University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Developmental Cognitive Neuroscience Laboratory - Faculty and Staff Publications | Developmental Cognitive Neuroscience Laboratory

1-1-2009

Growth Mixture Modeling of Academic Achievement in Children of Varying Birth Weight Risk

Kimberly Espy
University of Nebraska, kespy2@unl.edu

Julia Hua Fang
University of Nebraska, jfang2@unl.edu

David Charak
Learning Point Associates, Naperville, Illinois

Nori Minich
Case Western Reserve University

H. Gerry Taylor
Case Western Reserve University

Follow this and additional works at: http://digitalcommons.unl.edu/dcnlfacpub

Part of the Neurosciences Commons

Espy, Kimberly; Fang, Julia Hua; Charak, David; Minich, Nori; and Taylor, H. Gerry, "Growth Mixture Modeling of Academic Achievement in Children of Varying Birth Weight Risk" (2009). *Developmental Cognitive Neuroscience Laboratory - Faculty and Staff Publications.* Paper 47.
http://digitalcommons.unl.edu/dcnlfacpub/47

This Article is brought to you for free and open access by the Developmental Cognitive Neuroscience Laboratory at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Developmental Cognitive Neuroscience Laboratory - Faculty and Staff Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Published in *Neuropsychology* 23:4 (2009), pp. 460–474; doi: 10.1037/a0015676 Copyright © 2009 American Psychological Association. Used by permission. http://www.apa.org/pubs/journals/neu/ "This article may not exactly replicate the final version published in the APA journal. It is not the copy of record."

This research was supported in part by Grants HD026554 and HD050399 from the National Institute of Child Health and Development to the senior author and his colleague Maureen Hack, and by Grants R01 MH065668 from the National Institute of Mental Health, R01 DA014661 from the National Institute on Drug Abuse, and P01 HD038051 from the National Institute of Child Health and Development to the first author. We acknowledge Maureen Hack for her vital contributions to the design, conduct, and oversight of the longitudinal project. We also thank the participating families, project staff, and graduate students who assisted in various tasks associated with this study.

Submitted June 20, 2008; revised December 29, 2008; accepted February 17, 2009.

Growth Mixture Modeling of Academic Achievement in Children of Varying Birth Weight Risk

Kimberly Andrews Espy, Office of Research and the Department of Psychology, University of Nebraska–Lincoln

Hua Fang, Office of Research and the Department of Psychology, University of Nebraska–Lincoln

David Charak, Learning Point Associates, Naperville, Illinois

Nori Minich, Department of Pediatrics, Case Western Reserve University

H. Gerry Taylor, Department of Pediatrics, Case Western Reserve University, & Department of Pediatrics, Rainbow Babies & Children's Hospital, University Hospitals Case Medical Center, Cleveland, Ohio

Corresponding author – Kimberly Andrews Espy, Associate Vice Chancellor for Research, 303 Canfield Administration Building, University of Nebraska–Lincoln, Lincoln, NE 68688–0433; e-mail kespy2@unl.edu

Abstract

The extremes of birth weight and preterm birth are known to result in a host of adverse outcomes, yet studies to date largely have used cross-sectional designs and variable-centered methods to understand long-term sequelae. Growth mixture modeling (GMM) that utilizes an integrated person- and variable-centered approach was applied to identify latent classes of achievement from a cohort of school-age children born at varying birth weights. GMM analyses revealed 2 latent achievement classes for calculation, problem-solving, and decoding abilities. The classes differed substantively and persistently in proficiency and in growth trajectories. Birth weight was a robust predictor of class membership for the 2 mathematics achievement outcomes and a marginal predictor of class membership for decoding. Neither visuospatial-motor skills nor environmental risk at study entry added to class prediction for any of the achievement skills. Among children born preterm, neonatal medical variables predicted class membership uniquely beyond birth weight. More generally, GMM is useful in revealing coherence in the developmental patterns of academic achievement in children of varying weight at birth and is well suited to investigations of sources of heterogeneity.

Keywords: growth mixture modeling, latent class identification, prematurity, low birth weight, academic achievement

Modern advances in perinatal care (e.g., assisted ventilation techniques for neonates and maternal antenatal steroid medications) have led to the survival of increasing numbers of children who are born very preterm (<32 weeks gestational age) and/or with very low birth weight (VLBW; <1,500 g, 3 lbs. 5 oz). The increased survival of prematurely born infants has been most dramatic in children born at the lower extreme of birth weight (Hack & Fanaroff, 1999). Children with VLBW have higher rates of cognitive and academic deficits, behavior problems, and neurosensory and other health disorders than do term-born children of normal birth weights (Taylor, Klein, & Hack, 2000). These adverse sequelae are more common and more severe in children of the lower extreme of birth weight or gestational age (Hack, Klein, & Taylor, 1996; Klebanov, Brooks-Gunn, & McCormick, 1994). Although modern neonatal care has contributed to a lower rate of intraventricular hemorrhage (Wilson-Costello et al., 2007), increases in survival have not been accompanied by decreases in the rates of other major neonatal and postnatal medical complications. In fact, any benefit of improved neonatal care has been offset by the survival of higher risk, lower birth weight infants (Anderson & Doyle, 2003; Taylor, Klein, Drotar, Schluchter, & Hack, 2006).

Neuropsychological methods detect more subtle cognitive weaknesses or "hidden" handicaps that accompany these early complications. Even early in life, the adverse developmental effects of VLBW are evident, including lower overall mental and motor skills; reduced visual recognition memory; and poorer language, executive, and attentional skills (Espy et al., 2002; Goyen, Lui, & Woods, 1998; Landry, Smith, Miller-Loncar, & Swank, 1997; Sullivan & McGrath, 2003). A substantial body of literature has documented neuropsychological deficits during the preschool years and early primary school that persists throughout the school-age years (Anderson, Doyle, & Victorian Infant Collaborative Study Group, 2004; Botting, Powls, Cooke, & Marlow, 1998; Friske & Whyte, 1994; Luoma, Herrgard, Martikainen, & Ahonen, 1998; Msall, Buck, Rogers. & Catanzaro, 1992; Saigal, 2000; Taylor, Minich, Bangert, Filipek, & Hack, 2004; Wolke & Meyer, 1999). Deficits in nonverbal skills, perceptual-motor abilities, executive control, and attention cannot be attributed entirely to overall mental deficiency or neu-

rosensory handicaps (Anderson et al., 2004; Hack et al., 1992; Taylor, Klein, Minich, & Hack, 2000). Furthermore, these children have more learning and social-behavioral problems than term-born children (Bhutta, Cleves, Casey, Cradock, & Anand, 2002; Hille et al., 2001; Klein, Hack, & Breslau, 1989), including lower adaptive behavior skills and social competence (Saigal, Pinelli, Hoult, Kim, & Boyle, 2003); more internalizing and externalizing symptomatology (Botting et al., 1997; Breslau & Chilcoat, 2000; Szatmari, Saigal, Rosenbaum, Campbell, & King, 1990; Whitaker et al., 1997); and higher rates of learning disabilities. Mathematics disabilities are particularly prominent (Anderson & Doyle, 2003; Espy, McDiarmid, et al., 2004; Litt, Taylor, Klein, & Hack, 2005; Taylor, Hack, Klein, & Schatschneider, 1995). Children with VLBW also have lower levels of academic achievement and higher rates of grade repetition and special education than term-born controls (Klebanov, Brooks-Gunn & McCormick, 1994; Saigal, 2000; Taylor, Klein, Hack, 2000; Taylor, Klein, Minich, et al., 2000).

Neurobiological risks, such as the degree of low birth weight, abnormalities on neonatal cranial ultrasounds, chronic lung disease, septicemia, and composite biological risk indexes account for substantial variability in outcome (Hack et al., 1992; Hack et al., 2000; Koller, Lawson, Rose, Wallace, & McCarton, 1997; Landry, Fletcher, Denson, & Chapieski, 1993; Liaw & Brooks-Gunn, 1993; McGrath & Sullivan, 2002; Taylor et al., 2006; Taylor, Klein, Schatschneider, & Hack, 1998) Adverse outcomes are observed even in children without major neonatal complications, making it challenging to identify those at risk (Espy et al., 2002; Taylor et al., 2000, 2006). Environmental risks, including sociodemographic characteristics and financial disadvantage, as well as more "proximal" family influences such as family functioning, negative life events, and maternal psychological distress also are related to outcome (Bendersky & Lewis, 1995; Breslau, 1995; Breslau & Chilcoat, 2000; Taylor et al., 1998, 2006).

Adding to these complexities is the differences in outcome that unfold across development. Because cross-sectional designs are used most often in outcome studies, it is difficult to determine whether the pattern of observed weaknesses is stable over time, resolves with age, or worsens with development as the brain areas most compromised by early insult become more engaged for skill acquisition and maintenance. In some studies, a relatively stable pattern of weaknesses has been observed across school age into adolescence (Breslau et al., 2001; Powls, Botting, Cooke, & Marlow, 1995; Rickards, Ryan, & Kitchen, 1988). Other results suggest an exacerbation of impairment into adolescence (Botting et al., 1998; Cohen et al., 1996; O'Callaghan et al., 1996; Saigal, Hoult, Streiner, Stoskopf, & Rosenbaum, 2000; Taylor, Klein, Minich, & Hack, 2000; Zelkowitz, Papageorgiou, Zelazo, & Weiss, 1995). Finally, in one study, initial reductions on a test of vocabulary in young children with VLBW relative to term-born peers diminished across development, with the two groups obtaining similar scores by at age 8 years (Ment et al., 2003).

Although weaknesses in visuospatial-motor abilities, attention, executive control, memory, and academic achievement are commonly identified in studies of children born very early and at VLBW, these investigations have used "variable-centered" approaches (B. Muthén & Muthén, 2000), in which the goal is to relate pre-established risk factors to outcomes of interest. Studies exemplifying this approach are ones that examine neuropsychological proficiencies as a function of birth weight, neonatal complications, and environmental disadvantage. In contrast, "person-centered" approaches such as growth mixture modeling (GMM), although also incorporating the variable-centered approach, use cluster or latent class analyses. These approaches address questions on relations among individuals, in which the interest is to subgroup persons with similar outcomes and understand how subgroups differ from one another. Person-centered approaches are particularly well suited to study of outcomes of VLBW because individuals classified according to birth weight differ substantially in other risk factors, such as perinatal complications. This approach may also be ideal for teasing apart developmental trajectories associated with low base rate phenomena, such as the specific medical conditions that can accompany VLBW.

