Modeling and Testing Change: An Introduction to the Latent Growth Curve Model

Article *in* Understanding Statistics · August 2003

DOI: 10.1207/S15328031US0203_02

Barbara M. Byrne
University of Ottawa

UNDERSTANDING STATISTICS, *2*(3), 177–203

TEACHING ARTICLES

Modeling and Testing Change: An Introduction to the Latent Growth Curve Model

Barbara M. Byrne and Gail Crombie
School of Psychology
University of Ottawa

The purposes of this article are threefold: (a) to outline the basic concepts associated with latent growth curve (LGC) modeling; (b) to demonstrate the modeling and testing of LGC models based on three relatively simple, albeit increasingly complex, examples; and (c) to illustrate the modeling mechanism used in testing for the tenability of key statistical assumptions associated with LGC modeling. Based on 3-wave data comprising an original sample of 601 adolescents (Grades 8, 9, and 10) and using a multiple-sample approach that takes into account missing data resulting from time-related attrition, we "walk" the reader through the various stages of the model specification and testing processes. Based on self-rating scores of perceived ability as the outcome variable, we begin with a single-domain LGC model of perceived math ability and then follow with a more complex multiple-domain model that includes perceived ability in math, language, and science. Our final application extends the multiple-domain model to include the predictor variable of gender. We conclude by summarizing several advantages of LGC modeling over the more traditional methods used in the measurement of change.

Keywords: measuring change, latent growth curve modeling, structural equation modeling

Behavioral scientists have long been intrigued with the investigation of change. Questions such as, "Do the rates at which children learn differ in accordance with

Requests for reprints should be sent to Barbara M. Byrne, School of Psychology, University of Ottawa, Ottawa, Ontario, Canada K1N 6N5. E-mail: bmbch@uottawa.ca

their interest in the subject matter?" or "Are changes in teachers' levels of burnout related to administrative decision making policy?" exemplify this type of inquiry. Answers to questions of change necessarily demand repeated measurements on a sample of individuals at multiple points in time. Historically, researchers have typically based analyses of change on two-wave panel data, a strategy that Willett and Sayer (1994) deemed to be inadequate because of limited information. As Willett and Sayer noted, "When true development follows an interesting trajectory over time, 'snapshots' of status taken before and after are unlikely to reveal the intricacies of individual change" (p. 363). Addressing this weakness in longitudinal research, Willett (1988) and others (Bryk & Raudenbush, 1987; Rogosa, Brandt, & Zimowski, 1982; Rogosa & Willett, 1985) outlined methods of individual growth modeling that, in contrast, capitalized on the richness of multiwave data, thereby allowing for more effective testing of systematic interindividual differences in change.

In a novel extension of this earlier work, researchers (e.g., McArdle & Epstein, 1987, Meredith & Tisak, 1990; Muthén, 1991) have shown how individual growth models can be tested using the analysis of mean and covariance structures within the framework of structural equation modeling (SEM). Considered within this context, it has become customary to refer to such models as latent growth curve (LGC) models. Given its many appealing features (for an elaboration, see Willett & Sayer, 1994), together with the ease with which researchers can tailor its basic structure for use in innovative applications (see, e.g., Chou, Bentler, & Pentz, 1998; Duncan & Duncan, 1995; Hancock, Kuo, & Lawrence, 2001; Li, Duncan, & Duncan, 2001), it seems evident that LGC modeling has the potential to revolutionize analyses of longitudinal research. Consequently, we believe that it behooves researchers of longitudinal data to familiarize themselves with this important analytic strategy. Our intent here is to assist in this learning process by using a didactic approach to LGC modeling that enables us to "walk" readers through the various analytic stages involved.

In broad terms, then, the purpose of this article is to present an introductory-level overview of LGC modeling. Specifically, the aims are threefold: (a) to outline the basic concepts associated with LGC modeling; (b) to demonstrate the modeling and testing of LGC models based on three relatively simple, albeit increasingly complex, examples; and (c) to illustrate the modeling mechanism used in testing for the tenability of key statistical assumptions associated with LGC modeling.

MEASURING CHANGE IN INDIVIDUAL GROWTH OVER TIME: THE GENERAL NOTION

In answering questions of individual change related to one or more domains of interest, a representative sample of individuals must be observed systematically over time and their status in each domain measured on several temporally spaced occasions (Willett & Sayer, 1994). However, several conditions may also need to be met. First, the outcome variable representing the domain of interest must be of a continuous scale.[1] Second, although the time lag between occasions can be either evenly or unevenly spaced, both the number and the spacing of these assessments must be the same for all individuals. Third, when the focus of individual change is structured as a LGC model, with analyses to be conducted using a SEM approach, data must be obtained for each individual on three or more occasions. Finally, the sample size must be large enough to allow for the detection of person-level effects (Willett & Sayer, 1994). Moreover, when analyses entail SEM, a methodology grounded in large-sample theory that assumes data to be multivariate normal, sample size requirements become even more critical. Accordingly, one would expect minimum sample sizes of not less than 200 at each time point (see Boomsma, 1985; Boomsma & Hoogland, 2001).

In presenting a logically ordered introduction to the topic of LGC modeling, we structured the content of this article within the framework of several sections. We begin first by describing the example data on which our three applications are based. Next, we introduce the hypothesized model to be tested in the first application. This model comprises a single-domain output variable (Perceived Ability in Math) and provides the springboard for our discussion of two submodels that serve as the basic building blocks of the LGC model. Willett and Sayer (1994) termed these underpinning submodels as "Level 1" and "Level 2" models. The Level 1 model can be thought of as a "within-person" regression model that represents individual change over time with respect to (in this instance) a single outcome variable, Perceived Ability in Math. The Level 2 model can be viewed as a "between-person" model that focuses on interindividual differences in change with respect to the outcome variable. Following explanations bearing on the underlying concepts and specification of these two submodels, we then test for the validity of the total hypothesized LGC model and report the findings.

In the next section, we move to our second application (Model 2), a multiple-domain LCG model that builds on our original model by including two additional outcome variables—Perceived Ability in Language and Perceived Ability in Science. We close out this section by demonstrating how to test for the tenability of three important assumptions associated with LGC models: (a) the linearity (or nonlinearity) of individual growth trajectories over time, (b) the independence of,

[1]One reviewer noted that at least one recent SEM program (MPlus; Muthén & Muthén, 1998–2001) is capable of analyzing LGC models based on noncontinuous data. As such, a probit regression is considered for the outcomes as a function of both the growth factors and the covariates. This means that growth is modeled for individual differences in the development over time based on the probability of a certain outcome.

and (c) the homoscedasticity of measurement error variances associated with the outcome variable.[2]

Finally, in the last section, our third application (Model 3) extends the multiple-domain model to include the fixed predictor variable of gender. This series of presentations is considered to be logically ordered because we must be able to describe intraindividual change before moving on to investigate interindividual differences in change. Likewise, we must establish that interindividual differences in change exist before we can ask whether these differences can be related to some predictor variable.

