

# Early Identification of Reading Difficulties Using Heterogeneous Developmental Trajectories

Christy Kim Boscardin and Bengt Muthén  
University of California, Los Angeles

David J. Francis  
University of Houston

Eva L. Baker  
University of California, Los Angeles

Serious conceptual and procedural problems associated with current diagnostic methods call for alternative approaches to assessing and diagnosing students with reading problems. This study presents a new analytic model to improve the classification and prediction of children's reading development. Growth mixture modeling was used to identify the presence of 10 different heterogeneous developmental patterns. In all, 411 children in kindergarten through Grade 2 from 3 elementary schools in Texas were administered measures of phonological awareness, word recognition, and rapid naming skills 4 times a year. The mean ages were 5.8 years ( $SD = 0.35$ ) for the kindergartners, 6.9 years ( $SD = 0.39$ ) for Grade 1, and 8.0 years ( $SD = 0.43$ ) for Grade 2; the percentage of boys was 50%. The results indicate that precursor reading skills such as phonological awareness and rapid naming are highly predictive of word reading (word recognition) and that developmental profiles formed in kindergarten are directly associated with development in Grades 1 and 2. Students identified as having reading-related difficulties in kindergarten exhibited slower development of word recognition skills in subsequent years of the study.

*Keywords:* reading development, screening, reading skills, achievement, longitudinal studies

Recently, with mounting evidence for early detection as the key to remediation and prevention for later reading difficulties, there has been a growing interest and urgent call for earlier identification of children with reading difficulties. According to research funded by the National Institute of Child Health and Human Development (National Reading Panel, 2000), if intervention is delayed until 9 years of age (the time when most children with reading difficulties typically receive services), approximately 75% of children will continue to have reading difficulties in later grades (Lyon, 1998). The majority of children identified as having reading difficulties in early grades will continue to have problems in later grades without appropriate instructional intervention (Juel, 1988; Scarborough, 1998). Satz and Fletcher (1988) suggested that interventions are most effective if implemented prior to the overt manifestation of disability. Schenck, Fitzimmons, Bullard, Taylor, and Satz (1980) also concluded that children at high risk who received intervention

early demonstrated significant improvement in academic performance over time. Earlier studies have shown that older children who were identified as having reading difficulties would not have required learning disability status if their difficulties had been diagnosed and they had received intervention at an early age (De Hirsh, Jansky, & Langford, 1966; Strag, 1972).

Previous studies have shown that overall academic success in later grades can be predicted with reasonable accuracy by using reading outcomes at early grades (Slavin, 1994; Strag, 1972; Torgesen & Wagner, 2002). Previous longitudinal studies have also suggested that children at risk for reading difficulties can be identified much earlier than previously thought (Juel, 1988; Shaywitz, Escobar, Shaywitz, Fletcher, & Makuch, 1992).

## Identification With IQ Discrepancy-Based Method

Despite the importance of early detection, previous methods for identifying children with reading difficulties suffered from the lack of a theoretical foundation and supportive evidence for validity, which unnecessarily delayed identification (Lyon et al., 2001). Previously, children were identified as having reading difficulties if there was a substantial discrepancy between a child's aptitude, typically operationalized by IQ, and his or her reading performance (Francis, Fletcher, Shaywitz, Shaywitz, & Rourke, 1996; Gunning, 1998). Although the IQ discrepancy-based method has been the most widely used definition of reading difficulty, there were several conceptual and measurement problems that warranted an alternative method of identification of persons with dyslexia and other poor readers (Francis, Shaywitz, Stuebing, Shaywitz, &

---

Christy Kim Boscardin and Eva L. Baker, Graduate School of Education and Information Studies and Center for Research and Evaluation, Standards and Student Testing, University of California, Los Angeles; Bengt Muthén, Graduate School of Education and Information Studies and Department of Statistics, University of California, Los Angeles; David J. Francis, Department of Psychology and the Texas Institute for Measurement, Evaluation, and Statistics, University of Houston.