Fortunately, recent statistical advances have resulted in the application of both variable- and person-centered techniques to modeling of growth for longitudinally collected data. Conventional growth modeling (CGM) is a variable-centered approach that can be conducted using structural equations models (SEM), mixed linear models, or hierarchical linear (HLM) models, which are known generally as multilevel models (Hedeker & Gibbons, 1994; McCulloch & Searle, 2001; B. Muthén, 2004; Raudenbush, 2001; Raudenbush & Bryk, 2002; Singer & Willett, 2003; Skrondal & Rabe-Hesketh, 2004). Although CGM is accomplished somewhat differently in SEM and HLM methods, the results are identical when the same growth parameters are modeled. These approaches now are applied routinely to address developmental questions, including changes in brain responses to auditory stimuli across early development (Espy, Molfese, Molfese, & Modglin, 2004) and variation in longitudinal neuropsychological outcomes in children with VLBW (Taylor, Minich, Klein, & Hack, 2004).

The major objective of the present study was to apply person-centered growth modeling techniques to better understand individual variation in the development of academic skills in children with VLBW. Learning difficulties at school age increase risks for long-term problems in behavior adjustment and limited educational and vocational attainments (Ewing-Cobbs et al., 2004; Klebanov, Brooks-Gunn, & McCormick, 1994). Identifying different patterns of growth in academic achievement during the school-age years and risk factors associated with these patterns is thus critical to determine which children may need the most extensive early interventions to promote skill development. Previous studies have documented persistent deficits in academic skills (Hack, 2005), but it is unclear if these deficits are stable over time or if the gap in skills between children with VLBW and their term-born peers widens with age. Several longitudinal studies have suggested relatively constant deficits in academic achievement among VLBW cohorts across the school-age years (Breslau, Paneth, & Lucia, 2004; Schneider, Wolke, Schlagmuller, & Meyer, 2004), whereas others suggest that these deficits may become more pronounced with age (Saigal et al., 2000; Taylor, Klein, Minich, & Hack, 2000). It is also important to investigate factors other than VLBW that may contribute to risks for poor achievement and to determine whether growth in achievement is affected more adversely in some children than in others.

In previous reports on outcomes for the sample of children followed in the present study, Taylor and colleagues (Taylor et al., 1995, 2000) assessed achievement in two groups of children varying in the degree of VLBW (<750 g and 750 to 1, 499 g) and

term-born controls at the mean ages of 7 and 11 years. Results revealed lower scores for the <750 g group compared with term controls on reading and mathematics at both follow-up assessments. Taylor et al. (2000) also observed that the <750 g group made less positive gains in reading across the two assessments than the term group. However, changes in achievement across subsequent follow-ups at later ages were not examined in these earlier reports. Multiple assessments of achievement as part of this larger study provided a unique opportunity to examine longer term achievement outcomes into adolescence. Application of both variable- and person-centered GMM methods (B. Muthén & Muthén, 2000) also enabled better characterization of differences in growth of academic skills with age and the correlates of these individual differences.

Specific aims were to (a) empirically characterize the developmental trajectories of differing academic skills across childhood into adolescence; and (b) determine whether birth weight, nonverbal neuropsychological abilities, and environmental risk measured at early school age predicted different patterns of academic proficiencies. Based on past literature, we anticipated that birth weight would be a robust predictor of class membership, and that nonverbal neuropsychological skills and environmental risk in early childhood would contribute substantively to the prediction of class membership (Taylor et al., 1995). Finally, we hypothesized that among children with VLBW, neonatal medical conditions would predict class membership beyond birth weight alone (Taylor et al., 1998).

Method

Participants

The total sample consisted of 196 children, 67 children born at term (>36 weeks) of normal birth weight (>2,500 g) and 129 children born preterm and VLBW at <1,499 g. Most of the children were recruited into the study at early school age (Hack et al., 1996), with a few additional children recruited at the second follow-up assessment to maximize sample size for study of developmental change (Taylor, Minich, Klein, et al., 2004). Because there are fewer children born at the lowest end of the birth weight spectrum, children <750 g (n = 64) were the sampled "target" participant, representing 93% of the survivors in this range of birth weight born from July 1, 1982 through December 31, 1986 in the six-county region surrounding Cleveland, Ohio. Children in the 750 to 1,499 g (n = 65) and term-born groups were individually matched with a <750 g child based on birth date (within 3 months), race, and either the same hospital of birth (750 to 1,499 g children) or same school (term children).

Table 1 presents the sample background and perinatal characteristics. The number of girls and boys were approximately equal between term and higher (750 to 1,499 g) and lower (<750 g) weight children with VLBW, $\chi^2(2, N = 196) = 0.066$, $p > .97$. The number of children of minority race also did not differ by birth weight group, $\chi^2(2, N = 196) = 0.092$, $p > .95$. Finally, birth weight groups did not differ in socioeconomic status (SES), $F(2, 175) = 0.25$, $p > .78$, calculated using the Hollingshead index (Hollingshead, 1957) composite of parent education and occupation (reversed scored so that higher score reflected higher SES) and then normalized within the sample. Not surprisingly, the birth weight groups differed in length of hospitalization, $F(1, 126) = 43.28$, $p < .001$; days on the ventilator, $F(1, 126) = 32.10$, $p < .001$; and number of children with chronic lung disease, $\chi^2(1, N = 128) = 19.14$, $p < .001$; septicemia, $\chi^2(1, N = 128) = 5.16$, $p = .024$; and apnea $\chi^2(1, N = 129) = 5.44$, $p = .023$. Further sample information is provided in Taylor, Minich, Klein, and Hack (2004).

Table 1. Sample Characteristics

	Birth weight group		
Sample characteristics	Term[a] M/n (SD/%)	750 to 1,499 g[b] M/n (SD/%)	<750 g[c] M/n (SD/%)
Sex			
Boys	23 (34)	21 (33)	21 (32)
Girls	44 (66)	43 (67)	44 (68)
Race			
Black	34 (51)	31 (48)	33 (51)
White	33 (49)	33 (52)	32 (49)
SES	0.05 (1.01)	0.02 (1.06)	–0.07 (1.00)
LISRES–A Environmental Risk	50.33 (6.02)	49.21 (4.55)	49.79 (5.81)
VSPM Factor Score*	0.07 (0.96)	0.22 (1.02)	–0.29 (0.97)
Birth weight (g)***	3,370 (575)	1,179 (212)	670 (67)
Gestational age (weeks)***	—	29.5 (2.3)	25.8 (1.8)
Length of hospitalization (days)***	—	57.2 (36.59)	126.30 (75.67)
Days on ventilation***	—	9.7 (18.76)	43.38 (43.69)
Apnea*	—	47 (73.4)	58 (89.2)
Chronic lung disease***	—	6 (9.4)	27 (42.2)
Jaundice	—	25 (39.7)	18 (28.6)
Septicemia*	—	15 (23.4)	27 (42.2)
Necrotizing enterocolitis	—	5 (7.8)	4 (6.3)

N = 196. SES = socioeconomic status z score; LISRES–A = Life Stressors and Social Resources Inventory–Adult Form; VSPM = visuospatial/perceptual motor; Chronic lung disease = oxygen dependence for ≥ 36 weeks corrected age; jaundice = maximal indirect serum bilirubin > 10 mg/dL.

[a] n = 67. [b] n = 64. [c] n = 65.

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

A sequential panel design (Mehta & West, 2000; B. Muthén, Khoo, Francis, & Boscardin, 2003; Nesselroade & Baltes, 1979; Tonry, Ohlin, & Farrington, 1991) was used, where participants were enrolled in early elementary school, around age 7 years (M = 6.87 years, SD = 0.93, range 5.31 to 9.34 years). Children were assessed approximately 4 years after the initial visit and annually for four subsequent assessments. Given the variability in age at study entry, the respective follow-up intervals spanned from ages 10 through 16 years, shown in Table 2. Age-related change in children's academic achievement scores was modeled because the interest is in growth across development, not change between visits. Because variability in the age at entry was large relative to the other ages, the difference in the actual age at enrollment from age 7 was used as a covariate for this assessment only. Because of this sampling, subsequent assessment schedule, and some attrition (sample retention = 92%), there were a different number of assessments at each age period, resulting in an unbalanced design. At each age, however, the mean ages and sample sizes were approximately equal across the three birth weight groups (see Table 2).

Outcome Measures

Three subtests from the Woodcock–Johnson Psycho-Educational Battery–Revised (WJ–R) Tests of Achievement (Woodcock & Johnson, 1989) were administered to measure academic achievement outcome. Calculation requires the examinee to perform mathematic operations that vary in difficulty. In Applied Problems, examinees analyze and solve practical mathematics problems. Letter-Word Identification assesses reading decoding by requiring examinees to orally read a list of single words of increasing difficulty. These subtests were chosen for their demonstrated high reliability and validity and because the resultant W scores are Rasch model-derived values that represent equal-interval measurement both within and across individuals, where any given difference along the scale has the same implication for performance at any level or age, a desirable property for growth modeling. For simplicity, we refer to scores on Calculation, Applied Problems, and Letter-Word Identification as Calculation, Problem-Solving, and Decoding, respectively.

Predictors

The central predictor was birth weight (BWT), a continuous variable measured in grams. At the initial evaluation at study entry, children were administered a neuropsychological battery that included the tetrad short-form of the Kaufman Assessment Battery for Children (K–ABC; Kaufman & Applegate, 1988), as well as tests of picture naming, verbal short-term memory and verbal comprehension, perceptual-motor skills, and attention and executive function (Taylor et al., 1995). A principal axis factor analysis with varimax rotation was conducted on age-standardized scores, as described in Taylor, Burant, Holding, Klein, and Hack (2002), to reduce the number of predictors and to identify distinct cognitive constructs. Tests with low primary loadings or high cross-loadings were excluded after the initial analysis. The final factor analysis yielded two factors accounting for 63% of the variance in scores. Factor 1 had an eigenvalue of 4.46 and explained 50% of the variance in scores, and Factor 2 had an eigenvalue of 1.21 and explained 13% of the variance in scores. Tests loading on Factor 1, referred to as the visuospatial/perceptual-motor factor (VSPM), included the Developmental Test of Visual-Motor Integration (Beery, 1989), the short-form of the Test of Motor Proficiency (Bruininks, 1978), the Purdue Pegboard (Gardner, 1979), the Computerized Test of Attention (Murphy-Berman & Wright, 1987) and Triangles and Matrix Analogies subtests of the K–ABC. Tests loading on Factor 2, referred to as the verbal memory factor, included the Pseudoword Repetition Test (Taylor, Lean, & Schwartz, 1989), Recalling Sentences subtest of the Clinical Evaluation of Language Fundamentals–Revised (Semel, Wiig, Secord, & Sabers, 1987), and Word Order subtest of the K–ABC. Factor composites were computed by averaging the age-adjusted standard scores of the constituent tests. Because VSPM and the verbal memory factor were correlated significantly ($\rho = 0.32$, $p < .001$) and VSPM accounted for the greatest variance, only VSPM was retained as a predictor of class membership in the subsequent analyses.