Example Data

The three applications we present in this article are based on three-wave data comprising self-ratings of Perceived Academic Ability in Grades, 8, 9, and 10 for 601 adolescents. More specifically, the data represent subscale scores related to self-perceived ability in math, language, and science. Consistent with most longitudinal research, some participant attrition occurred over the 3-year period. In total, 101 cases were lost, leaving 500 complete-data cases. Unfortunately, the use of complete data resulting from listwise deletion of missing scores assumes that the data are "missing completely at random" (MCAR) and is therefore considered a questionable practice (Little & Rubin, 1987; Muthén, Kaplan, & Hollis, 1987; for more details on types of missing data, see Byrne, 2001; Little & Rubin, 1987; Muthén et al., 1987.) Given (a) the known problems associated with the use of listwise-deleted data; (b) the availability of SEM multiple-sample model specification that can take such missingness into account, yet still provide large-sample estimates (see Duncan & Duncan, 1994, 1995; Muthén et al., 1987); and (c) evidence that even if the MCAR condition is found to be untenable, the missing-data model noted in (b) is likely to reduce bias inherent in the use of listwise-deleted data (Muthén et al., 1987), we considered it most appropriate to employ this multiple-sample, missing-data model. Accordingly, we formulated three groups of participants based on our finding of three patterns of missing data and then placed equality constraints across the missing data groups to obtain unbiased and consistent estimates. Group 1 ($n = 500$) represents participants for whom complete data were available across the 3-year time span. Group 2 ($n = 543$) represents participants for whom data were available only for Years 1 and 2, whereas Group 3 ($n = 601$) represents participants for whom data were available only for Year 1 of the study.[3] Descriptive statistics related to each of the three time points for male and female adolescents are presented in Table 1.

[2]We considered it more instructive and practicable to illustrate these procedures based on the multiple-domain, rather than the single-domain, model.

[3]An elaborated description of this approach to the specification and estimation of a multiple-sample LGC model is beyond the scope of this article. For a more comprehensive explication and discussion, readers are referred to Duncan and Duncan (1994, 1995), Muthén et al. (1989), and Rubin (1976).

TABLE 1
Descriptive Statistics for Perceived Ability in Math, Language, and Science for Male and Female Adolescents Across Time

Variable	*M*	*SD*	*SK*	*KU*	*N*
Male adolescents					
Perceived Math Ability					
Year 1 (Grade 8)	5.292	1.198	–0.715	–0.142	290
Year 2 (Grade 9)	5.185	1.285	–0.766	–0.019	265
Year 3 (Grade 10)	4.888	1.513	–0.709	–0.378	246
Perceived Language Ability					
Year 1 (Grade 8)	4.663	1.110	–0.411	0.174	290
Year 2 (Grade 9)	4.655	1.128	–0.305	0.248	265
Year 3 (Grade 10)	4.581	1.138	–0.258	–0.220	246
Perceived Science Ability					
Year 1 (Grade 8)	4.807	1.200	–0.395	–0.259	290
Year 2 (Grade 9)	4.950	1.270	–0.700	0.336	265
Year 3 (Grade 10)	5.130	1.161	–0.796	0.431	246
Female adolescents					
Perceived Math Ability					
Year 1 (Grade 8)	4.741	1.384	–0.876	0.128	311
Year 2 (Grade 9)	4.769	1.357	–0.730	–0.002	278
Year 3 (Grade 10)	4.635	1.455	–0.636	–0.325	254
Perceived Language Ability					
Year 1 (Grade 8)	5.016	1.098	–0.442	0.119	311
Year 2 (Grade 9)	5.066	1.093	–0.574	0.273	278
Year 3 (Grade 10)	5.129	1.105	–0.732	0.412	254
Perceived Science Ability					
Year 1 (Grade 8)	4.603	1.253	–0.445	–0.149	311
Year 2 (Grade 9)	4.769	1.295	–0.786	0.357	278
Year 3 (Grade 10)	4.866	1.299	–0.859	0.682	254

Note. SK = skewness; KU = kurtosis.

Instrumentation. The Perceived Math Ability subscale, taken from the Student Attitude Questionnaire (see Eccles, 1983; Jacobs & Eccles, 1992), is designed to measure children's beliefs and attitudes about mathematics. Measurements of Perceived Ability in Language and Science were based on items from the Math subscale, albeit the term *math* was modified as *language arts* and *science,* respectively; the items are anchored with values of 1 (*not at all good, the worst,* and *much worse*) and 7 (*very good, the best,* and *much better*). The three 7-point Likert scaled items comprising each subscale assess students' perceptions of how good they are at math (language, science); (a) in an absolute sense, (b) relative to other students in their class, and (c) relative to other school subjects. Moderately high internal consistency reliability (*Mdn* $\alpha = .84$) for the math subscale has been reported (Eccles, Adler, & Meece, 1984; Fuligni, Eccles, & Barber, 1995; Jacobs & Eccles, 1992;

Parsons, Adler, & Kaczala, 1982). For this sample, similarly high median coefficient alpha values, over the 3-year period, were obtained for the math (.91), language (.88), and science (.92) subscales, and test–retest reliability, measured over a 1-month period for a randomly selected subsample of 62 students, was found to be .91, .88, and .92, respectively.

Procedure. Working within a classroom setting, trained researchers administered the questionnaires to small groups of 10 to 25 students during the months of April and May over a 3-year period comprising Grades 8, 9, and 10, respectively. Data collection at this time ensured that students were at least 80% through the school year and thus had a good sense of their ability in each of the targeted subject areas.

APPLICATION 1: A SINGLE-DOMAIN LGC MODEL (MODEL 1)

Modeling Intraindividual Change

The first step in building a LGC model, as noted previously, is to examine the within-person growth trajectory. In our case here, this task translates into determining, for each individual, the direction and extent to which his or her score in self-perceived Ability in Math changes from Grade 8 through Grade 10. Of critical import in most appropriately specifying and testing the LGC model, however, is that the shape of the growth trajectory be known a priori. If the trajectory of hypothesized change is considered to be linear, then the specified model will include two growth parameters: (a) an intercept parameter representing an individual's score on the outcome variable at Time 1 and (b) a slope parameter representing the individual's rate of change over the time period of interest. Within the context of our work here, the intercept represents an adolescent's Perceived Ability in Math at the end of Grade 8; the slope represents the rate of change in this value over the 3-year transition from Grade 8 through Grade 10. If, on the other hand, the growth trajectory were considered to be nonlinear, the hypothesized model would then include a third parameter representing curvature. For purposes of didactic simplicity, we assume the trajectory to be linear and then later (in the final section of Application 2) illustrate how to test for the validity of this assumption.

Of the many advantages in testing for individual change within the framework of a structural equation model over other longitudinal strategies, two are of primary importance. First, this approach is based on the analysis of mean and covariance structures and as such, distinction is made between observed and unobserved (or latent) variables in the specification of models. Second, the methodology allows for both the modeling and estimation of measurement error. With these

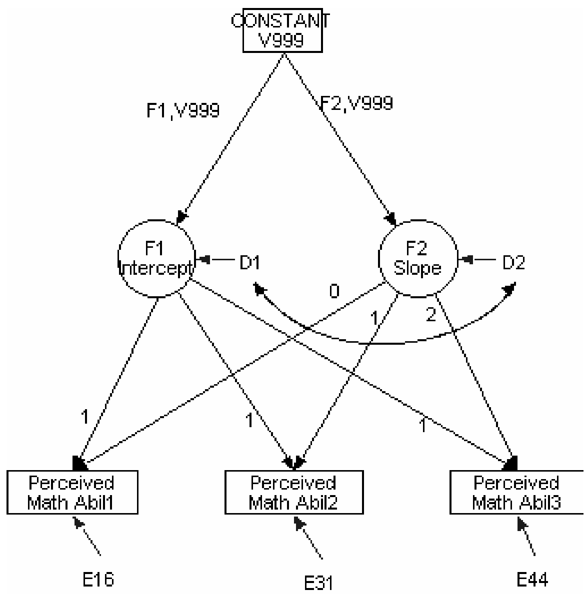

FIGURE 1 Hypothesized single-domain latent growth curve model of Perceived Math Ability (Model 1).

basic concepts in hand, we turn now to Figure 1 where the model to be tested in this first application is schematically presented.