Correspondence concerning this article should be addressed to Christy Kim Boscardin, University of California, Los Angeles Graduate School of Education and Information Studies, 300 Charles E. Young Drive North, Room 101C, 319 GSE&IS/Mailbox 951522, Los Angeles, CA 90095-1522. E-mail: christyk@ucla.edu

Fletcher, 1996; Shaywitz et al., 1992). Dissatisfaction with previous approaches to identification of children with reading difficulties led to consideration of alternative approaches such as Response to Intervention (RTI) in the recent reauthorization of the Individuals with Disabilities Education Act. The core concept of RTI is based on the premise that a student exhibiting a slower rate of development and failure to respond adequately to intervention may be identified as requiring special services and being at risk for learning disability (Fletcher, Coulter, Reschly, & Vaughn, 2004). With RTI, the focus is on screening, instructional intervention, and continual monitoring. Proponents of the RTI approach suggest that with continuous progress monitoring, the focus will be shifted to prevention and intervention rather than relying on the current "wait-to-fail" model facilitated by the use of IQ-discrepancy approaches (Compton, Fuchs, Fuchs, & Bryant, 2006). Early and accurate identification of students at risk for reading difficulties will be fundamental to successful implementation of the various RTI models proposed (Bradley, Danielson, & Hallahan, 2002; Compton et al., 2006; Fuchs, Mock, Morgan, & Young, 2003). As illustrated by use of the IQ discrepancy-based method in the past, using test scores from a single time point is often unreliable and insufficient for identification of reading difficulties.

Early identification of children with reading difficulties will require a system that accurately predicts which children are at risk for reading failure. However, previous screening procedures have yielded unreliable results due to high rates of classification and measurement errors (Fletcher & Satz, 1984; Jenkins & O'Connor, 2002; Scarborough, 1998). Speece (2005) suggested that one of the reasons for the rate of inaccuracy and the problems associated with early identification is lack of consideration and disregard for potential change and growth in the reading development process. Considering that growth and development are fundamental to the concept of learning, it seems only logical to consider growth (longitudinal) data as the primary source for the identification of reading problems.

### Predictors of Reading

Much progress has been made in the past 20 years in understanding the correlates of reading development and the key predictors of reading outcomes (National Reading Panel, 2000; National Research Council, 1998). One of the most significant predictors of early difficulties in acquiring accurate and fluent word recognition skills has been identified as the individual differences in phonological skill (Jenkins & O'Connor, 2002; Liberman, Shankweiler, & Liberman, 1989; Parrila, Kirby, & McQuarrie, 2004). There seems to be a wide consensus that deficits in phonological awareness are related to later reading difficulties (Catts & Kamhi, 1999; Stanovich & Siegel, 1994; Torgesen, Wagner, Rashotte, Burgess, & Hecht, 1997). More recently, some studies have suggested that deficits in both phonological awareness and in serial naming speed may produce more severe reading difficulties, providing evidence in favor of the double-deficit model (Morris et al., 1998; Schatschneider, Carlson, Francis, Foorman, & Fletcher, 2002; Wolf & Bowers, 2000). However, investigation of the relationship between these two skills on reading development over time and the relative importance of these two skills in early identification of children with reading difficulties

has been limited. In this study, we examined the relationship between the development of precursor skills such as phonological awareness and rapid naming and the development of word recognition skill in later grades.

In light of the current call for alternative approaches to identification of students with reading difficulties, we offer a novel approach to examining reading development through the application of a new longitudinal clustering technique called *growth mixture modeling*. The purpose of this study was to determine whether distinctive groups of students with various developmental profiles can be identified based on precursor reading skills and whether these profiles will help characterize the course of reading difficulties manifested during different developmental periods. Analysis of individual growth curves based on early reading skills is an alternative statistical approach that will potentially be helpful in earlier identification of students with reading difficulties. These individual growth curves may not only yield earlier diagnosis of reading difficulties but may also provide contextual data for further analysis of individual differences and problems.

