Research in Human Development
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t775653695

A Second-Order Growth Mixture Model for Developmental Research

Kevin J. Grimm[a]; Nilam Ram[b]
[a] University of California, Davis [b] The Pennsylvania State University,

To cite this Article Grimm, Kevin J. and Ram, Nilam(2009) 'A Second-Order Growth Mixture Model for Developmental Research', Research in Human Development, 6: 2, 121 — 143
To link to this Article: DOI: 10.1080/15427600902911221
URL: http://dx.doi.org/10.1080/15427600902911221

Research in Human Development, *6*(2–3), 121–143, 2009

ISSN: 1542-7609 print/1542-7617 online
DOI: 10.1080/15427600902911221

Psychology Press
Taylor & Francis Group

A Second-Order Growth Mixture Model for Developmental Research

Kevin J. Grimm
University of California, Davis

Nilam Ram
The Pennsylvania State University

Growth mixture modeling, a combination of growth modeling and finite mixture modeling, is a flexible, exploratory method for identifying and describing between-person heterogeneity in change. In this article we introduce a second-order growth mixture model that combines a longitudinal common factor model, measurement invariance constraints, latent growth model, and mixture model. This approach capitalizes on the benefits of multivariate measurement and the flexibility of mixtures for representing heterogeneity. We describe the model and illustrate its use with multi-reporter longitudinal data from the National Institute of Child Health and Human Development (NICHD) Study of Early Child Care and Youth Development tracking the development of children's externalizing behaviors through elementary school.

Developmental researchers are concerned with how individuals change. Whether the changes are quantitative or qualitative, the study is longitudinal or cross-sectional, observational or experimental, developmentalists seek to implicitly or explicitly describe and understand how individuals change over time. Even though within-person change is often approximated by between-person differences in cross-sectional studies, longitudinal (i.e., repeated measures) studies are necessary for describing how people actually change (i.e., identification of intraindividual change; Baltes & Nesselroade, 1979). An additional important aspect of studying

Address correspondence to Kevin J. Grimm, Psychology Department, University of California, Davis, One Shields Avenue, Davis, CA 95616. E-mail: kjgrimm@ucdavis.edu

change is describing and understanding heterogeneity in change and determining what constructs are related to that heterogeneity (i.e., determinants of between-person differences in change; Baltes & Nesselroade, 1979). The primary analytic techniques for examining how people change, how people differ in change, and the determinants of change has become the latent growth curve model (McArdle & Epstein, 1987; Meredith & Tisak, 1990; Preacher, Wichman, MacCallum, & Briggs, 2008; Rogosa & Willet, 1985).

The latent growth curve has roots in the 1950s (Rao, 1958; Tucker, 1958) and became a mainstay in longitudinal research when Meredith and Tisak (1990) showed how the latent growth curve could be fit as a restricted common factor model using available structural equation modeling software. The latent growth curve was an improvement over repeated measures ANOVA and difference score analysis largely because change variance could be distinguished from error variance—an issue that had been raised with difference scores (Cronbach & Furby, 1970; Nesselroade & Cable, 1974). In addition to structural equation modeling (SEM) programs, the latent growth curve model can be fit using multilevel (e.g., mixed-effects, random coefficient) modeling programs under the logic that the repeated observations are nested within participants (see Bryk & Raudenbush, 1992). The multilevel and SEM approaches to growth modeling have advantages and disadvantages (Ghisletta & Lindenberger, 2004), and identical models can be constructed using both approaches (Ferrer, Hamagami, & McArdle, 2004). Since its introduction, the latent growth curve model has been extended in a variety of ways to handle nonlinear patterns of development (Browne, 1993; Grimm & Ram, in press; Ram & Grimm, 2007), multiple groups (McArdle & Hamagami, 1996), changes in multiple variables (McArdle, 1988), lead–lag associations in multivariate change (McArdle, 2001), and mixture distributions (B. O. Muthén & Shedden, 1999). As we describe, the growth model has also been extended to include a lower-order factor model (Hancock, Kuo, & Lawrence, 2001; McArdle, 1988) or item response model (Curran, Edwards, Wirth, Hussong, Chassin, 2007; McArdle, Grimm, Hamagami, Bowles, & Meredith, in press).

Although very useful, the basic latent growth curve has some limitations for describing between-person differences in change. In particular, two key assumptions of the model may be restrictive. First, between-person differences change are assumed to be normally distributed. Second, it is assumed that all individuals follow the same pattern of change (e.g., all individuals exhibit linear change). The growth mixture model (GMM: Muthén, 2004; B. O. Muthén & Shedden, 1999; B. O. Muthén & Muthén, 2000), a finite mixture extension of the latent growth curve, can allow for greater flexibility in if and how these assumptions about the organization of between-person differences are invoked.

THE GROWTH MIXTURE MODEL

In brief, growth mixture modeling is a method for identifying multiple unobserved subsamples, describing longitudinal change within each unobserved subsample, and examining differences in change between and within unobserved subsamples (see Connell & Frye, 2006; B. O. Muthén & Muthén, 2000; Ram & Grimm, 2009, for an introduction). By allowing for the possibility that individuals belong to one of multiple unobserved groups (i.e., classes), the GMM does not restrict the between-person differences in change to follow a continuous normal distribution. Additionally, because these unobserved groups can have fundamentally different change patterns (e.g., one class exhibiting linear change, and another class exhibiting exponential change) the GMM is an important innovation that can be used to locate, describe, and understand additional sources of heterogeneity. For example, Odgers et al. (2008) examined trajectories of antisocial behaviors for males and females and found a four-class GMM to be the best representation of the data. The four classes had different levels of antisocial behavior at age 7 and different trajectories. The classes were indicative of (1) children with persistently low levels of antisocial behavior (persistent low), (2) children with high levels of antisocial behavior and a decreasing trajectory (childhood limited), (3) children with low levels of antisocial behavior and an increasing trajectory (adolescent onset), and (4) children with high levels of antisocial behavior and a flat trajectory (early onset persistent). Similarly, Small and Bäckman (2007) examined changes in the Mini-Mental Status Examination (MMSE) in a sample of older adults using GMMs. Small and Bäckman settled on a two-class model with quadratic trends. The two classes showed small initial differences, but one class (n = 112) showed sharp declines in the MMSE during the subsequent 7 years, whereas the second class (n = 416) showed significant, but comparatively small declines in the MMSE during the follow-up period.

Of course, the additional flexibility the GMM provides for modeling change must be carefully used. As pointed out by several authors, without good theory and thoughtful application, the benefits of the GMM might be overshadowed by inappropriate use (e.g., Bauer, 2007; Bauer & Curran, 2003, Hipp & Bauer, 2006). The questions raised about the GMM stem from five issues: (1) the possibility of incorrect solutions due to convergence at local, as opposed to global, maxima in the likelihood function; (2) lack of clearly defined rules for assessing model fit and making model comparisons (i.e., deciding on the appropriate number of latent classes); (3) the exploratory nature of the model; (4) how and when covariates are included in the model; and (5) the possibility that non-normal distributions in the outcome measure are incorrectly interpreted as arising from multiple populations.

Our primary interest is to propose the second-order growth mixture model (SOGMM) to take advantage of multivariate assessment (e.g., multiple reporters) when examining heterogeneity of developmental trajectories. This model may help to address the fifth issue. Before proceeding, though, we briefly mention some of the strategies that have been developed to help deal with the other four issues.