For readers who may not be familiar with the symbols associated with structural equation models, a brief explanation is in order. By convention, circles (or ellipses) represent unobserved factors, squares (or rectangles) represent observed variables, single-headed arrows (→) represent the impact of one variable on another, and double-headed arrows (↔) represent covariances or correlations between pairs of variables. In building a model for study, researchers use these symbols in one of four basic configurations, each of which represents an important component in the analytic process. We now examine these four configurations as they relate to the LGC model (Model 1) shown in Figure 1.

We focus first on the three rectangles at the bottom of the path diagram. The labels "Perceived Math Abil1" through "Perceived Math Abil3" represent the observed scores on Perceived Math Ability collected at each of three time points. The "E" associated with each of these observed variables represents random measurement error.[4] The two variables enclosed in circles represent the un-

[4]The labeling system applied in Figures 1, 2, and 3, is consistent with the notation of the EQS program. Accordingly, the numeral associated with each error term is consistent with the location of the related variable in the data set. For example, the variable "Perceived Math Ability" is the 16th variable in the data set; hence an error term labeled *E16.*

observed Intercept and Slope factors. Within the context of Model 1, they are designated as Factor 1 (F1) and Factor 2 (F2), respectively. The "D" associated with each of these factors represents a disturbance (or residual) term. In more general models, these variables represent residual error associated with the regression of unobserved factors on other unobserved factors in the model. Specific to the LGC model, however, these residuals represent individual differences in the intercept and linear growth trajectories, respectively. (In the interest of clarity and simplicity, the term *residual* shall be used throughout the remainder of this article.) Finally, the rectangle labeled "Constant" represents the specially designated variable, V999, within the framework of the EQS program (Bentler, 2000). Rooted in the Bentler–Weeks (Bentler & Weeks, 1980) structural equation representation system, this variable provides the mechanism by which a covariance structure is transformed into a mean and covariance (or moment) structure. (The functional importance of this transformation is addressed later in the article.)

We turn now to the symbols used to connect each of these observed and unobserved variables. Although each single-headed arrow represents a regression coefficient, its interpretative meaning varies according to whether or not it leads from a source variable. As such, the arrows leading from F1 and F2 to the three observed variables represent the regression of Perceived Math Ability scores onto both the Intercept and Slope at each of three time points. The arrows (labeled "F1,V999"; "F2,V999") leading from the Constant to each of the Intercept and Slope factors represent the average intercept (or starting point of the growth curve) and average linear growth coefficient, respectively. In contrast, the arrows leading from the Es to the observed variables, and from the Ds to the two Intercept and Slope factors, represent the influence of error as described earlier. The double-headed arrow linking D1 with D2 represents a covariance between these two residuals. The validity of this hypothesized relation is assumed in the specification of an LGC model. Finally, explanation regarding the numerical values assigned to the regression paths leading from the Intercept and Slope factors to the observed variables are explained later.

Recall that our primary focus in this section is to model intraindividual change. Within the framework of SEM, this focus is captured by what is termed the *measurement model,* the portion of a model that incorporates only linkages between the observed variables and their underlying unobserved factors. Of primary interest in any measurement model is the strength of the factor loadings or regression paths linking the observed and unobserved variables. (For a nonmathematical elaboration on fundamental concepts and applications related to SEM, readers are referred to Byrne, 1994, 1998, 2001.) Accordingly, the only parts of the model in Figure 1 that are relevant in the modeling of intraindividual change are the two factors (Slope, Intercept), the regression paths linking the three observed variables to these factors, and the measurement errors associated with the observed variables (Perceived Math Abil1 to Perceived Math Abil3).

We now consider this measurement model within a statistical framework. As such, its various components can be summarized by means of a regression equation. Expressed in matrix format, the measurement model related to Figure 1 is represented by the following regression equation:

(1)

Essentially, this equation states that each adolescent's Perceived Math Ability score, at each of three time points (Time 1 = 0; Time 2 = 1; Time 3 = 2) is a function of three distinct components: (a) a factor loading matrix of constants (1; 1; 1) and known time values (0; 1; 2) that remain invariant across all individuals multiplied by (b) an LGC vector containing individual-specific and unknown growth parameters (Intercept, Slope), plus (c) a vector of individual-specific and unknown errors of measurement. Whereas the LGC vector represents the within-person true change in Perceived Math Ability over time, the error vector represents the within-person noise that serves to erode these true change values (Willett & Sayer, 1994).

In preparation for a transition from the modeling of intraindividual change to the modeling of interindividual change, it is important that we review briefly the basic concepts underlying the analyses of mean and covariance structures in SEM. When population means are of no interest in a model, analysis is based on covariance structure parameters only; as such, all scores are considered to be deviations from their means and thus the constant term (represented as α in a regression equation) equals zero. Given that mean values played no part in the specification of the Level 1 (or within-person) portion of our LGC model, only the analysis of covariance structures was involved. However, in moving to Level 2, the between-person portion the model, interest focuses on mean values associated with the Intercept and Slope factors; these values derive from an analysis of mean structures. Because both levels are involved in the modeling of interindividual differences in change, analyses are based on both mean and covariance structures.

Modeling Interindividual Differences in Change

Level 2 argues that, over and above hypothesized linear change in Perceived Math Ability over time, trajectories will necessarily vary across adolescents as a consequence of different intercepts and slopes. Within the framework of SEM, this portion of the model reflects the "structural model" component that, in general, portrays relations among unobserved factors and postulated relations among their associated residuals. Within the more specific LGC model, however, this structure is limited to the regression paths linking the Constant to the Intercept and Slope factors (F1,V999; F2,V999), along with their related residuals, as reflected in the upper tier of the model shown in Figure 1. Expressed in simple matrix terms, this portion of the model can be summarized as follows:

(2)

The specification of these parameters, then, makes possible the estimation of interindividual differences in change. The key element in this specification is the Constant term because it provides the mechanism for transforming the covariance structure of the measurement (Level 1) model into the mean structure needed for analysis of the structural (Level 2) model. More specifically, in modeling and testing mean and covariance structures, as is the case here, analysis must be based on the moment matrix that is made possible by the inclusion of the Constant in the model specification. Each structural equation program differs in the way it addresses this transformation. In EQS (the program used here), the Constant represents the standard V999 variable, albeit the label "Constant" has been maintained for purposes of consistency in the field (Bentler, 2000). This variable is always taken as an independent variable that has no variance and no covariance with other variables in the model.

Within the usual context of SEM specification, the model in Figure 1 would convey the notion that both the Intercept and Slope factors are predicted by the Constant, but with some degree of error as captured by the residual terms (D1, D2); furthermore, these residuals are hypothesized to covary (D1,D2). However, in the special case of a LGC model, as noted earlier, the residuals call for a somewhat different interpretation. Specifically, they are residuals in the sense that they represent individual differences in the intercept and slope parameters. These differences derive from deviations between the individual growth parameters and their respective population means (or average intercept and slope values).

We now reexamine Equation 2, albeit in more specific terms, to clarify information bearing on possible differences in change across time. Within the context of this first application (Model 1), interest focuses on five parameters that are key to determining between-person differences in change: two factor means (F1,V999; F2,V999), two factor residual variances (D1, D2), and one residual covariance (D1,D2). The factor means represent the average population values for the Intercept and Slope and answer the question, "What is the population trajectory of true change in Perceived Math Ability from Grades 8 through 10?" The factor residuals represent deviations of the individual Intercepts and Slopes from their population means, thereby reflecting population interindividual differences in the initial (Grade 8) Perceived Math Ability scores and the rate of change in these scores, respectively. Addressing the issue of variability, these key parameters answer the question, "Are there interindividual differences in the growth trajectory of Perceived Math Ability in the population?" Finally, the residual covariance represents the population covariance between any deviations in initial status and rate of change and answers the question, "Is there any evidence of interindividual differences in the association between initial status and rate of change in perceived math ability across the time that spans Grade 8 through Grade 10?"