### Growth Modeling

Conventional growth modeling has been a useful technique for examining individual differences in learning development (Bryk & Raudenbush, 1992; Jennrich & Schluchter, 1986; Laird & Ware, 1982; Lindstrom & Bates, 1988; Muthén, Khoo, Francis, & Boscardin, 2002). To study development or change over time, the researcher represents individual outcomes using rates of growth over multiple time points. In growth curve models, individual reading development can be formulated in terms of initial reading level and the rate of learning or development of reading. Typically, the analysis of individual change is represented using random coefficient modeling. Individual change and growth over time can be represented through growth models in order to provide a more dynamic view of reading development.

### Growth Mixture Modeling

More recently, a new modeling technique that takes into consideration the effects of heterogeneity in a sample to provide more reliable predictions for later development has been introduced (Muthén, 2000; Muthén et al., 2002; Muthén & Muthén, 2000). Because the conventional growth modeling approach estimates the variation in growth curves as a function of growth factors, interactions of growth factors are difficult to detect and model. In contrast, in the growth mixture modeling framework, interactions between heterogeneous individual growth curves and background factors can be examined without the restrictions of conventional growth modeling techniques. Because growth mixture models allow for the effect of covariates to differ based on the developmental profiles, researchers can determine the relationship between rapid naming skills or a student's background on his or her developmental trajectory. Consequently, the application of growth mixture modeling is extremely useful for developing intervention protocols, as well as identifying specific problems related to development of particular reading skills and to other factors that may contribute to individual differences in development. Growth mixture modeling offers significant advantages over conventional

growth modeling techniques (Muthén, 2000; Muthén et al., 2002). Muthén et al. provided precursory examination of how growth mixture modeling can be used for the identification of several different developmental profiles. In the current study, we expanded on the previous studies on the utility of growth mixture modeling to determine whether distinctive groups of students with various developmental profiles can be identified. Specifically, given the importance of phonological awareness skills in early readers, we examined (a) the relationship between initial status and development of phonological awareness during early stages of reading development and of word recognition skills as well as (b) the contribution of rapid naming skills in differentiating poor readers.

## Method

### Sample

The sample for the current study was drawn from a larger study consisting of 945 students from a cohort-sequential longitudinal study designed to assess the development of early reading skills (Schatschneider et al., 2002; Schatschneider, Fletcher, Francis, Carlson, & Foorman, 2004; Schatschneider, Francis, Foorman, & Fletcher, 1999). The cohort-sequential longitudinal design for sample selection is presented in Table 1. The larger study sample represented a random selection of children in kindergarten through Grade 2 from three elementary schools in Texas. The students represented a random selection (80%) of students who had parental consent to participate in the study. Students with evidence of severe emotional problems, uncorrected vision problems, hearing loss, acquired neurological disorders, or classification at the lowest level of English as a second language based on school designation were excluded from the study. Children referred for special services in kindergarten were also excluded from participating in this study. Students with later referrals for special education services, however, were included in the study. Measurements of early reading skills were taken four times a year from kindergarten through Grade 2. During a given academic year, students were tested four times (October, December, February, and April) on measures related to reading skills.

For our research purposes, the 411 students with complete data for at least kindergarten and Grade 1 were selected out of the original 945 students for the present study. There were no statistically significant differences between the original sample and this subset on any of the measures. Out of these 411 students, only 208 students had complete data for all three grades. Loss of Grade 1 and Grade 2 in this subset was due to the termination of the study, which prevented us from following some students into Grade 2, and one school's decision not to participate in Year 4. For the

subset of children who had at least complete data in kindergarten and Grade 1 ( $n = 411$ ), about 50% of the sample were boys. The ethnic breakdown of the subset sample was 55% White, 17% African American, 16% Hispanic, 11% Asian, and 1% other ethnicity. The mean ages were 5.8 years ( $SD = 0.35$ ) for the kindergartners, 6.9 years ( $SD = 0.39$ ) for Grade 1, and 8.0 years ( $SD = 0.43$ ) for Grade 2. For the purposes of the present study, White and Asian students were considered to be nonminority, whereas Black, Hispanic, and other students were categorized as minority. The Hollingshead (1975) Four Factor Index of Social Status survey administered to parents was used to collect data on socioeconomic status (SES): (a) 8% were classified as lower class, (b) 40% as working class, (c) 45% as middle-upper class, and (d) the remaining 7% did not provide these data.