To assess the contribution of the child's social environment to class membership, the Life Stressors and Social Resources Inventory–Adult Form (LISRES–A; Moos & Moos, 1994) administered at study entry was used as an index of proximal life stressors and social resources (Taylor, Minich, Klein, et al., 2004). A summary score of environmental risk (ER) was created from the mean of the T scores for six stressors scales (health, work, spouse, extended family, friends, and negative life events). Because the daily care for a preterm child can contribute to perceived family stress, items pertinent to the child were removed in computing the summary score for all participants. The ER score was used as the predictor of class mem-

Table 2. Children's Age by Birth Weight Groups for Each Assessment

Age period	Actual age M (SD)	N at each age period	Term age M (SD) (n)	750 to 1,499 g age M (SD) (n)	<750 g age M (SD) (n)
7	6.87 (0.93)	178	7.00 (1.00) (52)	6.91 (0.89) (64)	6.72 (0.92) (62)
10	9.98 (0.33)	66	10.02 (0.31) (22)	9.92 (0.35) (22)	9.99 (0.32) (22)
11	11.02 (0.31)	107	11.04 (0.31) (38)	10.97 (0.31) (32)	11.03 (0.30) (37)
12	12.01 (0.30)	127	11.99 (0.31) (47)	11.96 (0.31) (36)	12.07 (0.27) (44)
13	13.02 (0.30)	142	12.98 (0.29) (52)	13.01 (0.33) (43)	13.05 (0.28) (47)
14	13.94 (0.31)	90	13.96 (0.33) (35)	13.91 (0.31) (27)	13.94 (0.28) (28)
15	15.15 (0.27)	61	15.12 (0.35) (20)	15.18 (0.22) (22)	15.15 (0.30) (19)
16	16.12 (0.37)	103	16.11 (0.37) (35)	16.14 (0.40) (36)	16.10 (0.34) (32)

Term = full-term; 750 to 1,499 g = 750 to 1,499 g birth weight; <750 g = <750 g birth weight.

bership, where higher scores reflected more stressful environments. There was no difference in the ER score among the birth weight groups, $F(2, 193) = 0.68$, $p = .51$.

Several neonatal medical variables were selected as predictors of class membership: two continuously distributed variables, Length of Hospital (in days) and Days of Ventilation, and four categorical variables, Apnea, Chronic Lung Disease, Jaundice, and Necrotizing Enterocolitis (coded 1 for children who experienced the medical condition and 0 for those who did not).

GMM

Like the conventional HLM and SEM approaches to growth modeling, GMM can be used to examine a mean growth trajectory and individual variation within a population, considering both the person-centered and variable-centered approaches (B. Muthén & Muthén, 2000). Unlike conventional growth models, however, GMM utilizes a "mix" of latent continuous and categorical variables, and is used to identify meaningful subpopulations within the larger population to examine the mean trajectories and individual variation across and within the subpopulations. More important, these subpopulations are not known a priori, but rather are determined empirically, termed "latent classes." Individuals are assigned to subpopulations or latent classes based on their posterior probabilities using multinomial logistic regression (B. Muthén & Shedden, 1999). Latent class analysis (Nagin, 1999; Nagin & Tremblay, 2001) is similar to GMM in terms of identifying the latent classes and modeling the growth trajectory across classes, but GMM has the added advantage of allowing for within-class variation of variable-centered methods (B. Muthén, 2006; Nagin & Tremblay, 2005). GMM is the "second generation" of SEM-parameterized growth models, which fully incorporates the multilevel approach to understand nested, individual variation (B. Muthén, 2001a, 2001b, 2002). GMM has the same advantages as CGM in accommodating missing data, and thus can be applied to unbalanced designs.

GMM starts from conventional growth models to identify the growth functions (e.g., linear or quadratic). In GMM, the CGM assumption that all participants are drawn from a single population with common population parameters (e.g., intercepts, slopes, or acceleration) then is relaxed. GMM uses latent categorical variables to allow for the parameter variation within and across unobserved latent classes. Figure 1 depicts a general diagram of the GMM as applied to our data. Double-arrowed curved lines represent growth factor covariance and single-arrowed lines represent estimated path values. Each growth factor in the respective circles has indicators Y_1 to Y_8 representing reading or math scores at the 8 ages (7, 10, 11, 12, 13, 14, 15, and 16 years), respectively. Residuals are represented by ε_{1-8} in squares, where Age7 in the square represents the age covariate for the initial assessment Y_1. The three predictors of interest BWT, VSPM, and ER, are shown in the lower left box. The categorical latent growth trajectory variable *c* is below the black line, representing the unobservable latent "class" of children, who are determined empirically to represent a coherent subgroup based on the pattern of variation in their growth trajectory.

Using CGM, unconditional models with only the growth coefficients (linear and quadratic), no predictors or covariates, and no latent classes were run first for Calculation, Problem-Solving, and Decoding across age. In these and all subsequent models, Age 13 was chosen as the centering point for greater measurement precision, as the majority (72%) of participants had completed assessments at this age. By setting the intercept at the age of 13, the intercept is the estimated subtest performance at age 13 and the variance estimate also reflects the value at Age 13. Because a quadratic model best fit the data,

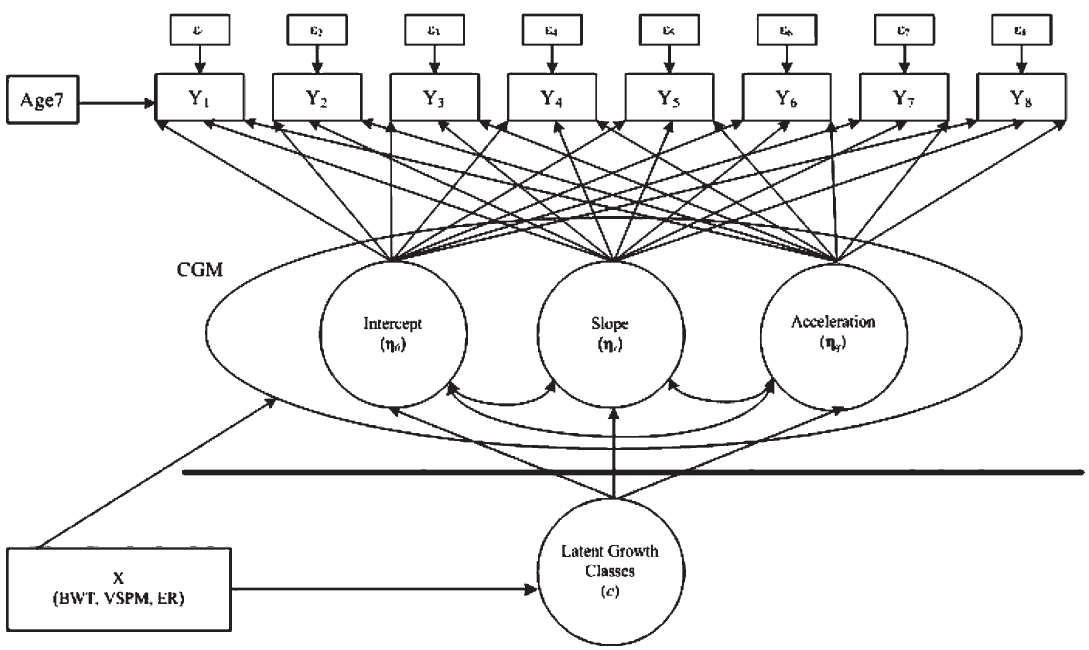


Figure 1. General diagram of growth mixture models for of birth weight risk. CGM = conventional growth modeling; BWT = birth weight; VSPM = visuospatial/perceptual-motor factor; ER = environmental risk; c = class.

the slope parameter is the estimated increase in achievement per unit of age at Age 13. Finally, the quadratic parameter is the estimated change in slope across the observation period, and identifies whether growth is "convex (troughed)" or "concave (peaked)." The positive sign of the quadratic parameter indicates accelerating growth and the negative sign indicates deceleration. For achievement outcomes, a "concave" pattern with a peak is expected as children's skills grow toward a maximum value. Using calculus, the age at which the quadratic function reaches its peak can be calculated by [13 – [1/2] (α_s/α_q)], where α_s and α_q represent the estimates for slope and quadratic parameters, although as with any polynomial, the timing of the peak may not fall within the range of the data. In general, these CGM models examine the latent growth factors and growth trajectory shape drawn from a single population, where individual trajectories were allowed to vary around a single population mean trajectory.

In GMM, the heterogeneity of population growth trajectories is captured by the latent categorical variable c with K classes, where the continuous latent growth variables for individuals in the kth class are related to $\boldsymbol{c}$ and the observed predictors $\boldsymbol{x}$. Here, the mixture models contain three growth factor means in the kth class (intercept, slope, acceleration), and the three fixed-effect coefficients (BWT, VSPM, and ER) of x_i on the three latent growth factors. A multinomial logistic regression model is applied in GMM to describe the relation between predictors and latent trajectory classes, where the probability of being in the kth class for child i is conditional on the predictors.

Model Estimation and Assessment

In this study, all models were estimated by maximum likelihood using the expectation maximization algorithm (Dempster, Laird & Rubin, 1977; Muthén & Muthén, 2001, 2006; B. Muthén & Shedden, 1999). This method is appropriate for analysis of data that are collected under conditions of "planned missingness" and are considered missing at random (Graham, Taylor, & Cumsille, 2001; Little & Rubin, 2002; Schafer & Graham, 2002), as was the case for this study. The maximum likelihood estimator with robust standard errors and chi-square likelihood ratio test allows missing data that is consistent with missing at random assumption, as well as nonnormal and nonindependence outcomes (Yuan & Bentler, 2000). Local maxima are encountered often in GMM, especially with an increasing number of latent classes (Muthén, 2004; Muthén & Muthén, 2006). For K ≥ 2, this study used 100 to 10,000 random sets of starting values at the initial stage and 5 to 20 optimizations at the final stage to avoid local maxima.

To determine the CGM trajectory shape, the chi-square test, based on maximum log-likelihood ratio (MLR) and scaling correction factors (Satorra, 2000), was used to compare relative fit among models that included linear and quadratic growth functions, respectively. The selected latent class model then was adopted in subsequent models examining the effects of predictors of class membership. The a priori models were tested by holding parameters invariant across classes, fixing or freeing parameters within and across classes, and then adding predictors. The chi-square likelihood ratio test is not appropriate for comparing models with different numbers of classes. An integrated approach was adopted here, where the number of latent classes was determined by the overall evaluation of the four criteria: (a) Bayesian information criteria (BIC; the smaller information criterion indicates better fit), (b) entropy, (c) bootstrap likelihood ratio test (BLRT); and (d) graphs of estimated class mean trajectories with and without covariates. BIC (Schwartz, 1978) was selected because it best identifies the correct number of classes in GMM using the common information criteria (Nylund, Asparouhov & Muthén, 2006). Entropy (E_k) was used to measure the classification quality based on participant's posterior class membership probabilities (Nagin, 1999; Ramaswamy, Desarbo, Reibstein, & Robinson, 1993), where entropy values closer to 1 indicate clear classification. BLRT uses bootstrap samples to estimate the distribution of the log-likelihood difference test statistic and was regarded as a powerful indication of the correct number of clusters (Nylund et al., 2006). A low p value of BLRT (e.g., $p < .05$) indicates the rejection of $k - 1$ classes in favor of k classes. Because the prediction of class membership is a key feature of GMM that permits testing empirically derived hypotheses, B. Muthén (2004) recommended that predictors be included in models to help determine the number of classes. Therefore, graphs of class mean trajectories were plotted with and without predictors to examine whether the latent classes were substantive. For example, the kth class mean trajectory may be so close to the (k −1)th class as to render the distinction between the two latent classes a trivial one.