Testing for Interindividual Differences in Change

Hypothesized model. In reviewing the hypothesized LGC model in Figure 1, we observe numerical values assigned to the paths flowing from the Intercept and Slope factors to the observed variables but no such values associated with those linking the Constant to these factors. The paths with assigned values represent fixed parameters that will not be estimated. The "1s" specified for the paths flowing from the Intercept factor to each of the observed variables indicate that each is constrained to a value of 1.0. This constraint reflects the fact that the intercept value remains constant across time for each individual (Duncan, Duncan, Stryker, Li, & Alpert, 1999). The values of 0, 1, and 2 assigned to the Slope parameters represent Years 1, 2, and 3, respectively. These constraints address the issue of model identification, a complex topic that goes beyond the boundaries of this article (for further elaboration of this concept, readers are referred to Bollen, 1989; Byrne, 1994, 1998, 2001); they also ensure that the second factor can be interpreted as a slope. Three important points are of interest with respect to these fixed slope values. First, technically speaking, the first path (assigned a zero value) is really nonexistent and therefore has no effect. Although it would be less confusing to simply eliminate this parameter, it has become customary to include this path in the model, albeit with an assigned value of zero (Bentler, 2000). Second, these values represent equal time intervals (1 year) between measurements; had data collection taken place at unequal intervals, the values would need to be calculated accordingly (e.g., 6 months = .5). Third, although the choice of fixed values assigned to the Slope factor loadings is arbitrary, it is important to realize that the Intercept factor is tied to a time scale (Duncan et al., 1999). Thus, any shift in fixed loading values on the Slope factor will necessarily modify the scale of time bearing on the Intercept factor, which in turn will influence interpretations related to the Intercept factor mean and variance. Finally, the fact that no numerical values have been assigned to the paths flowing from the Constant to each of the factors indicates that these parameters will be freely estimated. These parameters are typically unknown and, as noted earlier, represent the average intercept (or starting point) and average linear growth coefficients.

Statistical analyses. All analyses were conducted using Version 6 of the EQS program (Bentler, 2000). In testing each of the three hypothesized models, goodness of fit to the sample data was based on the comparative fit index (CFI; Bentler, 1990) and the root mean square error of approximation (RMSEA; Browne & Cudeck, 1993). The CFI ranges in value from 0 to 1.00, with a value of .90 serving as the rule-of-thumb, lower limit cut point of acceptable fit. The RMSEA takes into account the error of approximation in the population and asks the question, "How well would the model, with unknown but optimally chosen parameter values, fit the population covariance matrix if it were available?" (Browne & Cudeck, 1993, pp. 137–138). This discrepancy, as measured by the RMSEA, is expressed per degree of freedom, thus making it sensitive to model complexity; values less than .05 indicate good fit, and values as high as .08 represent reasonable errors of approximation in the population. Although we report the χ^2 statistic in our summaries of model fit, its known sensitivity to sample size necessarily precludes its use as an appropriate measure of goodness of fit (see Jöreskog & Sörbom, 1993).

Given (a) the assumption of multivariate normality demanded of SEM methodology, in general, and bivariate normality of latent growth vectors in LGC models, in particular, and (b) the fact that case outliers can often seriously distort model fit, it is always important to scrutinize data prior to any testing of a specified model. Our preanalysis of the data revealed the distribution of scores to be approximately multivariate normal (*M* skewness = –0.78; *M* kurtosis = 0.42; Mardia's normalized estimate = 5.37). In addition, tests using the robust maximum likelihood (ML) estimator available in EQS revealed no difference in overall model fit between the Satorra–Bentler χ^2 statistic (S–B χ^2; Satorra & Bentler, 1988) and the usual χ^2 statistic.[5] Finally, our analyses revealed evidence of one multivariate outlier in Group 1; this case was deleted from all subsequent analyses.

Results. Goodness-of-fit statistics related to Model 1, a single-domain (Perceived Ability in Math) LGC model, revealed an exceptionally well-fitting model,

[5]Mardia's normalized estimate of multivariate kurtosis behaves in large samples as a unit normal variable. Although Duncan and Duncan (1995) recommended a value of 2.00, there is really no known definitive cutoff point for this estimate. However, one means of verifying evidence of multivariate normality is to conduct the analyses twice as described here. The S–B χ^2 serves as a correction for the χ^2 statistic when distributional assumptions are violated. In the event that the data are non-normally distributed, the S–B χ^2 typically will be lower than the usual χ^2 value, and there will be differences between the standard errors of parameters derived from ML estimation and those derived from ML estimation using the robust statistics.

$\chi^2(7) = 7.842$, $p = .35$ (CFI = 1.00; RMSEA = .009). Although the CFI index, based on a larger model, might trigger concerns that the model was overfitted to the data, this was not the case here.[6] Rather, the exceptionally high CFI value is indicative of a well-fitting model but at the same time one that is very simplistic in structure (for similar findings, see, e.g., Duncan & Duncan, 1995, 1999.) For purposes of illustrating the specification and testing of our initial LGC model, this simplicity, of course, is intended.

We turn now to Table 2 in which the estimated values for all parameters of primary interest are reported. The first two parameters, represented in the path diagram as the structural paths leading from the Constant to the Intercept and to the Slope (see Figure 1), describe the average true mean trajectory for Perceived Ability in Math. Whereas the path coefficient associated with the Intercept indicates the average self-perceived Math Ability score at initial status (Grade 8), the path coefficient associated with the Slope represents the average rate of change over Grades 8, 9, and 10; both estimates were statistically significant as indicated by z values > 1.96. These estimates indicate that although the average self-reported score at the end of Grade 8 was 5.066, this score decreased, on average, by .135 in Grade 9 and again in Grade 10. From this information, we can determine that as adolescents progressed from Grade 8 through Grade 10, they perceived themselves as having increasingly less ability in math.

Turning next to the residuals, we focus first on their variances, both of which were statistically significant. Recall that these parameters are residuals in the sense that they represent deviation from the average intercept and slope and as such yield information on interindividual differences. Accordingly, these indicate the presence of important interindividual differences in both Grade 8 self-reports of Perceived Math Ability and in their change over time from Grade 8 through Grade 10. Such evidence provides strong justification for the incorporation of predictor variables in subsequent analyses in an effort to explain this variation. In this article, Model 3 addresses this issue.

The fifth entry in Table 2 (Intercept/slope) represents the covariance between the residuals associated with the intercept and rate of change in Perceived Math Ability across adolescents. As shown in Table 2, this association was found to be –0.329, which was statistically significant. This negative value suggests that adolescents who exhibit low Perceived Math Ability scores in Grade 8 tend to have greater rates of increase from Grade 8 through Grade 10 than adolescents exhibiting high levels of Perceived Math Ability in Grade 8.

[6]The term *overfit* is used in reference to a model for which additional parameters have been specified to improve the overall fit of the model to the data. Typically, these parameters represent trivial and substantively meaningless information (e.g., correlated measurement error variances that are relatively weak and otherwise unsubstantiated).

TABLE 2
Parameter Estimates for Single-Domain Model (Model 1)

Variable	*Parameter*	*Estimate*	*z Value*[a]
Factor means			
Intercept	F1,V999	5.066	160.710
Slope	F2,V999	–0.135	–5.283
Factor residuals			
Intercept	D1	1.456	13.588
Slope	D2	0.283	4.749
Intercept/slope	D2,D1	–0.329	–4.331
Error variances			
Perceived Math Abil1	E16	0.177	1.874
Perceived Math Abil2	E31	0.698	10.912
Perceived Math Abil3	E44	0.937	6.929

Note. $\chi^2(7) = 7.842$; CFI = 1.00; RMSEA = .009.
[a]z values > 1.96 indicate statistical significance ($p < .05$).