### Measures

A skills assessment battery was administered individually to each child in October, December, February, and April between kindergarten and Grade 2. Kindergarten students were assessed in phonological awareness and rapid naming. In first and second grades, students were assessed in word recognition. Descriptive statistics on the measures are shown in Table 2.

### Phonological Awareness

The phonological awareness test that was administered was an experimental version of the Comprehensive Test of Phonological Processes developed by Wagner, Torgesen, and Rashotte (1999). A detailed description of the test is provided in Schatschneider, Fletcher, Francis, Carlson, and Foorman (2004). The assessment consisted of seven subtests, including (a) phoneme segmentation, (b) phoneme elision, (c) sound categorization, (d) first sound comparison, (e) blending onset and rime, (f) blending phonemes into words, and (g) blending phonemes into nonwords.

**Phoneme segmentation.** This task required children to listen to words and then tell the interviewer "each sound you hear in the word in the order that you hear it." There were 4 practice items and 15 test items, consisting of one- and two-syllable words with two to five phonemes (e.g., *ate*, *up*, *jump*).

**Phoneme elision.** For the test for phoneme elision, the child was asked to say the word after deleting a specific phoneme (e.g., "Say the word *cup*. Now tell me what word would be left if I said *cup* without the /k/ sound"). There were 4 practice items and 15 test items. All phonemes deleted were consonants. The first 12 items were three-phoneme single-syllable words, for which the deletion was at the end of the word for the first 6 items and the beginning of the word for the other 6 items. The last three items were three- to five-phoneme two-syllable words for which the consonant to be deleted was in the middle (e.g., *ti[g]jer*).

**Sound categorization.** This task asked a child to select one of four words that did not share a phoneme with the rest (e.g., "In the set of *fun*, *pin*, *bun*, and *gun*, select the one that doesn't sound like the others").

**First sound comparison.** In this task, the child was asked to point to the picture of the word that begins with the same sound as the presented word. For example, a booklet with pictures of a rug, a saw, and ash was presented, with a target word of *rake*. The

Table 1  
*Cohort Structure for the Larger Study*

Cohort	Year 1	Year 2	Year 3	Year 4
1	K ( $n = 183$ )	1 ( $n = 170$ )	2 ( $n = 133$ )	
2		K ( $n = 210$ )	1 ( $n = 158$ )	2 ( $n = 91$ )

Note. K = kindergarten; 1 = Grade 1; 2 = Grade 2.

Table 2  
Descriptive Statistics for the Measures Used in the Study  
in IRT Scale

Variable	<i>M</i>	<i>SD</i>
Phonological awareness at Time 1	−1.22	.59
Phonological awareness at Time 2	−1.01	.65
Phonological awareness at Time 3	−0.79	.74
Phonological awareness at Time 4	−0.60	.82
Word recognition at Time 1	−0.98	.84
Word recognition at Time 2	−0.68	.85
Word recognition at Time 3	−0.40	.90
Word recognition at Time 4	−0.12	.90
Word recognition at Time 5	0.26	.76
Word recognition at Time 6	0.46	.77
Word recognition at Time 7	0.60	.77
Word recognition at Time 8	0.77	.77
Rapid naming at Time 4 (transformed)	−0.26	.74

Note. Test scores are item response theory (IRT) scores, standardized to have a mean of 0 and a standard deviation of 1.

correct response in this example is *rug*, which corresponds to the first sound of *rake* (Schatschneider et al., 2004).

**Blending onset and rime.** This task required the child to pronounce a word after the onsets (initial consonants or consonant cluster in a syllable) and rimes (vowels and remaining consonants of the word) had been combined. There were 15 test items, with the number of phonemes in the single-syllable words varying from three to four (e.g., *mouse*, *child*).