The estimation of GMMs is sensitive to starting values, and the estimation algorithm may converge on a local rather than the global (i.e., best) solution. To combat the possibility of not ending up at the optimal solution, is it recommended that the algorithm for estimating parameters and maximizing the likelihood function be started multiple times, beginning its search from different places in the parameter space. Replication of the solution with multiple sets of random starting values provides confidence in the obtained solution (see Hipp & Bauer, 2006; McLachlan & Peel, 2000; Muthén, 2004).

The second concern deals with model section. GMMs are not nested under growth models or other GMMs. With the usual indices for making nested model comparisons not available, researchers must use non-nested comparative fit indices (e.g., Akaike Information Criteria [AIC]; Bayesian Information Criteria [BIC]) for model selection, which may be highly sensitive to sample size. Additional indices that have been proposed (and evaluated with promising results) include the Lo-Mendell-Rubin Likelihood Ratio Test and Bootstrap Likelihood Ratio Test (see Muthén, 2004; Nylund, Asparouhov, & Muthén, 2007).

Growth mixture modeling is an exploratory, data-driven technique, and thus highly subject to the chance relationships existing in a particular set of data. Thus, replication of solutions across multiple studies is important for establishing confidence in and generalizability of results. In the absence of data drawn from multiple sources, users should consider attempting to replicate the findings using independent or random subsets of the available data (e.g., Bootstrap) as well as check results against theoretically based expectations regarding the number and characteristics of latent classes.

When and how covariates are included in the model is a tricky issue. On the one hand, covariates are important indicators of the substantive validity of the obtained classes. Demonstrating that the classifications map onto covariates in the expected manner or are predictive of important outcomes confirms their utility. On the other hand, when estimated simultaneously, these additional predictor or outcome variables can fundamentally affect the latent classes such that they (and which individuals are within them) are substantially different from the latent classes obtained without covariates. Additionally, covariates can be added as predictors of the latent classes and/or included within the class-specific models as predictors of the intercept and slope with any of these effects either differing or being constrained to be invariant among latent classes—all of which leads to different interpretations and implications (see Bauer & Curran, 2003; Lubke & Muthén, 2007; B. O. Muthén & Muthén, 2000).

Bauer and Curran (2003) highlighted concerns about the interpretation of results obtained via growth mixture modeling. Specifically, they cogently demonstrated the effects non-normality has on model selection when comparing GMMs with growth models. Bauer and Curran simulated normally distributed longitudinal data from a single population, transformed the data to be slightly skewed and kurtotic, and fit a series of growth and GMMs. The fit statistics universally favored the GMM over the growth model demonstrating how small deviations from normality can lead to the conclusion that multiple latent classes underlie the data even though the data generating mechanism was a single population. Non-normality (skew, kurtosis) may arise for many different reasons—not just from the mixing of multiple populations. Consider the possibility that the construct of interest is normally distributed in the population, but ceiling, floor, or other measurement anomalies have led to observed data that is non-normal. As Bauer and Curran demonstrated, the GMM will accommodate the non-normality by finding unobserved groups. The concern is that the GMM, by itself, does not distinguish among the reasons for non-normality. It simply fits the data as best it can—adding as many groups as necessary to represent the skew and kurtosis of the data. Non-normality resulting from measurement or other anomalies masquerades as latent classes with no red flags ever being raised.

As a consequence, it is imperative to take advantage of measurement models and multivariate assessment to provide a cleaner basis for the GMM. One approach is to make use of the benefits provided by multivariate assessments. As detailed by Edwards and Wirth (this issue), multiple measures and factor analytic models can be used to separate common variance from specific and error variance. The true score (e.g., factor score) distributions obtained from such models may provide a more precise and clear picture of how the construct is distributed in the sample (and population) and provide a stronger foundation for the GMM. For example, reports of externalizing behavior obtained from multiple sources (e.g., mother, father, and teacher) can be used in conjunction with factor or item response models to obtain a more reliable picture of a child's level of externalizing behavior. Similarly, models that specifically acknowledge and accommodate censored (ceilings or floors), zero-inflated, Poisson, Bernoulli, or other distributions can be used to obtain more precise estimates of the true distribution or frequencies of individuals' behavior. As has been done for the growth model (e.g., Hancock et al., 2001; McArdle, 1988), we propose taking advantage of the benefits of multivariate and other measurement models as one way to help deal with data anomalies that might inappropriately masquerade as multiple classes.

A SECOND-ORDER GROWTH MIXTURE MODEL

A SOGMM takes the benefits of multivariate measurement models and combines them with the GMM. Our purposes here are to introduce the model, as built from the longitudinal common factor model, measurement invariance constraints, second-order growth model, and the GMM, and to demonstrate its application to developmental data. First, we provide background on the context of our inquiry—between-child differences in change in externalizing behavior during childhood. Next, we present the model, its components, and illustrate how the model can be fit and evaluated using data from the NICHD Study of Early Child Care and Youth Development. Working knowledge of factor and growth models is assumed in the presentation of details, but not necessary for obtaining a broad overview of the method. Finally, we highlight some possible extensions of the SOGMM and reiterate that care and caution should be taken when using and interpreting GMMs.

RESEARCH ON INTRAINDIVIDUAL CHANGE IN EXTERNALIZING BEHAVIORS

Several studies have used growth curve modeling techniques to examine individual change in externalizing problems through early and middle childhood, adolescence, and emerging adulthood (Bongers, Koot, van der Ende & Verhulst, 2004; Bub, McCartney & Willett, 2007; Curran et al., 2007; Dekovic, Buist, & Reitz, 2004; Gilliom & Shaw, 2004; Keiley, Bates, Dodge & Pettit, 2000; Leve, Kim, & Pears, 2005; Miner & Clarke-Stewart, 2008). Prototypically, whether reported by parents, teachers, or other caregivers, studies have found significant decreases in externalizing behaviors over time and significant between-person variation in intercepts and linear slopes (rates of change). As applications of growth models, these studies investigated individual changes and between-person differences in change under the assumption that the samples were drawn from a single population.

Other studies have used group-based approaches (e.g., latent class, latent profile, & mixture models) that explicitly allow for multiple populations (e.g., Hill, Degnan, Calkins, & Keane, 2006; Moffitt, 1993; Moffitt, Caspi, Dickson, Silva, & Stanton, 1996; NICHD Early Child Care Research Network [ECCRN], 2004; Shaw, Gilliom, Ingoldsby, & Nagin, 2003). We selectively highlight a few of the findings as they might be used to form some preliminary views on how many and what types of change patterns might be expected in the forthcoming data example.