Finally, the last three entries in Table 2 (E16, E31, E44) provide estimates of the error variance associated with each occasion of measurement. Although the error terms associated with Years 2 and 3 were statistically significant, the one bearing on Year 1 (Grade 8) was found to be nonsignificant.

APPLICATION 2: A MULTIPLE-DOMAIN LGC MODEL (MODEL 2)

We move on now to our second application, which includes two additional outcome variables—self-perceived Language Ability and self-perceived Science Ability. The focus of our investigation here is to determine if, and to what extent, adolescents' self-perceptions of their ability in Math, Language, and Science change over the period extending from Grade 8 through Grade 10. This multiple-domain LGC model is schematically portrayed in Figure 2.

In reviewing Figure 2, you will quickly note that the model essentially represents an extended version of Model 1 (see Figure 1). Specifically, there are two major differences between Model 1 and Model 2. First, Model 2 specifies a 6-factor rather than a 2-factor structure, with each outcome variable having its own Intercept and Slope factors; relatedly, structural paths now lead from the Constant to each of the three Intercept and Slope factors. Second, consistent with Model 1, the residuals associated with the Intercept and Slope factors for each Perceived Ability construct are assumed to be correlated, as indicated by the double-headed arrows linking the Ds within each domain. Given the detailed description provided in Application 1 regarding the parameters comprising the LGC model, there is no

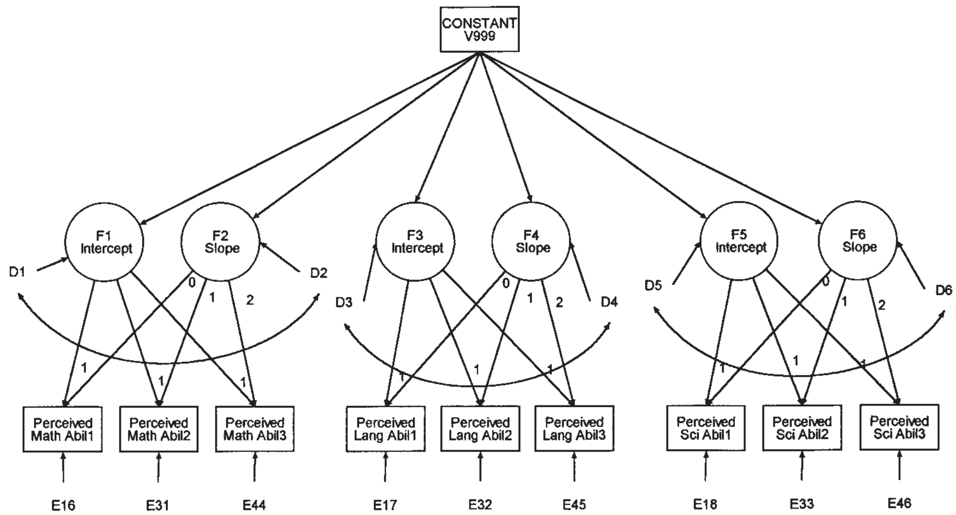

FIGURE 2 Hypothesized multiple-domain latent growth curve model of Perceived Math Ability, Perceived Language Ability, and Perceived Science Ability (Model 2).

need to repeat this information again with Model 2. Thus, we turn next to results derived from the test of this multiple-domain model.

Results

Findings from analysis of this multiple-domain model are presented in Table 3. Goodness-of-fit statistics revealed an excellent fit to the sample data, $\chi^2(63) = 134.255$ (CFI = .962; RMSEA = .026). Nonetheless, a review of the z values associated with the residual covariance between the Intercept and Slope factors representing Perceived Language and Perceived Science Abilities (D3,D4; D5,D6) revealed these parameters to be statistically nonsignificant; in the interest of parsimony, these parameters were subsequently deleted from the model. In addition, a review of the Lagrange Multiplier Test (LMTest) statistics[7] identified two between-domain residual covariances (D1,D5; D3,D5) that, if freely estimated, would lead to a better fitting model in the sense that it would represent the sample data more appropriately. Given the substantive reasonableness of these residual covariances, together with Willett and Sayer's (1996) caveat that in multiple-domain LGC models, covariation among the growth parameters across do-

[7]The LMTest is the procedure used in the EQS program to test, multivariately, the statistical viability of specified parameter constraints in a model.

TABLE 3
Parameter Estimates for Multiple-Domain Model (Model 2)

	Perceived Math Ability			*Perceived Language Ability*			*Perceived Science Ability*		
Variable	*Parameter*	*Estimate*	*z Value*[a]	*Parameter*	*Estimate*	*z Value*	*Parameter*	*Estimate*	*z Value*
Factor means									
Intercept	F1, V999	5.067	160.157	F3, V999	4.867	184.180	F5, V999	4.727	160.805
Slope	F2, V999	–0.136	–5.344	F4, V999	–0.007	–0.302	F6, V999	0.128	5.171
Factor residuals									
Intercept	D1	1.394	14.258	D3	0.536	14.479	D5	0.635	14.393
Slope	D2	0.249	4.435	D4	0.084	3.116	D6	0.094	3.241
Factor residual covariances									
Intercept/slope (within domain)	D2, D1	–0.288	–4.229						
Intercept/slope (between domains)	D1, D5	0.539	16.162	D3, D5	0.249	10.449			
Error variances									
Year 1 (Grade 8)	E16	0.225	2.992	E17	0.669	17.302	E18	0.869	19.103
Year 2 (Grade 9)	E31	0.663	11.279	E32	0.650	11.857	E33	0.790	12.516
Year 3 (Grade 10)	E44	1.003	7.759	E45	0.576	6.251	E46	0.587	5.821

Note. $\chi^2(63) = 134.255$; CFI = .962; RMSEA = .026.

[a] z values > 1.96 indicate statistical significance ($p < .05$).

mains (as reflected in their residual terms) should be considered, Model 2 was modified to address these concerns. Specifically, Model 2 was respecifed such that the two nonsignificant within-domain residual covariances were deleted, whereas the two between-domain residual covariances were freely estimated. Results presented in Table 3 are based on this final respecified model.

Turning, first, to the factor score means, we see that all were statistically significant except for the slope related to Perceived Language Ability. Not surprisingly, the results related to Perceived Math Ability mimic those reported for the single-domain model but with the exception that its error variance was now statistically significant. Findings related to the Intercepts revealed the average scores for both Perceived Language Ability (4.867) and Perceived Science Ability (4.727) were slightly lower than those for Perceived Math Ability (5.067). However, whereas the average change in adolescents' self-perceived Math ability decreased over a 3-year period from Grade 8 to Grade 10 (as indicated by the value of –0.136), the average change in self-perceived Science ability increased (0.128); the change in self-perceived Language ability was negligible (–0.007), as indicated by its statistical nonsignificance.

Turning next to the residual variance associated with the intercept and slope for each Perceived Ability domain (i.e., the Ds), we observed all parameters to be statistically significant. These findings revealed strong interindividual differences in both the initial scores of self-perceived ability in Math, Language, and Science at Time 1 and in their change over time as the adolescents progressed from Grade 8 through Grade 10. As noted in our review of Model 1, such evidence of interindividual differences provides powerful support for further investigation of variability related to the growth trajectories. In particular, the incorporation of predictors into the model can serve to explain their variability.