**Blending phonemes into words.** This task was identical to blending onset and rime except that in this task the child was asked to blend phonemes rather than onsets and rimes. The child was presented with a string of phonemes at a rate of two per second and asked to repeat them by putting the sounds together. There were 6 practice items and 15 test items (one- and two-syllable words) consisting of two to six phonemes (e.g., *i-f*, *t-oy*, *w-a-sh*, *b-a-m-b-oo*).

**Blending phonemes into nonwords.** This task was identical to blending phonemes into words except that nonwords were used in place of real words, with a parenthetical real word rhyme or near rhyme provided as a pronunciation key (e.g., *i-th* [*with*], *y-a-s* [*gas*], *th-u-ng* [*rung*]).

Internal consistency estimates for the subtests ranged from .85 to .95 on all occasions. With respect to concurrent validity, the subtests were correlated with the Lindamood Auditory Conceptualization Test (Lindamood & Lindamood, 1979). Correlations ranged from .41 to .75. For the analysis, we combined the scores on these seven phonological awareness tasks into one latent ability score. Instead of using raw total scores of phonological ability, we used item response theory (IRT) scores based on estimates of each person's latent phonological trait to represent phonological awareness with a mean of 0 and a standard deviation of 1.

### Word Recognition

Skills in word recognition were assessed in Grades 1 and 2 by asking students to read aloud 50 words on 4 × 6-in cards. The Grade 1 and Grade 2 lists each consisted of 50 words, with 16 words in common across the two grades. The 50 words included 36

single-syllable, 11 two-syllable, and 3 three-syllable real words. Words were matched for frequency of occurrence (Carroll, Davies, & Richman, 1971) and spanned first- through third-grade levels of difficulty. The internal consistency estimates exceeded .90 in the present study on all occasions. Concurrent and predictive validity for the word list were high, as evidenced by .80 correlations with the Letter Word and Word Attack subtests of the Woodcock–Johnson Psychoeducational Battery–Revised (Woodcock & Johnson, 1989). For the analysis, we used IRT scores based on estimates of each person's latent trait on word recognition skills as an indicator for word recognition ability. IRT estimates were scaled to have a mean of 0 and a standard deviation of 1.

### Rapid Naming

Denckla and Rudel's (1976) Rapid Automatized Naming Tests for Letters was administered in kindergarten. Rapid automatized naming letters were high-frequency lowercase letters (e.g., *a*, *d*, *o*, *s*, *p*). The stimuli consisted of five letters in a row, repeated 10 times in random sequences. The child was asked to name each letter as quickly as possible. The correct number of responses within 60 s was recorded. Test–retest reliability was .57 for kindergarten (reflecting variability in true change over this age range) and .77 for Grades 1 and 2. Children who did not know all five letters were not administered the test (Schatschneider et al., 2004). For the purposes of the current study, the scores obtained from April data collection were log-transformed and included in the analysis.<sup>1</sup>

Descriptive statistics based on the IRT scale on all of the measures are shown in Table 2. For phonological awareness, scores ranged from 1.22 *SD* below the mean to about 0.60 *SD* below the mean across the four time points in kindergarten. For word recognition, scores ranged from 0.98 *SD* below the mean to 0.77 *SD* above the mean across the eight time points during Grades 1 and 2.

### Analysis

The application of growth modeling with mixture components has been explored by other researchers (Nagin, 1999; Verbeke & LeSaffre, 1996); however, recent work introduced by Muthén (2000) and Muthén and Shedden (1999) has provided a much more flexible framework than previous models. General growth mixture modeling introduced by Muthén and colleagues provides technical advantages over conventional growth models by allowing greater flexibility in model specifications and assumptions. One of the theoretical assumptions of the conventional growth model is that the data come from a single population and that the single-population model accounts for all of the variation in the individual trajectories. As both the data and developmental theory suggest, however, there may be several heterogeneous subgroups within this population that require different sets of model specifications and assumptions. For example, as shown in Figure 1, although

<sup>1</sup> For the rapid naming variable, we combined the speed measure, RNL\_S (number correct/number of seconds) with total correct (RNL\_TR) to create TRNL (rapid naming), where  $TRNL = \log_2(RNL\_TR / RNL\_S + .1)$ .