Hill et al. (2006) fit latent profile models to longitudinal data collected during early childhood (i.e., ages 2 to 5). Heterogeneity of across-time profiles was described using four patterns: high-chronic (11% for girls, 9% of boys), high-decreasing (22% for girls, 39% for boys), moderate-decreasing (51% for girls, 41% for boys), and persistent-low (16% for girls, 11% for boys). Shaw et al. (2003) fit a semiparametric mixture model (e.g., Nagin, 1999) to longitudinal data on conduct problems obtained from children from ages 2 to 8. Heterogeneity was again described by four patterns of change: high-chronic (6 % of sample), high-decreasing (38%), moderate-decreasing (42%), and persistent low (14%). Using a similar approach, the NICHD ECCRN (2004) described heterogeneity of change in mother-rated aggressive behavior across repeated assessments from ages 2 to 9 using a five-class typology: two groups with low levels of aggression at age 2 and subsequent decreases (70% of sample), two groups with moderate levels of aggression at age 2 and subsequent decreases (27%,), and a small group that exhibited a high level of aggression at age 2 that stayed high to age 9 (3%). In all three studies there was a high-chronic class that exhibited stability and one or more classes whose problem behavior, although starting off with low, moderate, or high levels of problem behavior, decreased systematically with age. These empirical results suggest the possibility of multiple populations of children. Parsimoniously, and highlighting change more than level, the classes might represent a normative population whose problem behaviors decrease during childhood (from whatever level when first observed), and a relatively small clinical population who persistently exhibit high levels of problem behavior.

METHOD

Example Data

To illustrate the application and interpretation of a SOGMM we make use of multivariate-multioccasion data drawn from the National Institute of Child Health and Human Development Study of Early Child Care and Youth Development (NICHD-SECCYD). Briefly, SECCYD families were recruited through hospital visits to mothers shortly after the birth of a child in 1991 in 10 locations in the United States. During selected 24-hour intervals, all women giving birth (N = 8,986) were screened for eligibility. From that group, 1,364 families completed a home interview when the infant was 1 month old and became study participants (see e.g., NICHD ECCRN, 2002, for details). The analysis sample for the current study consisted of the 1,135 children for whom mother, father, and/or teacher reports of externalizing behavior were available for at least one of the grade 1, 3, 4, or 5 assessments.

Measures. Reports of externalizing behavior problems were obtained from mothers, fathers, and teachers during Grades 1, 3, 4, and 5 (coded 0, 2, 3, 4, respectively). At each occasion the child's mother and father completed the Child Behavior Checklist (CBCL/4-18; Achenbach, 1991a), a 118-item scale on which parents rate aspects of their child's behavior during the last 6 months on a 3-point scale (i.e., *not true, somewhat or sometimes true, very true or often true*). In parallel, the child's teacher completed the Achenbach's Teacher Report Form (TRF: Achenbach, 1991b), a 118-item scale on which the teacher rated aspects of the child's behavior. Ninety-three items on the TRF have direct counterparts on the CBCL/4-18 (the remaining items deal with situations specific to the school environment). Here we use the raw externalizing (i.e., sum) scores of the CBCL/4-18 (33 items) and the TRF (34 items), which have been shown to have good test-retest reliability and internal consistency (Achenbach, 1991a, 1991b). In sum, measures of the child's externalizing behavior were obtained at *four occasions* (i.e., Grades 1, 3, 4, 5) from three reporters (i.e., mother, father, & teacher).

Descriptive statistics. Summary statistics for the externalizing behavior scores are contained in Table 1. To accommodate missing data, the estimated means, standard deviations, and correlations were generated using full information maximum likelihood and are considered representative under missing at random conditions (Little & Rubin, 1987). Examining Table 1, it can be seen that the reports of externalizing behavior were strongly correlated within and across grades. The across-grade correlations were especially strong for reports obtained from the same informant (e.g., mother to mother, father to father, & teacher to

TABLE 1
Descriptive Statistics for the Externalizing Score from the Child Behavior Checklist in First, Third, Fourth and Fifth Grades

		First Grade			*Third Grade*			*Fourth Grade*			*Fifth Grade*		
	Correlations	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
	N	1009	668	1008	1007	637	982	992	611	914	993	631	927
First	1. Mother	1.00											
grade	2. Father	.51	1.00										
	3. Teacher	.34	.37	1.00									
Third	4. Mother	.74	.46	.34	1.00								
grade	5. Father	.52	.70	.35	.60	1.00							
	6. Teacher	.38	.41	.53	.40	.44	1.00						
Fourth	7. Mother	.72	.46	.30	.80	.59	.37	1.00					
grade	8. Father	.51	.71	.36	.54	.79	.38	.62	1.00				
	9. Teacher	.40	.33	.54	.39	.39	.64	.38	.42	1.00			
Fifth	10. Mother	.69	.44	.33	.73	.51	.34	.80	.55	.38	1.00		
grade	11. Father	.48	.66	.30	.53	.72	.40	.57	.77	.36	.60	1.00	
	12. Teacher	.30	.32	.49	.29	.31	.53	.31	.35	.60	.32	.34	1.00
	Mean	8.31	8.70	5.84	7.44	7.35	6.48	6.93	7.07	5.92	6.62	6.28	6.21
	Standard Deviation	6.65	6.49	8.32	6.37	5.82	9.39	6.20	6.56	8.99	6.35	6.48	9.25
	Skew[a]	1.23	1.11	2.07	1.23	1.14	2.22	1.40	1.88	2.38	1.55	1.94	2.23
	Kurtosis[a]	2.19	1.25	4.27	1.56	1.96	5.44	2.68	6.16	6.17	3.39	4.92	5.36
	Mode[a]	4	2	0	2	1	0	0	0	0	0	0	0

Note. Full information maximum likelihood estimates.
[a]Estimates were obtained from the sample with available data at each occasion.

teacher). The means tend to decrease across time, with the exception of the teacher reports, which remain relatively stable. Overall, though, the pattern suggests the possibility of normative decline. Additionally, skew and kurtosis were calculated using the subsample of children for whom data were available at each occasion. All scores were positively skewed and kurtotic—indicating non-normality was present in all measures.

Second-Order Growth Mixture Model

The SOGMM is built by combining (1) a longitudinal common factor model, (2) measurement invariance constraints, (3) a latent growth model, and (4) a mixture model. For clarity, we build the model in steps and attempt to explain the matrix algebra by highlighting the substantive implications of each component.

Longitudinal factor model. The longitudinal common factor model can be written as

$$\mathbf{y}_{it} = \nu_t + \mathbf{\Lambda}_t \eta_{it} + \varepsilon_{it}, \tag{1}$$

where $\mathbf{y}_{it}$ is a $p \times 1$ vector of observed variables for individual i at time t, ν_t is a $p \times 1$ vector of observed variable intercepts at time t, $\mathbf{\Lambda}_t$ is a $p \times q$ matrix of factor loadings at time t, η_{it} is a $q \times 1$ vector of latent factor scores for individual i at time t, and ε_{it} is a $p \times 1$ vector of residual scores for individual i at time t. The purpose of this component is to invoke the benefits of multivariate measurement to obtain a true representation of individuals' scores on the underlying construct. In our example, the multiple reports of children's externalizing behaviors are used to separate unique variance (ε_{it}; specific + measurement error) from children's true level of externalizing behavior (η_{it}) at each occasion. Mother, father, and teacher reports obtained in Grades 1, 3, 4, and 5 serve as the observed scores, $\mathbf{y}_{it}$. The factor scores, η_{it}, represent the children's true level of externalizing behavior at each grade and become our primary interest.