Results

The longitudinal achievement data was analyzed to address our objectives through the following steps: (a) establish the trajectory shape, (b) identify the latent classes, and (c) examine the prediction of class membership from the predictors of particular interest. To establish the trajectory shape, unconditional linear and quadratic functions were fit to the data using conventional growth models, respectively. The chi-square difference test based on MLR and scaling factors indicated a quadratic growth curve model best fit the data for Calculation, $\chi^2(4) = 55.54$, $p < .001$; Problem-Solving, $\chi^2(4) = 322.30$, $p < .001$; and Decoding, $\chi^2(4) = 337.46$, $p < .001$. The quadratic term was significant and negative in sign, indicating that with advancing age, the rate of linear growth was progressively smaller for the three achievement scores: Calculation $\alpha_q = -1.18$, $p < .001$; Problem-Solving $\alpha_q = -0.76$, $p < .001$; Decoding $\alpha_q = -1.28$, $p < .001$ (see Table 3). Figure 2 displays these quadratic growth curves for the entire sample, and Table 3 contains the growth statistics.

For GMM, the unconditional mixture models were fit to the achievement data by assuming that the three continuous latent growth coefficients were invariant within classes (L. Muthén & Muthén, 2007; Nagin, 1999). Because the variances of intercepts and residuals differed across classes based on the chi-square likelihood ratio tests for the model of the same classes, these terms were allowed subsequently to vary within classes. As initial exploratory modeling showed that the GMM with more than three classes ($k > 3$) contributed trivially to class identification, the models with $k > 3$ were not investigated further and only models with two- and three-classes were studied in comparison to the one-class conventional growth model.

Class Identification

To compare the models with different numbers of classes, the integrated criteria were applied initially for each of the

achievement models without the inclusion of any predictors, shown in Table 3. For the comparison of one-class CGM and two-class models, the smaller BIC and significant *p* value of the BLRT test indicated that the two-class mixture model better fit the data (see Table 3). The contrast of two- and three-class unconditional models showed some support for selection of the three-class model for Calculation. The BIC was smaller for the three-class model than for the two-class model (BIC = 6,997.87 vs.7,018.31). The LMR likelihood ratio test for the three-class model also indicated better fit, and the entropy value showed just slightly better classification quality (E_k = .94 vs. E_k = .93). However, examining the plots of estimated mean trajectories raised concern about the utility of the three-class model for Calculation, as the two higher mean class trajectories in the three-class model overlapped nearly entirely, indicating that there were not substantive differences in the Calculation trajectories between classes. The two-class model was thus adopted for Calculation based on the overall evaluation of the application of the four criteria. For the Problem-Solving models, all criteria indicated the two-class solution was preferred. For Decoding, entropy and BLRT suggested a three-class solution, whereas BIC and the plots of the estimated mean trajectories pointed to a two-class model as best fitting.

To confirm the selection of the two-class model, the predictors then were included in the respective achievement

Table 3. Growth Parameter Estimates of Three Conventional Growth Models

Calculation

CGM	Intercept α_0 (*SE*)	Slope α_s (*SE*)	Deceleration α_q (*SE*)	MLR log-likelihood
Linear[a]	507.12*** (2.18)	6.88*** (0.56)	N/A	–3,703.44 (1.57)
Quadratic[a]	511.09*** (2.04)	6.76*** (0.21)	–1.18*** (0.05)	–3,490.39 (1.37)

	Without covariates			With covariates		
GMM	BIC	Entropy	BLRT *k* – 1 versus *k* classes	BIC	Entropy	BLRT *k* – 1 versus *k* classes
One class	7,089.39	N/A	N/A	6,352.53	N/A	N/A
Two classes	7,018.31	0.93	0.00___	6,291.26	0.85	0.00***
Three classes	6,997.87	0.94	0.00___	6,350.53	0.76	0.43

Problem-solving

CGM	Intercept α_0 (*SE*)	Slope α_s (*SE*)	Deceleration α_q (*SE*)	MLR log-likelihood
Linear[a]	504.33*** (1.75)	8.53*** (0.23)	N/A	–3,080.35 (1.150)
Quadratic[a]	505.20*** (1.61)	4.93*** (0.35)	–0.75*** (0.06)	–2,994.73 (1.24)

	Without covariates			With covariates		
GMM	BIC	Entropy	BLRT *k* – 1 versus *k* classes	BIC	Entropy	BLRT *k* – 1 versus *k* classes
One class	6,084.47	N/A	N/A	6,062.02	N/A	N/A
Two classes	6,041.07	0.93	0.00***	6,057.63	0.93	0.00***
Three classes	6,044.80	0.89	0.05	6,074.81	0.71	0.08

Decoding

CGM	Intercept α_0 (*SE*)	Slope α_s (*SE*)	Deceleration α_q (*SE*)	MLR log-likelihood
Linear	505.26*** (2.36)	6.46*** (0.36)	N/A	–3,781.85 (2.79)
Quadratic[a]	509.65*** (2.20)	6.86*** (0.20)	–1.28*** (0.05)	–3,593.72 (2.30)

	Without covariates			With covariates		
GMM	BIC	Entropy	BLRT *k* – 1 versus *k* classes	BIC	Entropy	BLRT *k* – 1 versus *k* classes
One class	7,287.60	N/A	N/A	6,533.48	N/A	N/A
Two classes	7,190.26	0.75	0.00***	6,486.40	0.74	0.00***
Three classes	7,199.52	0.94	0.00***	6,534.88	0.74	0.07

CGM = conventional growth modeling; *SE* = standard error; MLR = maximum log-likelihood ratio; GMM = growth mixture modeling; BLRT = bootstrap likelihood ratio test; BIC = Bayesian Information Criteria.

[a] Linear and quadratic function with the covariate Age7 in the model.

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

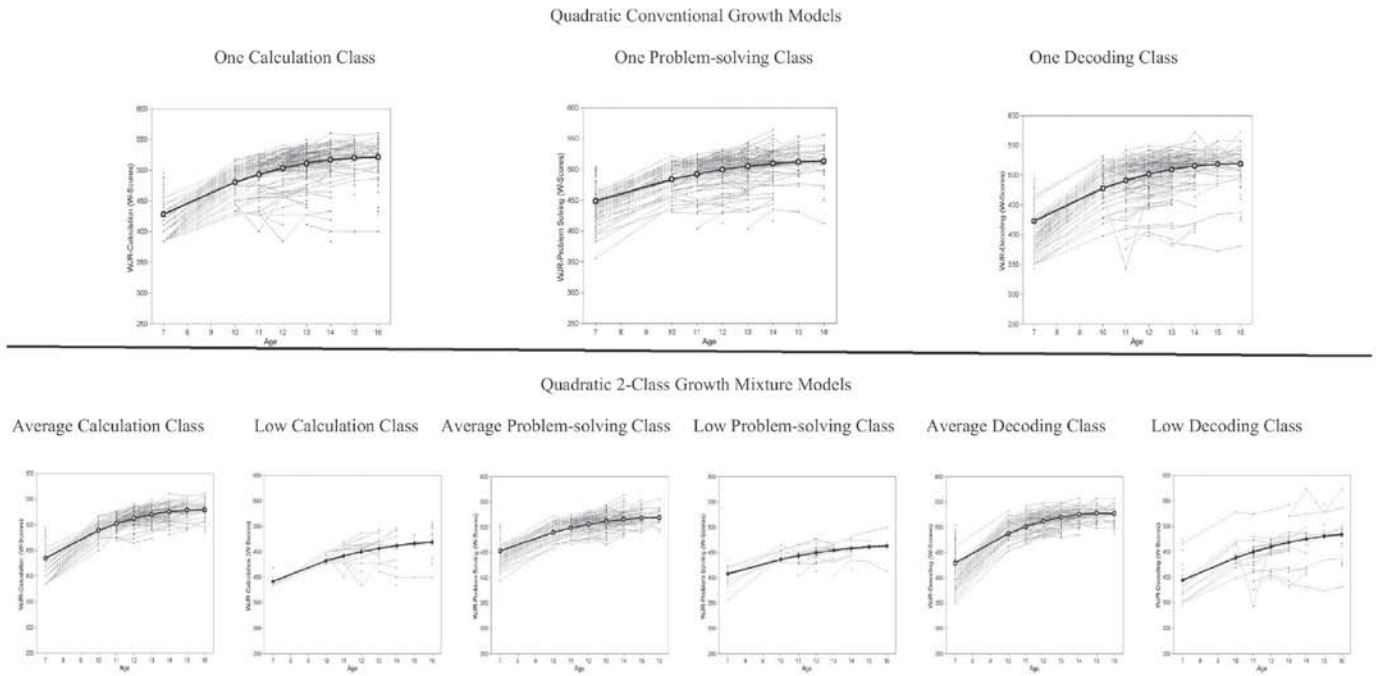


Figure 2. Individual curves are plotted around a single mean population trajectory and around two subpopulation trajectories for quadratic conventional growth model and two-class growth mixture models, respectively, for academic achievement outcomes.

GMMs. With the inclusion of BWT, VSPM, and ER, all fit criteria indicated that the two-class solution better described the observed Calculation, Problem-Solving, and Decoding achievement data, also depicted in Table 3. The BIC was smaller for the conditional two-class model (BIC = 6,291.26) than for the three-class model (BIC = 6,350.53). The BLRT also favored the two-class model, and the entropy value of the conditional two-class model (E_k = 0.85) was substantially larger than that for the conditional three-class model (E_k = 0.76). Table 4 shows the average probabilities of being in a class given a two-class solution. In support of this model, values on the diagonal were close to 1 and those on the off-diagonal were low and close to 0 (L. Muthén & Muthén, 2007). As expected based on the unconditional results, the conditional two-class Calculation model fit better than the conditional one-class CGM. A final consideration in selecting the two-class model for Calculation was that the plots for the two higher achieving classes were closely adjacent to each other in the three-class conditional model. Because these same criteria indicated a better fit of the two-class model for Problem-Solving and Decoding, the two-class model also was retained for these achievement outcomes.

