.77 (0.49)†	0.52 (0.56)	7.23 (0.88)	2.91 (0.11)†	0.41 (0.43)	7.06 (0.26)
Residual Covariance of						
Initial Status, Growth Rate	2.60	1.54	1.68	0.33	1.26	0.26
R^2						
Initial Status	0.18 (0.16)†			0.16 (0.15)†		
Growth Rate	0.01 (0.06)†			0.01 (0.2)†		
Residual Variance of *y*						
Grade 7	12.22	4.81	2.54	4.69	3.91	1.20
Grade 8	27.04	2.82	9.59	13.72	2.36	5.81
Grade 9	18.95	3.43	5.53	17.18	2.71	6.30
Grade 10	18.04	2.41	7.50	15.97	1.91	8.37
Grade 11	17.68	3.60	4.92	14.35	2.82	5.10
Grade 12	22.88	3.45	6.63	26.68	2.76	9.68
Residual covariance of *y*						
Grade 7, Grade 8	3.36	3.38	0.99	-4.08	2.83	-1.44
Grade 8, Grade 9	7.69	1.23	6.24	2.45	0.95	2.59
Grade 9, Grade 10	-11.66	2.68	-4.36	-5.92	2.13	-2.78
Grade 10, Grade 11	2.40	1.79	1.34	-0.78	1.44	-0.12
Grade 11, Grade 12	4.50	2.94	1.53	4.59	2.34	1.96

Note: *Fixed; †value for the older cohort in parentheses.

DISCUSSION

As the analyses of the mathematics achievement example have illustrated, growth modeling provides interesting ways to discover and describe individual differences in development over time. In contrast, the plot of averages for each grade shown in Figure 1 merely gives an aggregated view of achievement growth over time. Figure 1 did not show significant mean differences between males and females at any grade, but the growth model was nevertheless able to show significant gender differences in growth rate. The growth model is able to quantify the amount of individual differences in growth and relate the individual variation in growth to background variables.

Growth modeling in a latent variable analysis framework makes possible a very clear and flexible analysis. The chi-square testing procedure clearly highlights the assumptions that are imbedded in the model for example when assuming linear growth, when adding covariates, or when analyzing several cohorts and population subgroups. The model framework allows for easy generalizations of the basic growth model such as using multiple indicators and measures that are not continuous-normal, e.g. binary variables (see e.g., Muthén, 1983). Analysis approaches for randomized experiments are discussed in Muthén and Curran (1997). Analysis of cluster data giving rise to three-level modeling is described in Muthén (1997).

The fact that the latent variable approach utilizes mean and covariance structure modeling may, however, also lead to potential misuse. First, investigators may neglect to scrutinize their raw data with respect to such features as the shape of individual growth curves and outliers. Second, in some cases there may be competing models which fit the means and covariances roughly the same but may lead to different data interpretations.

Certain classic problems of growth modeling are, of course, still present in the latent variable analysis framework. For example, the problem of scale changes over time due to changes in content is relevant in the mathematics example. While Item Response Theory equating is used to put the achievement test items onto the same scale, the content emphasis of the test changes over time. In earlier grades arithmetic items dominate the achievement score while in later grades algebra and geometry items are also present. Interpretations of growth over time is problematic if the score does not have the same meaning across time and if growth over time is not homogeneous with respect to these different content areas. These potential problems are perhaps reduced when the aim is to compare growth across population subgroups such as in the present analysis of males and females.

Many further growth topics can be discussed in the latent variable analysis framework. Two will be merely mentioned here. One topic is the estimation of individual growth curves. Given an estimated growth model, such curves can be estimated by Empirical Bayes techniques. In the present framework this translates into estimation of factor scores, a topic which has a large literature in the latent variable modeling context. Another important topic is prediction of growth. For a given model, it is of interest to use an individual's background information to predict the individual's future growth.

ACKNOWLEDGMENTS: This research was supported by a grant from the Office of Educational Research and Improvement, Department of Education to the National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Graduate School of Education, UCLA. The authors with to thank Ginger Goff and Booil Jo for their excellent assistance.

REFERENCES

Allison, P. D. (1987). "Estimation of linear models with incomplete data." In *Sociological Methodology*, edited by C. Clogg. San Francisco: Jossey-Bass.

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

Little, R. J. & D. B. Rubin. (1987). *Statistical analysis with missing data*. New York: Wiley & Sons.

McArdle, J. J. & D. Epstein, D. (1987). "Latent growth curves within developmental structural equation models." *Child Development, 58*, 110–133.

Meredith, W. & J. Tisak. (1984). Tuckerizing curves. Paper presented at the Psychometric Society annual meetings, Santa Barbara, CA.

______. (1990). "Latent curve analysis." *Psychometrika, 55*, 107–122.

Muthén, B. (1983). "Latent variable structural equation modeling with categorical data." *Journal of Econometrics, 22*, 43–65.

Muthén, B. (1991). "Analysis of longitudinal data using latent variable models with varying parameters." Pp. 1–7 in *Best methods for the analysis of change; recent advances, unanswered questions, future directions*, edited by L. Collins & J. Horn. Washington DC: American Psychological Association.

______. (1993). "Latent variable modeling of growth with missing data and multilevel data." Pp. 199–210 in *Multivariate analysis: Future directions 2* , edited by C.R. Rao & C. M. Cuadras. Amsterdam: North-Holland.

______. (1997). "Latent variable modeling of longitudinal and multilevel data." Pp. 453–480 in *Sociological Methodology 1997*, edited by A. Raftery. Washington DC: American Sociological Association.

Muthén, B., D. Kaplan, & M. Hollis. (1987). "On structural equation modeling with data that are not missing completely at random." *Psychometrika, 42*, 431–462.

Muthén, L. & B. Muthén. (1998). *Mplus User's Guide.*

Wheaton, B., B. B. Muthén, D. Alwin, & G. Summers. (1977). "Assessing reliability and stability in panel models." Pp. 84–136 in *Sociological Methodology 1977* , edited by D.R. Heise. San Francisco: Jossey-Bass.

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