Measurement invariance constraints. The common factor model provides a framework for establishing or obtaining the distribution of scores on the latent construct of interest. In the longitudinal setting it is important to establish that the same construct has been measured at each occasion in the same metric. Tests of factorial invariance (Meredith, 1993; Meredith & Horn, 2001; Widaman & Reise, 1997) enable researchers to determine whether the same construct was measured in the same metric at each occasion. This is done by testing whether the relationships between the observed variables and the latent factor are the same at each occasion and whether changes in the observed measures can be carried by changes in the factor scores. Specifically, weak factorial invariance is established by imposing equality constraints on the factor loadings so that $\mathbf{\Lambda}_t = \mathbf{\Lambda}$ for all t. Strong factorial invariance, a necessary condition to examine change at the factor level (however, see Edwards & Wirth, this issue, for another view on measurement invariance constraints), adds equality constraints on the observed variable intercept so that $\nu_t = \nu$ for all t. The means of the factor scores can be estimated under the strong factorial invariance model. When strong invariance holds, Equation 1 simplifies to

$$\mathbf{y}_{it} = \nu + \mathbf{\Lambda}\eta_{it} + \varepsilon_{it}. \tag{2}$$

This first-order portion of the model insures that the measurement instrument was calibrated properly at each occasion and establishes a viable foundation on which to examine within-person changes and between-person differences in change.

Latent growth model. Having established a common and invariant measurement framework that takes advantage of multivariate measurement, we can begin examining how individuals' externalizing behavior changes across time and the between-child differences in those changes. Within-person change and between-person differences in change in the factor scores, η_{it}, can be examined using a *second-order growth model* (Ferrer, Balluerka, & Widaman, 2008; Hancock et al., 2001; McArdle, 1988). The second-order nature of the model means that a second layer of factors is built on the first layer given by the first-order longitudinal factor model given above (as opposed to growth models built on observed variables). Using the common factor approach (see Meredith & Tisak, 1990), a second-order latent growth curve can be written as

$$\eta_{it} = \mathbf{\Gamma}\xi_i + \zeta_{it}, \tag{3}$$

where $\mathbf{\Gamma}$ is a $q \times r$ matrix of second-order factor loadings, ξ_i is an $r \times 1$ vector of second-order factor scores (e.g., intercept and slope), and ζ_{it} is a $q \times 1$ vector of latent variable disturbance scores. The second-order factor scores (intercepts and slopes) can be written as deviations from the group mean, such as

$$\xi_i = \kappa + \omega_i, \tag{4}$$

where κ is an $r \times 1$ vector of latent factor means and ω_i is a $r \times 1$ vector of individual mean deviations. In our example, changes in the externalizing behavior factor (η_{it}), indicated by the mother, father, and teacher-reported observed externalizing scores at the first-order level, are modeled by second-order growth factors (ξ_i), typically, an intercept factor capturing between-person differences at the beginning (or any specific point) of the observation period, and a slope factor(s) capturing between-person differences in the rate of change across the series of measurements. These second-order growth factors are composed of means (κ) and individual deviations (ω_i) about the mean. Combining Equations 2, 3, and 4 we can write the second-order growth model as

$$\mathbf{y}_{it} = \nu + \mathbf{\Lambda\Gamma}\kappa + \mathbf{\Lambda\Gamma}\omega_i + \mathbf{\Lambda}\zeta_{it} + \varepsilon_{it}. \quad (5)$$

The expectations of the population mean vector, μ, and covariance matrix, $\mathbf{\Sigma}$, based on the second-order growth model are

$$\mu = \nu + \mathbf{\Lambda\Gamma}\kappa$$
$$\mathbf{\Sigma} = \mathbf{\Lambda}(\mathbf{\Gamma\Phi\Gamma}' + \mathbf{\Psi})\mathbf{\Lambda}' + \boldsymbol{\theta} \quad (6)$$

where $\mathbf{\Psi}$ is a $q \times q$ first-order latent variable residual covariance matrix, $\mathbf{\Phi}$ is an $r \times r$ second-order latent variable covariance matrix, and $\boldsymbol{\theta}$ is a $p \times p$ matrix of observed variable residual covariances. In the second-order growth modeling framework, $\mathbf{\Psi}$ is diagonal (first-order factor covariances are modeled by second-order growth model) with equivalent values in the diagonal following the homogeneity of variance assumption of latent growth modeling. An additional technical detail is the mean of the intercept factor is not identified (as the scale of the first-order factors is arbitrary). The needed identification constraint is obtained by fixing the mean of the intercept factor to 0.

Figure 1 is a path diagram of a second-order growth model with four occasions of measurement and three observed variables at each measurement occasion. The labels in the path diagram represent the matrices of Equation 6, and invariance constraints are indicated by common labels. For clarity, the observed variable intercepts, $\boldsymbol{\nu}$ (that are invariant across time), and unique covariances (e.g., $\theta_{1,4}$, $\theta_{1,7}$, $\theta_{1,10}$, $\theta_{4,7}$, $\theta_{4,10}$, $\theta_{7,10}$) are not shown in the path diagram. At the first-order level, the observed variables (squares) are indicators of factors (circles) using the longitudinal factor model and strong measurement invariant constraints. At the second-order level, the two factors represent an intercept (η_1) with unit loadings and a linear slope (η_2) with factor loadings following a linear change pattern with respect to measurement occasion (occasions 0, 2, 3, & 4). The mean of the linear slope (κ_2), variances of the intercept and slope ($\Phi_{1,1}$ & $\Phi_{2,2}$), and a covariance between the intercept and slope ($\Phi_{1,2}$) provide the model based description of the between-person differences in within-person change. Overall, the growth model is extremely flexible and allows for description and examination of a wide variety and types of linear and nonlinear change (see Grimm & Ram, in press; Ram & Grimm, 2007). Although not yet in widespread use, this second-order version has additional advantages, including the benefits of multivariate measurement (Hancock et al., 2001; McArdle, 1988) and increased statistical power (see Hertzog, Lindenberger, Ghisletta, von Oertzen, 2006).

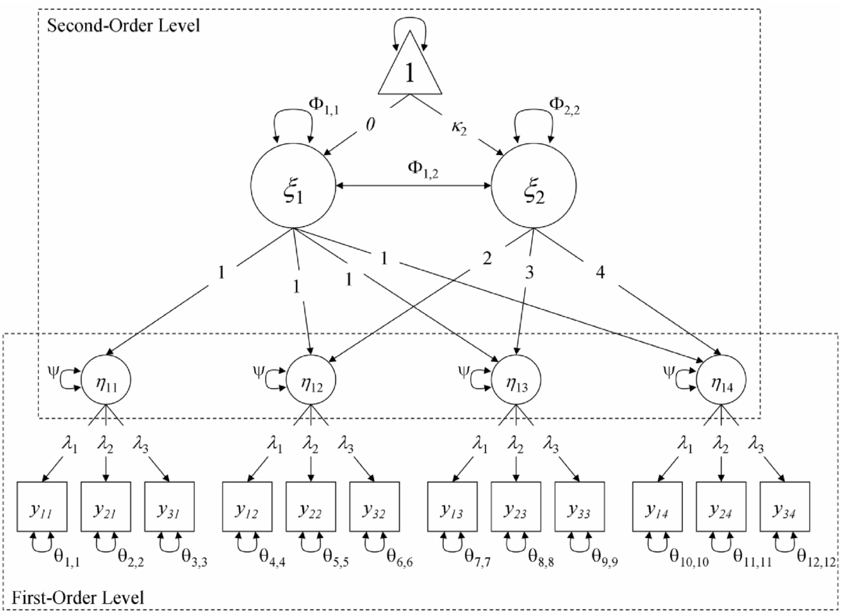

FIGURE 1 Path diagram of a second-order latent growth curve with four occasions of measurement of an unobserved factor indicated by three observed variables at each occasion.