Figure 3 displays the expected mean trajectory for Calculation, Problem-Solving, and Decoding for the two-class model. The mean trajectories of the two identified latent classes (termed *average* and *low*) are shown separately. The growth parameter estimates for these two classes for Calculation, Problem-Solving, and Decoding are shown in Table 5. For Calculation, the average class scored more than 60 W score points higher (intercept α_0) than the low class at Age 13 years and their linear change rate (slope $\boldsymbol{\alpha}_s$) also was faster at this age. Acceleration ($\boldsymbol{\alpha}_q$) for both groups was negative in sign, meaning that the rate of linear change was progressively slowing across age, with a larger magnitude of deceleration for the average class than for the low class. The estimated peak of the developmental trajectory for the average class was at age of 15.6 years compared with an estimated peak at 17.4 years for the low class, although the gap in achievement between low and average classes persisted across age, as evident in Figure 2. Furthermore, there was significant variation in the Calculation intercepts at Age 13 for both the average and low classes, as well as significant variation in slopes at Age 13 for the average class. Because variation in acceleration was nonsignificant for both classes, the quadratic parameters were fixed to 0. Participants in the average class performed better on Calculation than the low class at enrollment, and this performance difference between classes was evident at all ages. The average class, though, showed faster skill growth early in the developmental period and greater deceleration across age, which resulted in obtaining the maximal level of Calculation achievement at a younger age relative to those in the low class.

Table 4. Averaged Latent Class Probabilities by Classes

	Average class	Low class
Calculation		
Average class	0.94	0.06
Low class	0.04	0.96
Problem-solving		
Average class	0.99	0.01
Low class	0.08	0.92
Decoding		
Average class	0.92	0.08
Low class	0.06	0.94

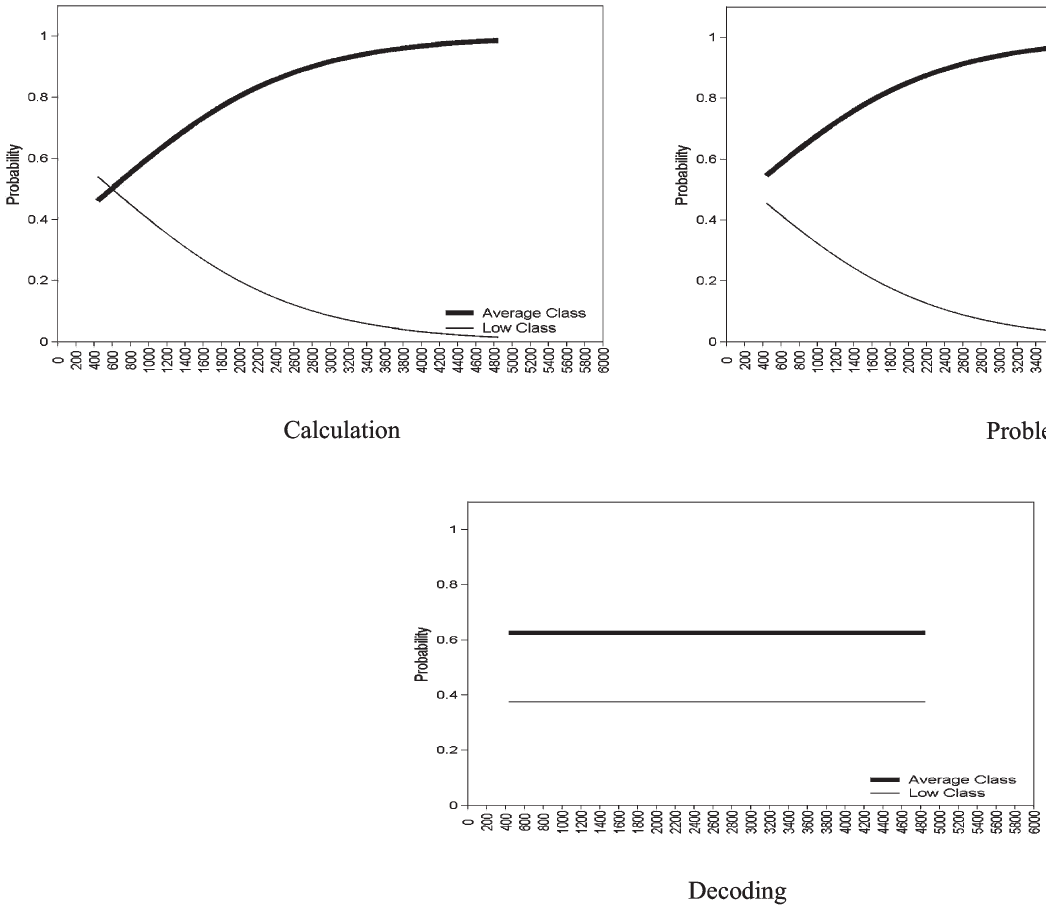


Figure 3. Probability for class membership against weight by latent classes.

The plots and growth parameter estimates for Problem-Solving was similar to those for Calculation, with a higher intercept and faster linear rate of change at Age 13 for the average class compared to the low class. Because growth deceleration also was faster for the average class, the estimated maximal Problem-Solving score was at Age 16.1 year for the average class and at 17.0 years for the low class. The pattern of individual differences in variances of the growth parameters for the two latent classes was identical to that for Calculation. Overall, for Problem-Solving, participants in the average class performed better than those in the low class across age. Similar to Calculation, children identified in the average class showed faster skill growth early in the developmental period, although the magnitude of the difference in early growth between average and low classes was smaller than for Calculation (estimated Calculation growth was 22.11 for the average class and 14.37 for the low class at Age 7, for Problem-Solving the respective values for these classes were 14.55 and 11.02). With greater growth deceleration, the average class obtained the maximal Problem-Solving achievement at a younger age compared to the low class.

For Decoding, the two latent classes differed in all three growth parameters. The average class was estimated to score more than 50 W points higher than the low class at Age 13 years. In contrast to the two mathematics scores, the linear rate of change for Decoding at Age 13 was higher for the low relative to the average class. Similar to the two mathematics scores, both latent classes showed deceleration in growth, with a faster rate of deceleration for the average class. The average class was estimated to reach maximal Decoding score at 15.4 years, whereas the low class was estimated to reach its peak at 17.8 years, again with a persistent gap in achievement scores noted across the observation period. The variance of the intercept for the low class was relatively larger than that for the average class, whereas the variance for the slope and quadratic parameters did not differ from 0 and thus was fixed. Similar to the two mathematics outcomes, participants in the average class performed better on Decoding than those in the low class across age, showed faster skill growth early in the developmental period, and greater growth deceleration, again resulting in a higher maximal level of Decoding achievement at a younger age compared to the low class. The early, age-related difference between the low and average classes in linear growth in Decoding were similar in magnitude to Calculation (at Age 7, estimated growth in Decoding was 23.55 for the average class and 17.24 for the low class).

Prediction of Class Membership

To characterize the relation between class membership and the predictors of interest, the categorical latent class variable, *c*, was regressed on BWT, VSPM, and ER simultaneously, which revealed the effects of each predictor controlling for the influences of the others. For Calculation and Problem-Solving, BWT was the only significant predictor of class membership

($\gamma_{BWT_calculation}$ = 0.001, p = .014; $\gamma_{BWT_problem\text{-}solving}$ = 0.001, p = .031). For Decoding, the effect of BWT on class membership was marginal ($\gamma_{BWT_decoding}$ = 0.001, p = .086). As BWT is continuously distributed, these coefficients indicate that the log odds of being in the average relative to the low class increases by .001 for each unit increase in BWT. The probability plots of BWT by the latent classes for each achievement outcome are displayed in Figure 3. For Calculation, the probability of being classified into the average achievement group increased from 0.46 to 0.54 as BWT increased from 439 to 750 g; from 0.54 to 0.71 as BWT increased from 750 to 1,499 g; and from 0.71 to 0.99 as BWT increased over 1,500 g. The probability plot indicates that children born at less than 600 g had a greater probability of being assigned to the low class relative to the average for Calculation (i.e., where the low and average class probability plots cross, as the probability of low class assignment > 0.50). For Problem-Solving, the probability of being classified into the average group increased from 0.54 to 0.62 as BWT increased from 439 to 750 g; from 0.62 to 0.77 as BWT increased from 750 to 1,499 g; and from 0.77 to 0.99 as BWT increased beyond 1,500 g. In contrast to the pattern evident for Calculation, children across the BWT spectrum were more likely to be identified as average achieving for Problem-Solving. For Decoding, the coefficient was not significant; therefore, the probability of identification in the average class relative to the low was a constant (0.63) for children irrespective of BWT. Neither VSPM nor ER predicted class membership beyond BWT for Calculation ($\gamma_{vspm_calculation}$ = 0.279, p = .258; $\gamma_{ER_calculation}$ = 0.05, p = .325), Problem-Solving ($\gamma_{vspm_problem\text{-}solving}$ = 0.279, p = .258; $\gamma_{ES_problem\text{-}solving}$ = 0.05, p = .325), or Decoding ($\gamma_{vspm_decoding}$ = 0.236, p = .492; $\gamma_{ES_decoding}$ = 0.067, p = .195).

The final step in analysis was to examine the impact of the neonatal medical variables on group membership after controlling for BWT. Because these conditions were coded only for children with VLBW, the analyses were conducted without inclusion of term children, although the class membership derived from the full sample GMM analyses was utilized. Days on ventilation significantly predicted class membership for all three achievement outcomes ($\gamma_{vent_calculation}$ = 0.015, p = .048, $\gamma_{vent_problem\text{-}solving}$ = 0.015, p = .034, and $\gamma_{vent_decoding}$ = 0.017, p = .028). Children with VLBW who required more ventilation days were more likely to be classified in the low class relative to the average class on Calculation, Problem-Solving, and Decoding, as the log odds increases by 0.015 to 0.017 for a unit increase in day on ventilation. The probability plots of the significant neonatal variables by latent classes for the three achievement outcomes are shown in Figure 4. The plot for Calculation demonstrates that children receiving ventilation for 34 days or more (again where the low and average class probabilities cross and the probability of low class assignment > 0.50), were more likely to be classified into the low relative to average class. Similarly, children who received 43 or more days of ventilation were more likely to be in the low relative to average class on Problem-Solving; and those who received 70 or more days of ventilation were more likely to be in the low compared to average class on Decoding.