In reviewing results related to the within-domain residual covariance between the intercept and slope for Perceived Math Ability, we found the estimate, as expected, was statistically significant. As with Model 1, this negative value (–0.288) suggests that for adolescents whose self-perceived scores in math ability were high in Grade 8, their rate of increase in scores over the 3-year period from Grade 8 through Grade 10 was lower than it was for adolescents whose self-perceived math ability scores were lower at Time 1.

Turning to the first Table 3 entry for between-domain residual covariances (D1,D5), we observed a strong association between the intercepts of Perceived Ability in Math and in Science, $r = .57$, a finding that would appear to be quite reasonable. Likewise, the residual covariance between the intercepts for Perceived Language and for Science Ability was also moderately strong, $r = .38$, albeit a somewhat more curious finding than was the link between perceived ability related to Math and Science.

Finally, as is evidenced in Table 3, all error variances were found to be statistically significant. In general, although error variability associated with Perceived

Language and Science Abilities appeared to be relatively consistent across time (i.e., homoscedastic), those associated with Perceived Ability in Math were less so. However, the validity of this purely visual assessment is either confirmed or disconfirmed in the next section in which we test the assumption of homoscedasticity.

Testing for the Tenability of Statistical Assumptions

At the beginning of this article, we identified three critical assumptions underlying LGC modeling: (a) that the individual growth trajectories are linearly defined, (b) that measurement error variances are independent, and (c) that measurement error variances are homoscedastic. Our task now is to demonstrate how the tenability of these assumptions can be tested.

Assumption of linearity. In testing Model 2, we hypothesized that individual change in self-perceived Math, Language, and Science abilities over a 3-year period (Grade 8 to Grade 10) are linear. However, it may well be that this change over time is nonlinear rather than linear. The tenability of this assumed linearity can be tested by comparing the overall fit of Model 2 with a modified version of Model 2 (Model 2a) for which curvilinearity is hypothesized. Given that Model 2 is nested within Model 2a, we can compare the difference between the two models by assessing the difference in χ^2 values, which is distributed as χ^2 with degrees of freedom equal to the difference in degrees of freedom. Statistical significance related to this χ^2 difference value ($\Delta\chi^2$) can then be determined. More specifically, we test the null hypothesis that addition of a quadratic term to the linear growth model (Model 2a) does not improve the fit of the simpler representation (Model 2), with evaluation being based on a standard "decrement-to-chi-square" test by which the respective χ^2 values are compared (Willett & Sayer, 1994).

Typically, the most common approach to nonlinear trajectories is the use of polynomials (Duncan et al., 1999). As such, the specification of curvilinearity in LGC models is addressed by means of a third factor, termed a *quadratic factor* that allows for the inclusion of quadratic (or cubic) effects in the hypothesized model. The factor loadings of this third factor can then be constrained to represent a quadratic function of the observed time metric. In other words, the fixed values assigned to the factor loadings of the quadratic factor would represent the square of each point in time. Specific to our three-wave example data, Time 1 would have a fixed value of zero (0^2), Time 2 a fixed value of 1 (1^2), and Time 3 a fixed value of 4 (2^2).

The testing of Model 2a yielded a $\chi^2(43)$ value of 566.207 (CFI = .741; RMSEA = .086). Comparison of this nonlinear model with Model 2 (the linear model) yielded a $\Delta\chi^2(20)$ value of 431.952, which of course is statistically significant ($p <$

.001). These results indicate that the addition of a quadratic parameter leads to an extremely sharp erosion of overall fit in modeling change across time. Based on these findings, we can therefore conclude that, for our example data, change over time is most appropriately represented by a linear growth model.

Assumption of measurement error variance independence. In accordance with classical test theory, researchers typically assume that measurement error variances are independent. However, when analyses involve repeated tests across closely spaced time waves, such assumptions may be untenable and indeed quite unrealistic. Memory effects, for example, can often contribute to the autocorrelation of error variances associated with assessments derived from the same instrument at different points in time.

Testing for the independence of measurement error variances is easily accomplished using the same model-comparison technique demonstrated for the previous assumption. As such, we first specify a model (Model 2b) in which temporally adjacent pairs of measurement error variances are allowed to covary and then compare this model with one in which they are specified as independent parameters. As before, the difference in χ^2 values and related degrees of freedom between the two models determines the tenability of the independence assumption.

The testing of Model 2b yielded goodness-of-fit statistics that represented a somewhat poorer fit to the data, $\chi^2(78) = 188.695$ (CFI = .953; RMSEA = .030) than was the case for Model 2. Furthermore, comparison of fit between the two models revealed a difference in χ^2 values that was statistically significant, $\Delta\chi^2(15) = 54.440$, $p < .001$. Provided with this finding, we conclude the assumption of independent measurement error variances to be tenable. That is to say, Model 2 in which measurement error variances are specified as independent of one another would appear to describe, most appropriately, the example data.

Assumption of measurement error variance homoscedasticity. Based again on classical test theory, researchers typically assume that error variability remains constant for repeated measurements taken at different periods of time; in other words, it is assumed to be homoscedastic. However, as Willett and Sayer (1994) noted, "there is no reason to believe a priori that the precision with which an attribute can be measured is identical at all ages and so measurement errors may be heteroscedastic" (p. 373).

As was the case for the two previous assumptions, homoscedasticity too can be tested using a comparison of covariance structure models. In testing for the validity of Model 2, the linear growth model under study, the error variances were specified as freely estimated parameters. As such, Model 2 allowed for heteroscedasticity. However, in specifying a model in which the error variances are deemed to be homoscedastic, one would constrain these error parameters to be equivalent across the three time periods for each outcome variable. For example, to specify homoscedasticity for Perceived Math Ability, we would constrain its related error variance value at both Time 2 and Time 3 to be equivalent to its estimated value at Time 1.

Testing of a model in which the error variances were assumed to be homoscedastic (Model 2c), yielded a decided decrement in overall fit, $\chi^2(96)$ = 256.504 (CFI = .933; RMSEA = .032). Comparison with Model 2 yielded a $\chi^2(33)$ difference value of 122.249, which again is statistically significant ($p < .001$). Examination of the LMTest of equality constraints revealed three such equivalencies to be untenable: (a) Perceived Math Ability across Times 1 and 2, (b) Perceived Math Ability across Times 1 and 3, and (c) Perceived Science Ability across Times 1 and 3.

In summary, findings from tests of the three key assumptions associated with LGC modeling confirmed that for our hypothesized model (a) the growth trajectory is linear (rather than curvilinear); (b) the measurement errors are independent (rather than autocorrelated); and (c) the measurement error variances are only partially homoscedastic, with the variability of error associated with the assessment of Perceived Math Ability fluctuating significantly between Grades 8 and 9 as well as between Grades 8 and 10, and those associated with the measurement of Perceived Science Ability fluctuating between Grades 8 and 10. More specifically, whereas the variability of error variances associated with the measurement of Perceived Language Ability remained stable across Grades 8, 9, and 10, the variability of those associated with the measurement of Perceived Math Ability increased significantly between Grades 8 and 9 (.270 and .660, respectively) and between Grades 8 and 10 (.270 and 1.005, respectively). Likewise, although error variance was relatively constant across measurements from Grades 8 through 9 and again from Grades 9 through 10, for Perceived Science Ability, it demonstrated a significant decrease between Grade 8 (.870) and Grade 10 (.589). These oscillations in error variance, of course, ultimately bear importantly on the reliability of the measuring instruments.