Note. Intercepts for observed variables (one-headed arrows from the constant to the observed variables) and correlated uniquenesses were omitted for clarity. Invariance constraints are noted by a common label. Lower and upper dotted boxes represent the first- and second-order levels, respectively.

Growth mixture model. The GMM (B. O. Muthén & Muthén, 2000; B. O. Muthén & Shedden, 1999) is a combination of the growth model and the finite mixture model (McLachlan & Peel, 2000). The idea is that the observed sample was drawn from K subpopulations. The problem is that it is not known who belongs to each class or subpopulation, each of which follows its own growth model. Interindividual differences within each class are normally distributed, but when combined or mixed together they appear as a single non-normal distribution. The GMM incorporates a categorical latent variable into the growth model that allows for a probabilistic separation of individuals into K classes.

Mathematically, the expectations from the second-order growth model in Equation 6 are written at the class level, such as

$$\mu_k = \nu_k + \mathbf{\Lambda}_k \mathbf{\Gamma}_k \kappa_k$$
$$\mathbf{\Sigma}_k = \mathbf{\Lambda}_k (\mathbf{\Gamma}_k \mathbf{\Phi}_k \mathbf{\Gamma}'_k) \mathbf{\Lambda}'_k + \mathbf{\Lambda}_k \mathbf{\Psi}_k \mathbf{\Lambda}'_k + \boldsymbol{\theta}_k. \quad (7)$$

Each matrix is subscripted by k to denote class specific parameters. The GMM described above is general such that every matrix is subscripted by k to denote class noninvariance. However, several constraints can be imposed for interpretability. At the first-order, $\nu_k = \nu$, $\mathbf{\Lambda}_k = \mathbf{\Lambda}$, $\mathbf{\Psi}_k = \mathbf{\Psi}$, and $\mathbf{\theta}_k = \mathbf{\theta}$ for all k to denote the observed variable intercepts, the factor loadings for the measurement model, the latent variable disturbances, and the residual variances of the observed variables are invariant over latent classes (e.g., measurement invariance constraints). In principle, these constraints are not necessary, in that classes can differ in any aspect of the model but make interpretation of the differences in change more straightforward and estimation simpler. Here we examine mean (κ_k), covariance ($\mathbf{\Phi}_k$), and structure ($\mathbf{\Gamma}_k$) differences across latent classes (i.e., in the parameters most relevant for representing growth).

Model Fitting

Elsewhere we have outlined four steps for conducting growth mixture modeling analyses (1) problem definition, (2) model specification, (3) model estimation, and (4) model selection and interpretation (see Ram & Grimm, 2009). Our main purpose here has been to present the second-order growth model. Thus, we provide only abbreviated descriptions of how these steps were implemented in the current analysis.

Problem definition. Using the empirical findings reviewed here, we formulated some initial GMM hypotheses. Specifically, we expect there would be two or three latent classes. The classes could differ in the mean amount of change, the extent of between-person difference in change, and in the pattern of change. Specifically, a two-class representation might distinguish a relatively heterogeneous group with a normative linear decline pattern, and a relatively homogenous clinical group with high scores that remain stable over time. Alternatively, a three-class model might additionally identify another group with low *and* stable levels of externalizing behavior.

Model specification, estimation, & selection. A series of two- and three-class SOGMMs were specified with different between-class equality constraints (κ_k, $\mathbf{\Phi}_k$, & $\mathbf{\Gamma}_k$). As can be seen in the details described in Table 2, the models differed in the parameters that were invariant over classes. Model $M1_2$ (subscript 2 for two classes) allowed for between-class differences in the means of the growth factors; Model $M2_2$ allowed for between-class differences in the means and in the variance/covariance matrix of the growth factors; Model $M3_2$ allowed for between-class differences in the means and variance/covariance matrix of the growth factors as well as the pattern of change. In particular, the pattern of change for one class was defined to capture stability (i.e., intercept-only model; expected clinical stability), whereas the pattern of change for the other class was

TABLE 2
Second-Order Growth Mixture Models fit to the Externalizing Factors

Model	*Linear Growth Model*	*Growth Mixture Models*					
		Two-Class Models					
	M0₁ → $M0_1$	$M1_2$		$M2_2$		$M3_2$	
		k = 1	*k = 2*	*k = 1*	*k = 2*	*k = 1*	*k = 2*
Latent variable means							
Intercept Mean	= 0	κ_1	= 0	κ_1	= 0	κ_1	= 0
Slope Mean	κ_2	κ_2	κ_2	κ_2	κ_2	= 0	κ_2
Slope loadings							
Grade 1	= 0	= 0		= 0		= 0	= 0
Grade 3	= 2	= 2		= 2		= 0	= 2
Grade 4	= 3	= 3		= 3		= 0	= 3
Grade 5	= 4	= 4		= 4		= 0	= 4
Latent variable covariances							
Intercept variance	$\Phi_{1,1}$	$\Phi_{1,1}$		$\Phi_{1,1}$	$\Phi_{1,1}$	$\Phi_{1,1}$	$\Phi_{1,1}$
Slope variance	$\Phi_{2,2}$	$\Phi_{2,2}$		$\Phi_{2,2}$	$\Phi_{2,2}$	=0	$\Phi_{2,2}$
Intercept-slope covariance	$\Phi_{1,2}$	$\Phi_{1,2}$		$\Phi_{1,2}$	$\Phi_{1,2}$	=0	$\Phi_{1,2}$
Residual factor variance	ψ	ψ		ψ		ψ	

defined to capture linear change (e.g., expected normative decline). The series of three-class SOGMMs allowed for differences among classes in similar manner.

The goal of the model selection step is to determine which model provides the best and most reasonable representation of the observed data. Given the exploratory nature of the GMM, this is often an iterative process guided by consideration of statistical fit and theoretical expectations. The statistical criteria included a collection of information regarding fit (e.g., BIC, Lo-Mendell-Rubin Likelihood Ratio Test), the separability of the latent classes (e.g., entropy, average latent class probabilities for most likely latent class membership), and the percentage of participants categorized into each class.

All models were fit to data using M*plus* 5.0 (L. K. Muthén & Muthén, 2007), and GMMs were fit using ten sets of random starting values. An annotated selection of programming scripts can be found at http://psychology.ucdavis.edu/labs/Grimm/personal/downloads.html.

RESULTS

The results are presented in four sections corresponding to the parts from which the model was built: longitudinal factor model, measurement invariance constraints, second-order growth modeling, and second-order growth mixture modeling. The first three sections are brief because they represent necessary precursors to second-order growth mixture modeling but were not the primary focus. Here we state what was found and concluded. In the last section, we describe the fit statistics, model choice decisions for the SOGMMs, and the chosen model.

Longitudinal Factor Model

The mother, father, and teacher reports of externalizing behavior were moderately to strongly related to the latent construct with standardized factor loadings ranging from .43 to .81. The factor loadings for teacher-reported behavior were lower than those for the mother- and father-reported externalizing behaviors, which may be due to the context of the measurement (e.g., school vs. home). The externalizing behavior factors were strongly correlated across elementary school with between-occasion correlations ranging between .86 and .93. The correlations among the uniquenesses were moderate to strong, ranging from .38 to .64, suggesting that a sizable part of the unique variance was specific variance as opposed to error variance.