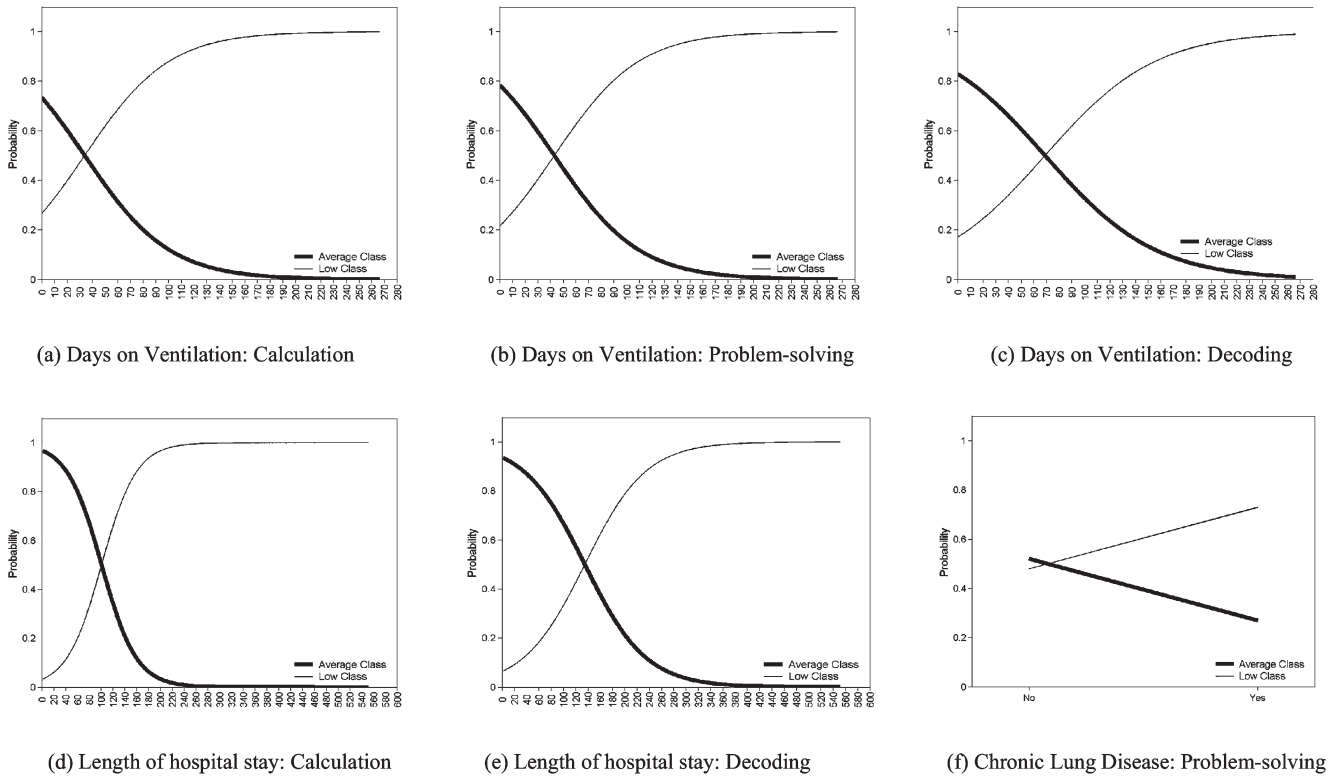
Controlling for BWT, length of hospitalization predicted class membership for the Calculation and Decoding variables, ($\gamma_{hosp_calculation}$ = 0.012, p = .040, $\gamma_{hosp_decoding}$ = 0.013, p = .023). Specifically, the respective odds ratio of being classified in the low versus average class for Calculation and Decoding was 1.012 and 1.013 for a unit (i.e., 1 day) increase in hospitalization. For Calculation, children who were hospitalized for 101 days or more were more likely to be classified in the low class than in the average class. For Decoding, the probability of a child being classified in the low compared to aver-

Table 5. *Growth Parameter Estimates*

Parameter	Average class (*SE*)	Low class (*SE*)
Calculation		
α_0	520.15 (1.40)***	456.58 (7.86)***
α_s	6.63 (0.23)***	6.09 (0.79)***
α_q	–1.29 (0.05)***	–0.69 (0.18)***
$V(\zeta_0)$	224.21 (45.53)***	224.20 (45.52)***
$V(\zeta_s)$	1.54 (0.66)*	4.44 (2.36)
$V(\zeta_q)$	—	—
Problem-solving		
α_0	511.48 (1.38) ***	454.43 (7.98) ***
α_s	4.95 (0.37) ***	4.42 (1.04)
α_q	–0.80 (0.07) ***	–0.55 (0.24) ***
$V(\zeta_0)$	183.97 (26.50) ***	183.98 (26.47) ***
$V(\zeta_s)$	6.03 (2.82)*	6.02 (2.83)*
$V(\zeta_q)$	0.25 (0.11)*	0.30 (0.10)*
Decoding		
α_0	519.96 (1.60) ***	468.19 (8.18) ***
α_s	6.63 (0.21) ***	7.64 (0.59) ***
α_q	–1.41 (0.05) ***	–0.80 (0.10) ***
$V(\zeta_0)$	192.16 (33.92) ***	1,444.82 (361.92) ***
$V(\zeta_s)$	—	—
$V(\zeta_q)$	—	—

An em dash (—) is fixed to 0, α_0, α_s, and α_q are the mean estimates, and $V(\zeta_0)$, $V(\zeta_s)$ and $V(\zeta_q)$ are the variance estimates of growth intercepts, slopes, and deceleration rates, respectively.

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



(a) Days on Ventilation: Calculation (b) Days on Ventilation: Problem-solving (c) Days on Ventilation: Decoding

(d) Length of hospital stay: Calculation (e) Length of hospital stay: Decoding (f) Chronic Lung Disease: Problem-solving

Figure 4. Probability for class membership against significant neonatal medical variables by latent classes with birth weight controlled.

age class exceeded .50 when the length of hospitalization was 136 days or more.

Again controlling for BWT, chronic lung disease was associated with class membership for Problem-Solving ($\gamma_{lung_problem\text{-}solving} = 1.072$, $p = .032$). Children with VLBW with chronic lung disease had a probability of being assigned to the low compared to the average class for Problem-Solving of .58, whereas children with VLBW without chronic lung disease had a probability of being assigned to the average compared to low class of .80.

Discussion

GMM, an integrated person- and variable-centered approach, was applied to longitudinal academic achievement data from a large cohort of children born early and at VLBW to better understand the heterogeneity of outcome across development. Two latent classes were identified empirically, defining subgroups of average- and low-achieving participants whose pattern of longitudinal growth was similar enough to successfully result in cohesive classification. These results demonstrate how GMM extends CGM approaches by incorporating the concept of latent classes into growth variation, therefore explaining more variance of the individual growth trajectories. The findings confirmed our hypotheses that the degree of VLBW is related to suboptimal growth in Calculation, Problem-Solving, and Decoding skills across childhood and into adolescence.

The GMM models revealed important information regarding the nature of academic achievement skill growth in children of varying BWT. With two latent classes, a coherent, low-performing group was identified empirically that differed from average-achieving children, calculation, problem-solving, and decoding skills. The consistency of the two class solution across differing achievement skills, as well as when term-born children are included in the models or not (results available from the author), suggests a coherent person-level group structure from school age into adolescence. A key feature of GMM is the assumption that individual differences also exist within each of the empirically identified groups. For example, some of the lightest children with VLBW performed as well as those born at term at certain ages, and concomitantly, some term-born children scored as poorly as the VLBW children across ages. Less than half of the lightest VLBW children (<750 g) were identified as members of the low class for Calculation (28 of 60, 47%) and for Problem-Solving (24 of 60, 40%), and about one third (21 of 60, 35%) were classified in the low class for Decoding (see Table 6). These findings confirm the degree of heterogeneity in academic outcomes, as even among the lightest of children with VLBW, the majority does not show suboptimal patterns of academic achievement skill growth. The GMM approach that includes CGM and latent classes can be utilized to effectively capture the variation within and across groups of children in developmental periods of interest.

More important, the difference between the low- and average-achieving classes was apparent across the developmental period from Age 7 to 16 years for Calculation, Problem-solving, and Decoding, with the average class outscoring the low class by a substantial margin. However, the GMMs revealed

a more complicated developmental picture. Given the faster linear rates of change early in the observation period, as well as in the greater growth deceleration, children in the average-achieving class were estimated to obtain maximal achievement performance between 15 and 16 years (depending on outcome type), whereas low-achieving children were estimated to achieve their maximal scores at later ages, beyond age 17, although a persistent gap in achievement for the low class was evident across age. Furthermore, the early, age-related growth and the differences between the low and average classes in linear growth both were greatest for Decoding and Calculation compared to Problem-Solving, which is consistent with the early, rapid acquisition and application of sound-symbol relations and mathematic operations for typically developing children.

For low-achieving children, this developmental pattern suggests that early achievement skill acquisition is fundamentally disrupted early in life that persists and skill growth is further impaired across development. These findings extend the earlier achievement findings of Taylor, Klein, Hack (2000) and Taylor, Klein, Minich, et al. (2000) by revealing continued and even increasing deficits in achievement over time in children with more extreme VLBW compared with term-born controls. The results are also consistent with studies that have either found stable deficits across development in children with VLBW (Breslau et al., 2001; Powls et al., 1995; Rickards et al., 1988) or that have raised the possibility of slower age-related acquisition in some skills (Botting et al., 1998; Cohen et al., 1996; O'Callaghan et al., 1996; Zelkowitz et al., 1995). The observed pattern of deficits between low- and average-achieving classes identified here with GMM approach that includes meaningful subgroups further refines our knowledge of individual variation in academic skill developmental patterns. The different trajectory patterns suggests that there are distinct biologically determined upper limits of achievement skill acquisition, and that children identified in the low class can continue to benefit from learning inputs and academic instruction beyond the age at which academic learning typically levels off in average-achieving children to achieve maximal proficiency. Of course, a different developmental pattern might have been observed during the transition from preschool to school age, or into adulthood.

Consistent with previous research, poorer achievement outcomes were related to several biological risk factors (Taylor et al., 2006; Taylor, Minich, Bangert, et al., 2004). What is newly demonstrated here using the GMM approach is the impact of these risks at the person level. More than half of the respective low-achieving classes for Calculation (28 of 38; 74%), Problem-Solving (24 of 43; 56%), and Decoding (21 of 36; 58%) were the lightest VLBW children (<750 g). These findings were reinforced by GMM analyses that identified the individuals within the sample who were at highest risk for suboptimal achievement growth. For Calculation, children under 600 g were more likely than not to be classified as low achieving. For the other two academic skills, Problem-Solving and Decoding, the probability of identification in the average-achieving group relative to the low-achieving group was greater across the entire BWT spectrum. For these measures though, the specificity of more extreme VLBW was low, as less than half of the lightest VLBW children were identified as members of the low classes.

Contrary to expectations, nonverbal skills as assessed by the VSPM Factor score were not related to latent class membership for Calculation, Problem-Solving, or Decoding when controlling for BWT. The failure of the VSPM score to predict poorer outcomes is surprising given the well-documented associations of these skills with academic achievement (Grunau, Whitfield, & Davis, 2002; Taylor et al., 2002). A likely explanation is that the influence of nonverbal skills on academic achievement was mediated in large part by BWT. The GMM results enrich this interpretation by including person-level effects, suggesting that impact of both BWT and nonverbal skills overlaps substantially at the person level and therefore leaves little systematic variation to predict achievement growth patterns. Environmental risk also failed to contribute to group membership beyond the influence of BWT for any of the achievement outcomes, contrary to extant findings (Bendersky & Lewis, 1995; Breslau & Chilcoat, 2000; Taylor et al., 1998; 2006). Including the person-level suggests that the neurobiological consequencea of very preterm birth at weights at the extreme of the continuum are the more significant contributors to suboptimal patterns of academic achievement in school age and into adolescence (Taylor et al., 1998, 2006).