APPLICATION 3: INTRODUCING GENDER AS A TIME-INVARIANT PREDICTOR OF CHANGE

As noted earlier, provided with evidence of interindividual differences, we can then ask whether, and to what extent, one or more predictors might explain this heterogeneity. For our purposes here, we ask whether statistically significant heterogeneity in the individual growth trajectories (i.e., intercept and slope) of Perceived Ability in Math, Language, and/or Science can be explained by gender as a time-invariant predictor of change. As such, two questions that we might ask are "Do self-perceptions of ability in Math, Language, and/or Science differ for adoles-

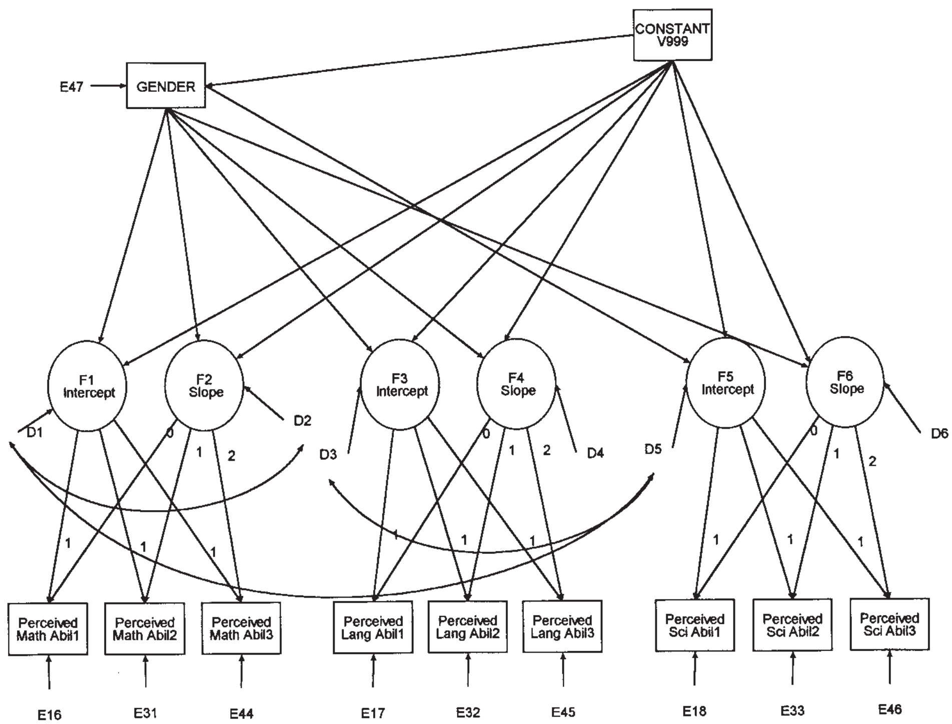

FIGURE 3 Hypothesized multiple-domain predictor latent growth curve model of Perceived Math Ability, Perceived Language Ability, and Perceived Science Ability (Model 3).

cent boys and girls at Time 1 (Grade 8)?" and "Does the rate at which self-perceived ability in Math, Language, and/or Science change over time differ for adolescent boys and girls?" To answer these questions, the predictor variable gender must be incorporated into the Level 2 (or structural) part of the model. This predictor model (Model 3) represents an extension of our final best fitting Model 2 and is shown schematically in Figure 3.

Of import regarding the path diagram displayed in Figure 3 are the newly added components. The first of these is the regression path leading from the Constant to the predictor variable of Gender. Essentially, this path represents the mean on the variable of Gender. However, a more intuitive understanding of this path can be gained by reviewing the standard regression equation[8]

$$y = a + b1x1 + b2x2 + \ldots + e, \qquad (3)$$

[8]We are indebted to Peter M. Bentler for providing us with this insightful explanation.

where the intercept a is the coefficient for the regression of a variable on a constant. As such, we can rewrite Equation 3 equivalently as

$$y = a1 + b1x1 + b2x2 + \ldots + e, \tag{4}$$

where 1 is the constant "variable."

The regression of a variable on a constant is an intercept when the variable is an independent variable; otherwise, it is the mean of the variable. Relatedly, the path flowing from the Constant to the variable of Gender in Figure 3 essentially represents a simpler version of the standard regression equation that can be stated as

$$y = a1 + e, \tag{5}$$

where y = gender; a = the path leading from the Constant to Gender with the Constant being equal to 1; and E47 = e.

The second set of new components comprise the regression paths that flow from Gender to the Intercept and Slope factors associated with each Perceived Ability domain. These regression paths are of primary interest in Model 3 as they hold the key in answering the question of whether the trajectory of Perceived Ability in Math, Language, and/or Science differs for adolescent boys and girls. Also worthy of note is the fact that with the addition of a predictor variable to the model, interpretation of the residuals necessarily changes; these residuals now represent variation remaining in the Intercepts and Slopes after all variability in their prediction by gender has been explained (Willett & Keiley, 2000). Rephrased within a comparative framework, we note that for Model 2 in which no predictors were specified, the residuals represent deviations between the factor Intercepts and Slopes and their population means. In contrast, Model 3 in which a predictor variable is specified, the residuals represent deviations from their conditional population means. As such, these residuals represent the adjusted values of factor Intercepts and Slopes after partialling out the linear effect of the predictor of change (Willett & Keiley, 2000).

Results

Findings related to the analysis of Model 3 are summarized in Table 4. As reported in the table footnote, goodness-of-fit statistics related to this predictor model indicated a relatively good fit to the data. Consistent with Table 3, the first two rows of the table present the average factor Intercept and Slope estimates for each outcome variable, albeit in the present case, controlling for gender. As you will note, these values are similar to those reported in Table 3.

Lines 3 and 4 yield information that is essentially of most interest in the analysis of Model 3. As indicated in Table 4, these estimates represent the regression of in-

TABLE 4
Parameter Estimates for Multiple-Domain Predictor Model (Model 3)

	Perceived Math Ability			*Perceived Language Ability*			*Perceived Science Ability*		
Variable	*Parameter*	*Estimate*	*z Value*[a]	*Parameter*	*Estimate*	*z Value*	*Parameter*	*Estimate*	*z Value*
Factor means									
Intercept	F1, V999	5.007	146.976	F3, V999	4.890	170.179	F5, V999	4.703	146.995
Slope	F2, V999	–0.107	–3.477	F4, V999	0.049	1.690	F6, V999	0.112	3.614
Factor regression of individual	F1, Gender	0.397	4.495	F3, Gender	–0.186	–2.562	F5, Gender	0.176	2.177
Change on gender	F2, Gender	–0.131	–2.354	F4, Gender	–0.109	–2.204	F6, Gender	0.018	0.334
Factor residuals									
Intercept	D1	1.363	14.048	D3	0.520	14.255	D5	0.626	14.287
Slope	D2	0.238	4.259	D4	0.072	2.753	D6	0.093	3.217
Factor residual covariances									
Intercept/slope (within domain)	D1, D2	–0.272	–4.011						
Intercept/slope (between domains)	D1, D5	0.527	15.989	D3, D5	0.252	10.659			
Error variances									
Year 1 (Grade 8)	E16	0.270	3.175	E17	0.680	17.617	E18	0.870	19.139
Year 2 (Grade 9)	E31	0.660	11.230	E32	0.648	11.907	E33	0.788	12.525
Year 3 (Grade 10)	E44	1.005	7.764	E45	0.588	6.483	E46	0.589	5.852

Note. $\chi^2(77) = 254.508$; CFI = .919; RMSEA = .037.

[a] z values > 1.96 indicate statistical significance ($p < .05$).

dividual change on Gender. Turning first to results for Perceived Math Ability, we see that Gender was found to be a statistically significant predictor of both initial status (.397) and rate of change (–.131). Given a coding of "0" for girls and "1" for boys, these findings suggest that whereas self-perceived ability in Math was, on average, higher for boys than for girls at Time 1 by .397, the rate of change in this perception for boys from Grade 8 through Grade 10 was slower than it was for girls by a value of .131. (The negative coefficient indicates that the boys had the lower slope.)