Measurement Invariance Constraints

A series of longitudinal factor models were fit to test whether the relationship between the latent factor and the observed mother, father, and teacher reports held at all occasions and whether changes in the observed means could be carried by changes at the factor level. Without substantial loss of fit compared to the baseline (free) model, the strong invariance model fit the data well (χ^2 = 137, *df* = 48, CFI = .985, TLI = .976, RMSEA = .045 (.036 – .053)). Thus, we obtained confidence that the same latent construct had been measured at each occasion in the same metric.

Second-Order Growth Model

The second-order latent growth models were fit to establish a baseline model for comparison with the SOGMMs. Across a series of second-order growth models (e.g., intercept only, intercept plus linear slope, intercept plus latent basis slope) we found the linear growth model to be the best fitting (see Table 3 Model $M0_1$; χ^2 = 155, *df* = 40, CFI = .983, TLI = .978, RMSEA = .043 (.035 – .051), AIC = 64,929, BIC = 64,129, Adjusted BIC = 64,003, # of parameters = 40). The linear model had a significant decreasing trajectory (κ_2 = –.41) with significant between-child differences in the intercept ($\Phi_{1,1}$ = 18.88) and slope ($\Phi_{2,2}$ = .22).

TABLE 3
Fit Statistics for the Second-Order Linear Growth Model ($M0_1$) and Second-Order Growth Mixture Models ($M1_2$ – $M3_3$)

	$M0_1$	$M1_2$	$M2_2$[a]	$M3_2$[a,b,c]	$M1_3$[c]	$M2_3$[a]	$M3_3$[a]
Sample size							
N_1	1135	104.36	574.40	563.66	104.36	540.87	562.95
N_2	—	1030.64	560.60	571.34	341.24	221.13	282.68
N_3	—	—	—	—	689.40	373.00	289.37
Fit statistics							
Parameters	40	43	46	43	46	52	48
AIC	63,929	63,682	62,572	62,747	63,688	63,252	63,349
BIC	64,130	63,898	62,803	62,964	63,920	63,514	63,590
ABIC	64,003	63,762	62,657	62,827	63,773	63,348	63,437
Entropy	—	.883	.817	.812	.401	.576	.518

Note. AIC = Akaike Information Criteria; BIC = Bayesian Information Criteria; ABIC = sample-size Adjusted Bayesian Information Criteria. Sample sizes are the final class counts and proportions for the latent classes based on the estimated model.

[a]Convergence issues (e.g., negative variances, correlations > |1|) were encountered.

[b]Log likelihood was not replicated.

[c]Standard errors were not trustworthy.

Second-Order Growth Mixture Model

Iterative fitting and examination of statistical fit criteria for the baseline, two-, and three-class models, as well as theoretical expectations, suggested the presence of multiple classes (see Table 3). This process included wrestling with convergence issues (i.e., negative variances & correlations > |1|) and implications for model selection (issues 1 and 2 above). Model $M1_2$, a two-class model with class differences in the mean of the intercept and linear slope, was chosen as the best representation of the within-person changes in externalizing behaviors and the between-person differences therein. Among our considerations were that: $M1_2$ had a lower BIC than the one-class model, $M0_1$; the Bootstrap likelihood ratio test (p value < .001) favored $M1_2$ over $M0_1$, the entropy of $M1_2$ was high (.883), there was convergence and replication of the solution across multiple sets of random starting values, and the substantive interpretation of the parameters was acceptable.

Model $M1_2$ describes two classes. One class ($k = 2$) contained the vast majority (91%; $n = 1030.64$) of the sample. The average pattern of change of this normative group was characterized by a lower level of externalizing behavior in first grade, arbitrarily located at $\kappa_1 = 0$ for model identification purposes, that decreased significantly across time, $\kappa_2 = -.47$. The other class, $k = 1$, contained approximately 9% of the sample (n = 104.36). This class was characterized by a higher level of externalizing behavior at the first-grade observation, $\kappa_1 = 8.22$, and a stable, or nonsignificantly increasing trajectory across time $\kappa_2 = .68$. This class contained children with high and stable trajectories. Note that, although fixed around an arbitrary zero point given by the first class' mean, these parameters are in the metric of the original externalizing behavior measure, specifically the mothers' reports on the CBCL/4-18. The variance components of the model indicate that, within-class, there were significant between-person variance in the intercept, $\Phi_{1,1} = 10.20$, but relative homogeneity in the rate of change—decreasing for the normative class, and stability for the high stable class, $\Phi_{2,2} = .10$. These variance components and the nonsignificant negative correlation between the intercept and linear slope were invariant across latent classes. The expected mean trajectories (bold lines), and the 95% confidence intervals for the within-class between-person differences in change (shaded areas) are plotted in Figure 2. Through the visual representation it is possible to see the overlap between the latent classes present in first grade, and the emergence of normative and high stable groups as the children progressed through elementary school.

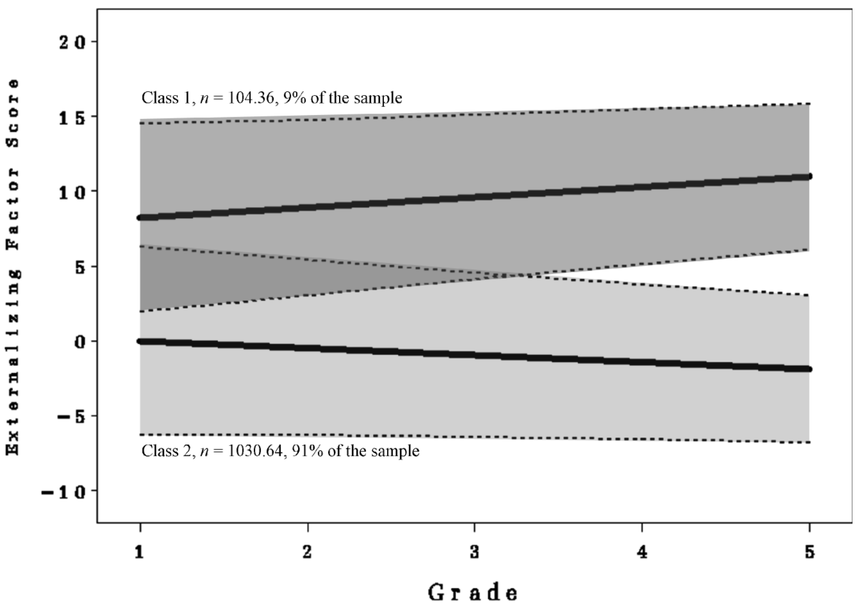

FIGURE 2 Expected mean trajectories and 95% confidence boundaries of expected within-class between-person differences for the two-class second-order growth mixture model, $M1_2$.

One way to help understand the classes obtained from a mixture model is to examine how they are related to other variables. Thus, child-level covariates were added to the model as predictors of class membership. These covariates included gender, maternal education, income status, maternal depression, and maternal sensitivity (measured when the child was 54 months old). Males, children whose mothers obtained less education, children whose mothers were less sensitive, and low-income children were more likely to be in the high stable class. Males were 134% more likely to be in the high stable class; low-income children were 36% more likely to be in the high stable class. Further class membership was associated with level of mother's education and sensitivity, such that additional units (years) of maternal education or maternal sensitivity were associated with a 20% and 12%, respectively, reduction in the odds of being in the high-stable class. Maternal depression was not significantly predictive of class membership.