Table 6. Counts of Term, 750 to 1,499 g, and <750 g Children by Latent Growth Classes

	Birth weight group			
	Term	750 to 1,499 g	<750 g	Total
Calculation				
Average class	57	51	32	140
Low class	2	8	28	38
Total	59	59	60	178
Problem-solving				
Average class	53	46	36	135
Low class	6	13	24	43
Total	59	59	60	178
Decoding				
Average class	54	49	39	142
Low class	5	10	21	36
Total	59	59	60	178

Although BWT was the central predictor; it is only a proxy for the impact of prematurity and related neonatal medical complications on the developing brain. Among the children with VLBW, risk factors that marked a more complicated neonatal course, including more days on ventilation, a longer period of neonatal hospitalization, and chronic lung disease predicted class membership for one of more of the achievement outcomes. Although consistent with previous findings (D'Angio et al., 2002; Short et al., 2007; Taylor, 1998, 2006), these results extend this literature by indicating that these associations are independent of BWT. Calculation was the most sensitive to disruption among the two mathematics domains, as the probability of classification as low- compared to average-achieving for Calculation was higher than that for Problem-Solving at the fewer days of ventilation (34 days compared to 43 days with BWT controlled). Decoding skills were more "resilient," in which identification into the low relative to average class for this outcome exceeded .5 after 70 days of ventilation. Chronic lung disease and days on ventilation also predicted class membership for Problem-Solving, potentially pointing to the impacts of prolonged oxygen imbalances on mathematics applications. Although these neonatal medical variables are more specific indicators of neurobiologic risk than BWT, they do not directly or specifically quantify the impact on the central nervous system. Because complications vary across children and likely have variable impact on the brain, neuroimaging methods may provide one means for further distinguishing the level of individual risk. More refined methods of assessing brain status thus may thus improve our understanding of the sources of variability in achievement skill growth across development. Larger multisite samples also may be required in investigations of the impact on achievement outcome of low base rate neonatal complications.

Subgroups of average- and low-achieving children were identified empirically and were related to predictors of interest, demonstrating the flexibility and capacity of the GMM approach to understand heterogeneity of outcome. This two-group classification likely does not reflect what would be determined in a representative sample of all children between ages 7 and 16 years, as such a sample would include on average over 90% of children born term at weights >2,500 and very few children at weights <750 g. Indeed, the sampling used here was ideal to determine empirically whether differing classes of children with distinctive developmental trajectories could be identified to understand heterogeneity in academic outcomes among children at-risk due to VLBW. If a different sampling strategy was used, it is likely that other coherent subgroups would have emerged. GMM extends conventional models that are predicated on a variable-centered orientation, and then incorporates latent class growth analysis to identify the number of coherent classes of individuals using a person-centered approach. Other analytic approaches also can be accommodated by GMM. For example, using Markov or piecewise GMM analysis and assuming transition stages or critical time points, the class membership shifting of children of varying BWT groups across latent classes can be modeled to address whether some children are average-achieving early in life and then shift to a low achieving subgroup during a critical period. The flexibility of GMM makes it an ideal framework with which to better model the complexities and heterogeneities of preterm and other at-risk children as they develop across time.

References

Anderson, P., & Doyle, L. W. (2003). Neurobehavioral outcomes of school-age children born extremely low birth weight or very preterm in the 1990s. *Journal of the American Medical Association*, *289*, 3264–3272.

Anderson, P., Doyle, L. W., & Victorian Infant Collaborative Study Group. (2004). Executive functioning in school-aged children who were born very preterm or with extremely low birth weight in the 1990s. *Pediatrics*, *114*, 50–57.

Beery, K. E. (1989). *Revised administration, scoring, and teaching manual for the Developmental Test of Visual-Motor Integration*. Cleveland, OH: Modern Curriculum.

Bendersky, M., & Lewis, M. (1995). Effects of intraventricular hemorrhage and other medical and environmental risks on multiple outcomes at age three years. *Developmental and Behavior Pediatrics*, *16*, 89–96.

Bhutta, A. T., Cleves, M. A., Casey, P. H., Cradock, M. M., & Anand, K. J. S. (2002). Cognitive and behavioral outcomes of school-age children who were born preterm. *Journal of the American Medical Association*, *288*, 728–737.

Botting, N., Powls, A., Cooke, R., & Marlow, N. (1998). Cognitive and educational outcome of very low birth weight children in early adolescence. *Developmental Medicine & Child Neurology*, *40*, 652–660.

Botting, N., Powls, A., Cooke, R. W. I., Marlow, N. (1997). Attention deficit hyperactivity disorders and other psychiatric outcomes in very low birth weight children at 12 years. *Journal of Child Psychology and Psychiatry*, *8*, 931–941.

Breslau, N. (1995). Psychiatric sequelae of low birth weight. *Epidemiologic Reviews*, *17*, 96–106.

Breslau, N., & Chilcoat, H. D. (2000). Psychiatric sequelae of low birth weight at 11 years of age. *Biological Psychiatry*, *47*, 1005–1011.

Breslau, N., Chilcoat, H. D., Susser, E., Matte, T., Liang, K., & Peterson, E. (2001). Stability and change in children's intelligence quotient scores: A comparison of two socioeconomically disparate communities. *American Journal of Epidemiology*, *154*, 711–717.

Breslau, N., Paneth, N. S., & Lucia, V. C. (2004). The lingering academic deficits of low birth weight children. *Pediatrics*, *114*, 1035–1040.

Cohen, S., Beckwith, L., Parmelee, A., Sigman, M., Asarnow, R., & Espinosa, M. (1996). Prediction of low and normal achievement in early adolescents born preterm. *Journal of Early Adolescence*, *16*, 46–70.

D'Angio, C. T., Sinkin, R. A., Stevens, T. P., Landfish, N. K., Merzbach, J. L., Phelps, D. L., et al. (2002). Longitudinal, 15-year follow-up of children born at less than 29 weeks gestation after introduction of surfactant therapy into a region: Neurologic, cognitive, and educational outcomes. *Pediatrics*, *46*, 812–815.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithim. *Journal of the Royal Statistical Society*, *39*(B), 1–8.

Espy, K. A., McDiarmid, M. D., Cwik, M. F., Senn, T. E., Hamby, A., & Stalets, M. M. (2004). The contributions of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology*, *26*, 465–486.

Espy, K. A., Stalets, M. M., McDiarmid, M. D., Senn, T. E., Cwik, M. F., & Hamby, A. (2002). Executive functions in preschool children born preterm: Application of cognitive neuroscience paradigms. *Child Neuropsychology*, *8*, 83–92.

Ewing-Cobbs, L., Barnes, M., Fletcher, J. M., Levin, H. S., Swank, P. R., & Song, J. (2004). Modeling of longitudinal academic achievement scores after pediatric traumatic brain injury. *Developmental Neuropsychology*, *25*, 107–133.

Frisk, V., & Whyte, H. (1994). The long-term consequences of periventricular brain damage on language and verbal memory. *Developmental Neuropsychology*, *10*, 313–333.

Gardner, R. (1979). *The objective diagnosis of minimal brain dysfunction*. Cresskill, NJ: Creative Therapies.

Goyen, T., Lui, K., & Woods, R. (1998). Visual-motor, visual-perceptual, and fine motor outcomes in very-low-birth weight children at 5 years. *Developmental Medicine and Child Neurology*, *40*, 76–81.

Graham, J., Taylor, B., & Cumsille, P. (2001). Planned missing-data designs in analysis of change. In L.Collins & A.Sayer (Eds.), *New methods for the analysis of change* (pp. 335–353). Washington, DC: American Psychological Association.

Grunau, R. E., Whitfield, M. F., & Davis, C. (2002). Pattern of learning disabilities in children with extremely low birth weight and broadly average intelligence. *Archives of Pediatric and Adolescent Medicine*, *156*, 615–620.

Hack, M. (2005). Young adult outcomes of very-low-birth-weight children. *Seminars in Fetal & Neonatal Medicine*, *11*, 127–137.

Hack, M., Breslau, N., Aram, D., Weissman, B., Klein, N., & Borawski-Clark, E. (1992). The effect of very low birth weight and social risk on neurocognitive abilities at school age. *Journal of Developmental and Behavioral Pediatrics*, *13*, 412–420.

Hack, M., & Fanaroff, A. A. (1999). Outcomes of children of extremely low birth weight and gestational age in the 1990s. *Early Human Development*, *53*, 193–218.

Hack, M., Klein, N., & Taylor, H. G. (1996). School-age outcomes of children of extremely low birth weight and gestational age. *Seminars in Neonatology*, *1*, 227–288.

Hack, M., Wilson-Costello, D., Friedman, H., Taylor, H. G., Schluchter, M., & Fanaroff, A. A. (2000). Neurodevelopment and predictors of outcomes of children with birth weights of less than 1000g. *Archives of Pediatric and Adolescent Medicine*, *154*, 725–731.

Hedeker, D., & Gibbons, R. D. (1994). A random-effects ordinal regression model for multilevel analysis. *Biometrics*, *50*, 933–944.

Hille, E. T. M., den Ouden, A. L., Saigal, S., Wolke, D., Lambert, M., Whitaker, A., et al. (2001). Behavioural problems in children who weigh 1000 g or less at birth in four countries. *Lancet*, *357*, 1641–1643.

Hollingshead, A. (1957). *Two-factor index of social position*. New Haven, CT: Yale University.

Kaufman, A., & Applegate, B. (1988). Short forms of K–ABC Mental Processing and Achievement Scales at age 4 to 12–1/2 years for clinical and screening purposes. *Journal of Clinical Child Psychology*, *17*, 359–369.

Klebanov, P., Brooks-Gunn, J., & McCormick, M. (1994). School achievement and failure in very low birth weight children. *Developmental and Behavioral Pediatrics*, *15*, 248–256.

Klein, N., Hack, M., & Breslau, N. (1989). Children who were very low birth weight: Developmental and academic achievement at 9 years of age. *Journal of Developmental & Behavioral Pediatrics*, *10*, 32–37.

Koller, H., Lawson, K., Rose, S., Wallace, I., & McCarton, C. (1997). Patterns of cognitive development in very low birth weight children during the first six years of life. *Pediatrics*, *99*, 383–389.

Landry, S. H., Fletcher, J. M., Denson, S. E., & Chapieski, M. L. (1993). Longitudinal outcome for low birth weight infants: Effects of intraventricular hemorrhage and bronchopulmonary dysplasia. *Journal of Clinical and Experimental Neuropsychology*, *15*, 205–218.