Results related to Perceived Language Ability again revealed Gender was a significant predictor of both the Intercept and the Slope. In this case, however, boys exhibited not only significantly lower perceptions of Language Ability than girls initially (–0.186) but also a slower rate of change in this perception (–0.109) over the 3-year period leading up to Grade 10.

Finally, Gender once again was a significant predictor of Perceived Science Ability in Grade 8, with boys showing higher scores on average by a value of .176 than girls. On the other hand, rate of change was found to be indistinguishable between boys and girls as indicated by their nonsignificant estimates.

DISCUSSION

Thanks to the pioneering work of McArdle and Epstein (1987), Meredith and Tisak (1990), and Muthén (1991), researchers interested in the investigation of change can now address the issue of individual growth modeling using the analysis of covariance structures. Within this general framework, such models are typically termed *latent growth curve models.* The key feature of this approach to investigating change is that it fuses within-person and between-person models of individual growth within the same structural framework, an impossible feat using traditional longitudinal methods. Importantly, these Level 1 and Level 2 components of the LGC model correspond directly with the measurement and structural components of a full structural equation model in the analysis of mean and covariance structures. Our intent in this article was to introduce the reader to LGC modeling by (a) presenting an overview of its fundamental concepts and statistical framework, (b) demonstrating the specification and testing of hypothesized models using example data, and (c) illustrating the means by which the tenability of key assumptions underlying LGC models can be tested. In essence, however, the growth curve models considered in this article are also known as *linear mixed models* or *random coefficient models.* Software for mixed models can handle both missing data (MAR) and the case in which different participants have data at different sets of time points.

Drawing from the work of Willett and Sayer (1994, 1996), we now highlight several important features captured by the LGC modeling approach to the investigation of change. First, the methodology can accommodate anywhere from 3 to 30 waves of longitudinal data equally well. Willett (1988, 1989) has shown, however, that the more waves of data collected, the more precise will be the estimated growth trajectory and the higher will be the reliability for the measurement of change. Second, there is no requirement that the time lag between each wave of assessments be equivalent. Indeed, LGC modeling can easily accommodate irregularly spaced measurements but with the caveat that all participants are measured on the same set of occasions. Third, individual change can be represented by either a linear or nonlinear growth trajectory. Although linear growth is typically assumed by default, this assumption is easily tested and the model respecified to address curvilinearity if need be. Fourth, in contrast to traditional methods used in measuring change, LGC models allow not only for the estimation of measurement error variances but also for their autocorrelation and fluctuation across time in the event that tests for the assumptions of independence and homoscedasticity are found to be untenable. Fifth, multiple predictors of change can be included in the LGC model. They may be fixed, as in the specification of gender in this article, or they may be time varying (see, e.g., Willett & Keiley, 2000). Finally, the three key statistical assumptions associated with LGC modeling (linearity, independence of measurement error variances, homoscedasticity of measurement error variances) can be easily tested via a comparison of nested models, as demonstrated in this article.

We have presented an illustrative introduction to LGC modeling and have emphasized its many important features and advantages over the more traditional approaches to the measurement of change. In doing so, we hope that our paradigmatic applications will not only encourage other researchers to venture forth in their use of this relatively sophisticated methodology but will also provide a springboard that makes their initial foray less onerous.

REFERENCES

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238–246.

Bentler, P. M. (2000). *EQS 6 structural equations program manual.* Encino, CA: Multivariate Software.

Bentler, P. M., & Weeks, D. G. (1980). Linear structural equations with latent variables. *Psychometrika, 45,* 289–308.

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

Boomsma, A. (1985). Nonconvergence, improper solutions, and starting values in LISREL maximum likelihood estimation. *Psychometrika, 52,* 345–370.

Boomsma, A., & Hoogland, J. J. (2001). The robustness of LISREL modeling revisited. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural equation modeling: A festschrift in honor of Karl Jöreskög* (pp. 139–168). Lincolnwood, IL: Scientific Software.

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). Newbury Park, CA: Sage.

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

Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS/windows: Basic concepts, applications, and programming.* Thousand Oaks, CA: Sage.

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Chou, C.-P., Bentler, P. M., & Pentz, M. A. (1998). Comparisons of two approaches to study growth curves: The multilevel model and the latent curve analysis. *Structural Equation Modeling, 5,* 247–266.

Duncan, T. E., & Duncan, S. C. (1994). Modeling incomplete longitudinal substance use data using latent variable growth curve methodology. *Multivariate Behavioral Research, 29,* 313–338.

Duncan, T. E., & Duncan, S. C. (1995). Modeling the processes of development via latent growth curve methodology. *Structural Equation Modeling, 3,* 187–205.

Duncan, T. E., Duncan, S. C., Stryker, L. A., Li, F., & Alpert, A. (1999). *An introduction to latent variable growth curve modeling.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Eccles, J. S. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75–146). San Francisco: Freeman.

Eccles, J. S., Adler, T. F., & Meece, J. L. (1984). Sex differences in achievement: A test of alternate theories. *Journal of Personality and Social Psychology, 46,* 26–43.

Fuligni, A. J., Eccles, J. S., & Barber, B. L. (1995). The long-term effects of seventh-grade ability grouping in mathematics. *Journal of Early Adolescence, 15,* 58–89.

Hancock, G. R., Kuo, W.-L., & Lawrence, F. R. (2001). An illustration of second-order latent growth models. *Structural Equation Modeling, 8,* 470–489.

Jacobs, J. E., & Eccles, J. S. (1992). The impact of mothers, gender-role stereotypic beliefs on mothers, and children's ability perceptions. *Journal of Personality and Social Psychology, 63,* 932–944.

Jöreskog, K., & Sörbom, D. (1993), *LISREL 8: Structural equation modeling with the SIMPLIS command language.* Chicago: Scientific Software.

Li, F., Duncan, T. E., & Duncan, S. C. (2001). Latent growth modeling of longitudinal data: A finite growth mixture modeling approach. *Structural Equation Modeling, 8,* 493–530.

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

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

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

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* (pp. 1–17). Washington, DC: American Psychological Association.

Muthén, B., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika, 52,* 431–462.

Muthén, B. O., & Muthén, L. K. (1998–2001). *Mplus user's guide.* Los Angeles: Author.

Parsons, J. E., Adler, T. F., & Kaczala, C. M. (1982). Socialization of achievement attitudes and beliefs: Parental influences. *Child Development, 53,* 310–321.

Rogosa, D. R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to to the measurement of change. *Psychological Bulletin, 90,* 726–748.

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

Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63,* 581–592.

Satorra, A., & Bentler, P. M. (1988). Scaling corrections for chi-square statistics in covariance structure analysis. In *American Statistical Association 1988 proceedings of the business and economics section* (pp. 308–313). Alexandria VA: American Statistical Association.

Willett, J. B. (1988). Questions and answers in the measurement of change. In E. Z. Rothkopf (Ed.), *Review of research in education* (Vol. 15, pp. 345–422). Washington, DC: American Educational Research Association.

Willett, J. B. (1989). Some results on reliability for the longitudinal measurement of change: Implications for the design of studies of individual growth. *Educational and Psychological Measurement, 49,* 587–602.

Willett, J. B., & Keiley, M. K. (2000). Using covariance structure analysis to model change over time. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 665–694). San Diego, CA: Academic.

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

Willett, J. B., & Sayer, A. G. (1996). Cross-domain analyses of change over time: Combining growth modeling and covariance structure analysis. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 125–157). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.