DISCUSSION

Second-Order Growth Mixture Model and Developmental Science

Studying developmental change is complicated. Measurement of unobservable constructs is difficult, change patterns are unknown, and people differ from one another in unknown ways. We have proposed the SOGMM as useful tool for describing within-person change and between-person differences in change (cf. Baltes & Nesselroade, 1979). The model combines the benefits provided by the longitudinal common factor model, measurement invariance constraints, the latent growth model, and the mixture model. First, the longitudinal factor model allows for simultaneous use of information from multiple measures and allows for the separation of error/specific variance from true/common variance (Equation 1). Second, adding measurement invariance constraints allows for greater confidence that the same construct has been measured at all occasions and protects against changes in the meaning and scale of the factor across time (Equation 2). These portions of the model provide access to the many benefits derived from psychometric precision. Third, the latent growth model provides a framework for modeling of within-person change and the between-person differences therein (Equation 3). Finally, the mixture model allows greater flexibility for modeling the heterogeneity in change by incorporating a categorical latent variable (Equation 7) into the model. Specifically, the mixture allows for and accommodates more of the nonlinearity and non-normality present in many developmental processes (see also Masyn, this issue). Brought together, the components of the model provide immense flexibility in how within-person change and between-person differences in change can be examined and described.

Model Components

We presented and illustrated the application of the SOGMM as a combination of four specific model components. At the first-order level these included a longitudinal confirmatory factor model with three indicator variables assumed to be normally distributed along a continuous scale, and strong measurement invariance constraints. At the second-order level, we made use of a linear growth model and two- and three-class GMMs. For simplicity of presentation and in pursuing our analysis objectives we used only one of many possible combinations. As mentioned by Ram and Gerstorf (this issue) researchers may consider how each component could be replaced with other models of the same type. For example, rather than using a common factor model in the first-order level, other measurement models could be considered (e.g., item response model) as appropriate, depending on the distributions of the measured variables (e.g., dichotomous, polytomous, censored, ordinal, binary, count, zero-inflated Poisson, etc.). At the second-order level other types of change models could be invoked (Hoffman & Stawski, this issue; Masyn, this issue; Selig & Preacher, this issue). All such possibilities for extending the multi-component SOGMM should be explored and put to use.

Cautions and Caveats

Although the flexibility and possibilities provided by growth mixture modeling seem great, it is important to reiterate that these benefits come with a set of cautions and concerns that should not be ignored. As noted at the outset, GMMs, second-order or otherwise, should not be fit haphazardly without direction. The search for heterogeneity should be conducted in a principled and disciplined way. Sets of models and logical alternatives should be formulated and interpreted carefully. Even still, as noted in our model selection process, issues must be wrestled with and borderline decisions made.

At present, it is unknown how robust the SOGMM is to assumption violations. Simulation studies are necessary to identify the particular data conditions and circumstances that bound its use. However, comparing the results described here with results from fitting (first-order) GMMs directly to the three observed externalizing behavior scores (mother, father, & teacher reports) suggest that the second-order model is more conservative than the first-order model—fewer groups were obtained when using the SOGMM. Our cautious interpretation is that the first-order factor model and invariance constraints were able to help deal with some of the anomalies that might have masqueraded as groups in the usual application of growth mixtures to observed data.

One major concern with the use of GMMs is that they always identify groups. As illustrated and highlighted by Bauer and Curran (2003), even if the data have not been collected from or generated by multiple unobserved classes, the GMM can provide evidence of multiple groups. Pushing beyond the usual fitting of the means, variances, and covariances of the observed data (first- and second-order moments), GMMs allows for description of higher order moments: skew and kurtosis. This is appealing: The models can provide more complete descriptions of additional aspects of the data. On their own, though, better descriptions of the data do not necessarily bring us closer to the underlying mechanisms that caused the data. Of particular concern is the possibility that the skew and kurtosis being described by GMMs may in reality be the result of measurement or other anomalies and incorrectly attributed to a typology of individuals.

CONCLUDING REMARKS

We forward the SOGMM as a promising possibility for examining developmental change, one that capitalizes on the benefits of multivariate assessment and enables researchers to relax (and test) some of the assumptions underlying the usual applications of growth curve models, in particular, that all participants have the same change pattern and that the between-person differences in change are continuous and normally distributed. We are confident that, by focusing on measurement issues and dealing with them at the first-order level, the SOGMM can do its job better—helping developmentalists untangle the mechanisms and processes that underlie how and why individuals change in different and interesting ways.

ACKNOWLEDGMENT

We would like to thank Jack McArdle, John Nesselroade, Fumiaki Hamagami, Bob Pianta, and our colleagues at the Center for the Advanced Study of Teaching and Learning, the Jefferson Psychometric Laboratory, and the Center for Developmental and Health Research Methodology at the University of Virginia for their helpful comments on this work. The first author would also like to thank John Horn for the conversations we had during the Society of Multivariate Experimental Psychology Conference in 2003 about growth mixture models and possible extensions – one of which was a second-order model. The first author was supported by the National Center for Research in Early Childhood Education, Institute of Education Sciences, U.S. Department of Education (R305A06021). The opinions expressed are those of the authors and do not represent views of the U.S. Department of Education.

REFERENCES

Achenbach, T. M. (1991a). *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Burlington: University of Vermont, Department of Psychiatry.

Achenbach, T. M. (1991b). *Manual for the Teacher Report Form and 1991 Profile*. Burlington: University of Vermont, Department of Psychiatry.

Baltes, P. B., & Nesselroade, J. R. (1979). History and rationales for longitudinal research. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (pp. 1–39). New York: Academic Press.

Bauer, D. J. (2007). Observations on the use of growth mixture models in psychological research. *Multivariate Behavioral Research*, *42*, 757–786.

Bauer, D. J., & Curran, P. J. (2003). Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychological Methods, 8*, 338–363.

Bongers, I. L., Koot, H. M., van der Ende, J., & Verhulst, F. C. (2004). Developmental trajectories of externalizing behavior in childhood and adolescence. *Child Development*, *75*, 1523–1537.

Browne, M. W. (1993). Structured latent curve models. In C. M. Cuadras & C. R. Rao (Eds.), *Multivariate analysis: Future directions 2* (pp. 171–198). Amsterdam: North-Holland.

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

Bub, K. L., McCartney, K., & Willett, J. B. (2007). Behavior problem trajectories and first-grade cognitive ability and achievement skills: A latent growth curve analysis. *Journal of Educational Psychology*, *99*, 653–670.

Connell, A. M., & Frye, A. A. (2006). Growth mixture modelling in developmental psychology: Overview and demonstration of heterogeneity in developmental trajectories of adolescent antisocial behaviour. *Infant and Child Development*, *15*, 609–621.

Cronbach, L. J., & Furby, L. (1970). How we should measure “change”: Or should we? *Psychological Bulletin*, *74*, 68–80.

Curran, P. J., Edwards, M. C., Wirth, R. J., Hussong, A. M., & Chassin, L. (2007). The incorporation of categorical measurement models in the analysis of individual growth. In T. D. Little, J. A. Bovaird, & N. A. Card (Eds.), *Modeling contextual effects in longitudinal studies* (pp. 89–120). Mahwah, NJ: Erlbaum.

Dekovic, M., Buist, K. L., & Reitz, E. (2004). Stability and changes in problem behavior during adolescence: Latent growth analysis. *Journal of Youth and Adolescence*, *33*, 1–12.