Landry, S. H., Smith, K. E., Miller-Loncar, C. L., & Swank, P. R. (1997). Predicting cognitive-language and social growth curves from early maternal behaviors in children at varying degrees of biological risk. *Developmental Psychology*, *33*, 1040–1053.

Liaw, F., & Brooks-Gunn, J. (1993). Patterns of low-birth-weight children's cognitive development. *Developmental Psychology*, *29*, 1024–1035.

Litt, J., Taylor, H. G., Klein, N., & Hack, M. (2005). Learning disabilities in children with very low birth weight: Prevalence, neuropsychological correlates, and educational interventions. *Journal of Learning Disabilities*, 847–856.

Little, R. J., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). New York: Wiley.

Luoma, L., Herrgard, E., Martikainen, A., & Ahonen, T. (1998). Speech and language development of children born at <32 weeks gestation: A 5-year prospective follow-up study. *Developmental Medicine and Child Neurology*, *40*, 380–387.

McCulloch, C. E., & Searle, S. R. (2001). *Generalized, linear, and mixed models*. New York: John Wiley: Sons, Inc.

McGrath, M., & Sullivan, M. (2002). Birth weight, neonatal morbidities, and school age outcomes in full-term and preterm infants. *Issues in Comprehensive Pediatric Nursing*, *25*, 231–254.

Mehta, P. D., & West, S. G. (2000). Putting the individual back into individual growth curves. *Psychological Methods*, *5*, 23–43.

Ment, L., Vohr, B., Allan, W., Katz, K., Schneider, K., Westerweld, M., et al. (2003). Change in cognitive function over time in very low-birth-weight infants. *Journal of the American Medical Association*, *289*, 705–711.

Moos, R., & Moos, B. (1994). *Family Environment Scale manual: Development, applications and research* (3rd ed.). Palo Alto, CA: Consulting Psychologists.

Msall, M. E., Buck, G. M., Rogers, B. J., & Catanzaro, N. L. (1992). Kindergarten readiness after extreme prematurity. *American Journal of Diseases of Children*, *146*, 1371–1378.

Murphy-Berman, V., & Wright, G. (1987). Measures of attention. *Perceptual & Motor Skills*, *64*, 1139–1143.

Muthén, B. (2001a). Latent variable mixture modeling. In G. A.Marcoulides & R. E.Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 1–33). Mahwah, NJ: Erlbaum.

Muthén, B. (2001b). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In L.Collins & A.Sayer (Eds.), *New methods for the analysis of change* (pp. 291–322). Washington, DC: American Psychological Association.

Muthén, B. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, *29*, 81–117.

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

Muthén, B. (2006). *The potential of growth mixture modelling*. Retrieved November 5, 2006, from http://www.statmodel.com/download/482.pdf

Muthén, B., Khoo, S-T., Francis, D. J., & Boscardin, C. K. (2003). Analysis of reading skills development from kindergarten through first grade: An application of growth mixture modeling to sequential processes. In S. P.Reise, & N.Duan (Eds.), *Multilevel modeling: Methodological advances, issues, and applications* (pp. 71–89). Mahwah, NJ: Erlbaum.

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

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

Muthén, L., & Muthén, B. (2007). *Mplus user's guide* (5th ed.). Los Angeles: Author.

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

Nagin, D. S., & Tremblay, R. E. (2001). Analyzing developmental trajectories of distinct but related behaviors: A group-based method. *Psychological Methods*, *6*, 18–34.

Nagin, D. S., & Tremblay, R. E. (2005). Developmental trajectory groups: Fact or a useful statistical fiction?*Criminology*, *43*, 873–904.

Nesselroade, J. R., & Baltes, P. B. (1979). *Longitudinal research in the study of behavior and development*. San Diego, CA: Academic.

Nylund, K. L., Asparouhov, T., & Muthén, B. (2006). 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.

Olsen, M. K., & Schafer, J. L. (2001). A two-part random-effects model for semicontinuous longitudinal data. *Journal of the American Statistical Association*, *96*, 730–745.

O'Callaghan, M., Burns, Y., Gray, P., Harvey, J., Mohay, H., Rogers, Y., et al. (1996). School performance of ELBW children: A controlled study. *Developmental Medicine & Child Neurology*, *38*, 917–926.

Powls, A., Botting, N., Cooke, R., & Marlow, N. (1995). Motor impairment in children 12 to 13 years old with birth weight of less than 1250 g. *Archives of Diseases in Childhood*, *72*, F62–F66.

Ramaswamy, V., Desarbo, W., Reibstein, D., & Robinson, W. (1993). An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Marketing Science*, *12*, 103–124.

Raudenbush, S. W. (2001). Toward a coherent framework for comparing trajectories of individual change. In L. M.Collins & A.Sayer (Eds.), *New methods for the analysis of change* (pp. 35–64). Washington, DC: American Psychological Association.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. London: Sage.

Rickards, A., Ryan, M., & Kitchen, W. (1988). Longitudinal study of very low-birth weight infants: Intelligence and aspects of school progress at 14 years of age. *Developmental and Behavioral Pediatrics*, *22*, 19–23.

Semel, E., Wiig, W., & Sabers, D. (1987). *CELF-R: Clinical Evaluation of Language Functions-Revised (technical manual)*. New York: Psychological Corporation.

Saigal, S. (2000). Follow-up of very low birth weight babies to adolescence. *Seminars in Neonatology*, *5*, 107–118.

Saigal, S., Hoult, L., Streiner, D., Stoskopf, B., & Rosenbaum, P. (2000). School difficulties at adolescence in a regional cohort of children who were extremely low birth weight. *Pediatrics*, *105*, 325–331.

Saigal, S., Pinelli, J., Hoult, L., Kim, M. M., & Boyle, M. (2003). Psychopathology and social competencies of adolescents who were extremely low birth weight. *Pediatrics*, *111*, 969–975.

Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In R. D. H.Heijmans, D. S. G.Pollock, & A.Satorra (Eds.), *Innovations in multivariate statistical analysis* (pp. 233–247). London: Kluwer Academic.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*, 147–177.

Schneider, W., Wolke, Deiter, D., Schlagmuller, M., & Meyer, R. (2004). Pathways to school achievement in very preterm and full term children. *European Journal of Psychology of Education*, *19*, 385–406.

Schwarz, G. (1978). Estimating the dimensions of a model. *Annals of Statistics*, *6*, 461–464.

Semel, E.Wiig, W., & Sabers, D. (1987). *CELF–R: Clinical Evaluation of Language Functions–Revised (technical manual)*. New York: Psychological Corporation.

Short, E. J., Kirchner, H. L., Asaad, G. R., Fulton, S. E., Lewis, B. A., Klein, N., et al. (2007). Developmental sequelae in preterm infants having a diagnosis of brochopulmonary dysplasia. *Archives of Pediatric and Adolescent Medicine*, *161*, 1082–1087.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis*. New York: Oxford University Press.

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*. New York: Chapman & Hall/CRC.

Sullivan, M. C., & McGrath, M. M. (2003). Perinatal morbidity, mild motor delay, and later school outcomes. *Developmental Medicine & Child Neurology*, *45*, 104–112.

Szatmari, P., Saigal, S., Rosenbaum, P., Campbell, D., & King, S. (1990). Psychiatric disorders at 5 years among children with birth weights <1000 grams: A regional perspective. *Developmental Medicine in Child Neurology*, *32*, 954–962.

Taylor, H. G., Burant, C., Holding, P. A., Klein, N., & Hack, M. (2002). Sources of variability in sequelae of very low birth weight. *Child Neuropsychology*, *8*, 164–178.

Taylor, H. G., Hack, M., Klein, N., & Schatschneider, C. (1995). Achievement in children with birth weights less than 750 grams with normal cognitive abilities: Evidence for specific learning disabilities. *Journal of Pediatric Psychology*, *20*, 703–719.

Taylor, H. G., Klein, N., Drotar, D., Schluchter, M., & Hack, M. (2006). Consequences and risks for <1000 g birth weight for neuropsychological skills, achievement, and adaptive functioning. *Journal of Developmental and Behavioral Pediatrics*, *27*, 459–469.

Taylor, H. G., Klein, N., & Hack, M. (2000). School-age consequences of <750 g birth weight: A review and update. *Developmental Neuropsychology*, *17*, 289–321.

Taylor, H. G., Klein, N., Minich, N., & Hack, M. (2000). Middle school-age outcomes in children with very low birth weight. *Child Development*, *71*, 1495–1511.

Taylor, H. G., Klein, N., Schatschneider, C., & Hack, M. (1998). Predictors of early school age out-comes in very low birth weight children. *Developmental and Behavioral Pediatrics*, *19*, 235–243.

Taylor, H. G., Lean, N., & Schwartz, S. (1989). Psendoword repetition ability in learning disabled children. *Applied Psycholinguistics*, *10*, 203–219.

Taylor, H. G., Minich, N., Bangert, B., Filipek, P. A., & Hack, M. (2004). Long-term neuropsychological outcomes of very low birth weight: Associations with early risks for periventricular brain insults. *Journal of the International Neuropsychological Society*, *10*, 987–1004.

Taylor, H. G., Minich, N., Klein, N., & Hack, M. (2004). Longitudinal outcome of very low birth weight: Neuropsychological findings. *Journal of the International Neuropsychological Society*, *10*, 149–163.

Tonry, M., Ohlin, L. E., & Farrington, D. P. (1991). *Human development and criminal behavior: New ways of advancing knowledge*. New York: Springer-Verlag.

Whitaker, A. H., Van Rossem, R., Feldman, J., Schonfeld, I., Pinto-Martin, J., Torre, C., et al. (1997). Psychiatric outcomes in low-birth-weight children at age 6 years: Relation to neonatal cranial ultrasound abnormalities. *Archives of General Psychiatry*, *54*, 847–856.

Wilson-Costello, D., Friedman, H., Minich, N., Siner, B., Taylor, G., Schluchter, M., et al. (2007). Improved neurodevelopmental outcomes for extremely low birth weight infants in 2000–2002. *Pediatrics*, *119*, 37–45.

Wolke, D., & Meyer, R. (1999). Cognitive status, language attainment, and prereading skills of 6-year-old very preterm children and their peers: The Bavarian Longitudinal Study. *Developmental Medicine & Child Neurology*, *41*, 94–109.

Woodcock, R., & Johnson, M. B. (1989). *Woodcock–Johnson Tests of Achievement–Revised – Standard and supplemental batteries*. Allen, TX: DLM Teaching Resources.

Yuan, K. H., & Bentler, P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. In M. E.Sobel & M. P.Becker (Eds.), *Sociological methodology 2000* (pp. 165–200). Washington, DC: American Sociological Association.

Zelkowitz, P., Papageorgion, A., Zelazo, P. R., & Weiss, M. (1995). Behavioral adjustment in very low and normal birth weight children. *Journal of Clinical Child Psychology*, *24*, 21–30.