Edwards, M. C., & Wirth, R. J. (2009). Measurement and the study of change. *Research in Human Development*, *6*(2–3), 74–96.

Ferrer, E., Balluerka, N., & Widaman, K. F. (2008). Factorial invariance and the specification of second-order latent growth models. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *4*, 22–36.

Ferrer, E., Hamagami, F., & McArdle, J. J. (2004). Modeling latent growth curves with incomplete data using different types of structural equation modeling and multilevel software. *Structural Equation Modeling*, *11*, 452–483.

Ghisletta, P., & Lindenberger, U. (2004). Static and dynamic longitudinal structural analyses of cognitive changes in old age. *Gerontology*, *50*, 12–16.

Gilliom, M., & Shaw, D. S. (2004). Codevelopment of externalizing and internalizing problems in early childhood. *Developmental and Psychopathology*, *16*, 313–333.

Grimm, K. J., & Ram, N. (in press). Non-linear growth models in Mplus and SAS. *Structural Equation Modeling.*

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

Hertzog, C., Lindenberger, U., Ghisletta, P., & von Oertzen, T. (2006). On the power of multivariate latent growth curve models to detect correlated change. *Psychological Methods*, *11*, 244–252.

Hill, A. L., Degnan, K. A., Calkins, S. D., & Keane, S. P. (2006). Profiles of externalizing behavior problems for boys and girls across preschool: The role of emotion regulation and inattention. *Developmental Psychology*, *42*, 913–928.

Hipp, J. R., & Bauer, D. J. (2006). Local solutions in the estimation of growth mixture models. *Psychological Methods*, *11*, 36–53.

Hoffman, L., & Stawski, R. S. (2009). Persons as contexts: Evaluating between-person and within-person effects in longitudinal analysis. *Research in Human Developmen*, *6*(2–3), 97–120.

Keiley, M. K., Bates, J. E., Dodge, K. A., & Pettit, G. S. (2000). A cross-domain growth analysis: Externalizing and internalizing behavior during 8 years of childhood. *Journal of Abnormal Child Psychology*, *28*, 161–179.

Leve, L. D., Kim, H. K., & Pears, K. C. (2005). Childhood temperament and family environment as predictors of internalizing and externalizing trajectories from ages 5 to 17. *Journal of Abnormal Child Psychology*, *33*, 505–520.

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

Lubke, G. H., & Muthén, B. O. (2007). Performance of factor mixture models as a function of model size, covariate effects, and class-specific parameters. *Structural Equation Modeling*, *14*, 26–47.

Masyn, K. E. (2009). Discrete-time survival factor mixture analysis for low-frequency recurrent event histories. *Research in Human Development*, *6*(2–3), 165–194.

McArdle, J. J. (1988). Dynamic but structural equation modeling of repeated measures data. In J. R. Nesselroade & R.B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (Vol. 2, pp. 561–614). New York: Plenum.

McArdle, J. J. (2001). A latent difference score approach to longitudinal dynamic structural analyses. In R. Cudeck, S. du Toit, & D. Sorbom (Eds.), *Structural equation modeling: Present and future* (pp. 342–380). Lincolnwood, IL: Scientific Software International.

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

McArdle, J. J., Grimm, K. J., Hamagami, F., Bowles, R. P., & Meredith, W. (in press). Modeling lifespan growth curves of cognition using longitudinal data with changing measures. *Psychological Methods*.

McArdle, J. J., & Hamagami, F. (1996). Multilevel models from a multiple group structural equation perspective. In G. A. Marcoulides & R. E. Schumaker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 89–124). Mahwah, NJ: Erlbaum.

McLachlan, G. J., & Peel, D. (2000). *Finite mixture models*. New York: John Wiley & Sons.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*, 525–543.

Meredith, W., & Horn, J. L. (2001). The role of factorial invariance in modeling growth and change. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 203–240). Washington, DC: American Psychological Association.

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

Miner, J. L., & Clarke-Stewart, K. A. (2008). Trajectories of externalizing behavior from age 2 to age 9: Relations to gender, temperament, ethnicity, parenting, and rater. *Developmental Psychology*, *44*, 771–786.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*, 674–701.

Moffitt, T. E., Caspi, A., Dickson, N., Silva, P., & Stanton, W. (1996). Childhood-onset versus adolescent-onset antisocial conduct problems in males: Natural history from ages 3 to 18 years. *Development and Psychopathology*, *8*, 399–424.

Muthén, B. (2004). Latent variable analysis. Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (ed.), *The Sage handbook of quantitative methodology for the social sciences* (pp. 345–368). Thousand Oaks, CA: Sage.

Muthén, B. O., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analysis: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*, *24*, 882–891.

Muthén, L. K., & Muthén, B. O. (2007). *Mplus user's guide, fifth edition*. Los Angeles: Muthén & Muthén.

Muthén, B. O., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, *55*, 463–469.

Nagin, D. S. (1999). Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychological Methods*, *4*, 139–157.

Nesselroade, J. R., & Cable, D. G. (1974). "Sometimes it's okay to factor difference scores" – The separation of state and trait anxiety. *Multivariate Behavioral Research*, *9*, 273–282.

NICHD Early Child Care Research Network. (2002). The relation of global first-grade classroom environment to structural classroom features and teacher and student behaviors. *Elementary School Journal*, *102*, 367–387.

NICHD Early Child Care Research Network. (2004). Person-centered trajectories of physical aggression. *Monographs of the Society for Research in Child Development*, *69*, 42–49.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, *14*, 535–569.

Odgers, C. L., Moffitt, T. E., Broadbent, J. M., Dickson, N., Hancox, R. J., Harrington, H., et al. (2008). Female and male antisocial trajectories: From childhood origins to adult outcomes. *Development and Psychopathology*, *20*, 673–716.

Preacher, K. J., Wichman, A. L., MacCallum, R. C., & Briggs, N. E. (2008). *Latent growth curve modeling*. Thousand Oaks, CA: Sage.

Ram, N., & Gerstrof, D. (2009). Methods for the study of development – developing methods: Introduction to a special issue. *Research in Human Development*, *6*(2–3), 61–73.

Ram, N., & Grimm, K. J. (2007). Using simple and complex growth models to articulate developmental change: Matching theory to method. *International Journal of Behavioral Development*, *31*, 303–316.

Ram, N., & Grimm, K. J. (2009). *Growth mixture modeling: A method for identifying differences in longitudinal change among unobserved groups*. Manuscript submitted for publication.

Rao, C. R. (1958). Some statistical methods for the comparison of growth curves. *Biometrics*, *14*, 1–17.

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

Selig, J. P., & Preacher, K. J. (2009). Mediation models for longitudinal data in developmental research. *Research in Human Development*, *6*(2–3), 144–164.

Shaw, D. S., Gilliom, M., Ingoldsby, E. M., & Nagin, D. S. (2003). Trajectories leading to school-age conduct problems. *Developmental Psychology*, *39*, 189–200.

Small, B. J., & Bäckman, L. (2007). Longitudinal trajectories of cognitive change in preclinical Alzheimer's disease: A growth mixture modeling analysis. *Cortex*, *43*, 826–834.

Tucker, L. R (1958). Determination of parameters of a functional relation by factor analysis. *Psychometrika*, *23*, 19–23.

Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Application in the substance use domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281–324). Washington, DC: American Psychological Association.