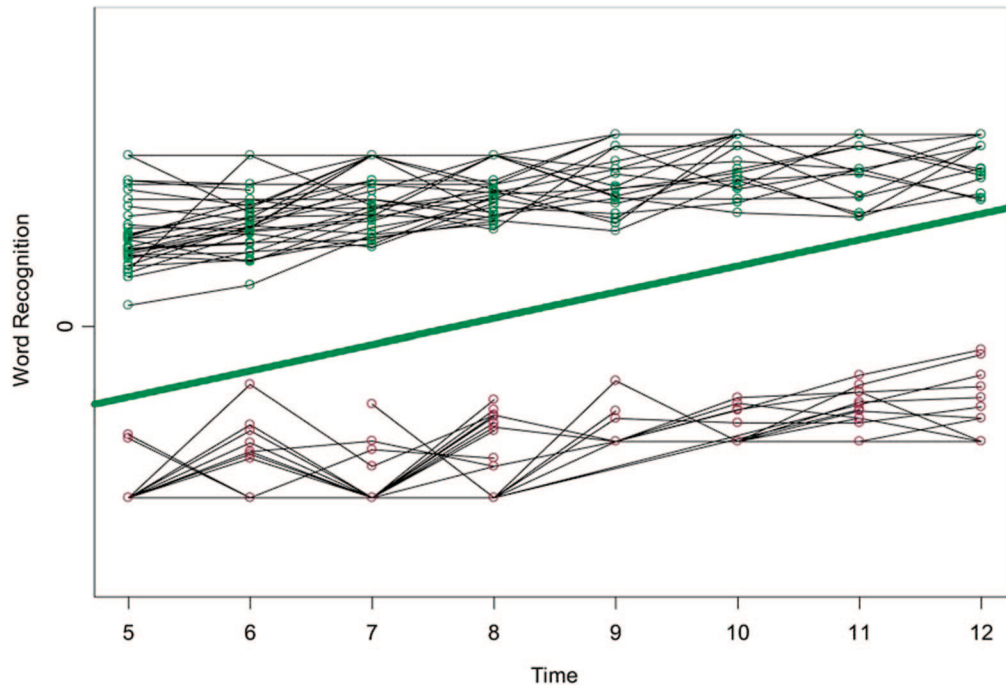


Figure 1. Individual growth trajectories for word recognition development. The two colors represent two distinct growth profiles of reading development.

there may be two different groups of students with very different reading development trajectories, with the conventional growth modeling techniques it would be difficult to detect the misspecification of the model. With the conventional growth models, the estimation of growth is determined by a single collection of growth trajectories with a single vector of means and covariance parameter estimates for intercept and slope parameters. In contrast, growth mixture modeling allows the identification of different subgroups in the model that represent the different collections of reading development trajectories. Hypothetically, the kindergarten growth parameters for a group of students with high intercept (exit level at kindergarten) and growth may influence reading development in Grade 1 differently than for students with low intercept and growth in kindergarten. Accordingly, individuals with similar exit levels in kindergarten may belong to different subgroups with different rates of development. An accelerated growth rate of precursors in kindergarten may suggest that these students are qualitatively different from students with no significant growth during the same time period. Accelerated growth can also be interpreted as a higher aptitude for learning, which could have a greater influence on reading development in Grade 1. Additionally, one could hypothesize heterogeneity in the influence of covariates on the different developmental trajectories.

Growth mixture modeling generalizes conventional growth modeling by allowing heterogeneity of different subgroups in the population through the use of a categorical latent variable (Muthén, 2000, 2001; Muthén & Shedden, 1999). These categorical latent variables or latent classes can represent multiple groups with different developmental trajectories for which group membership is unknown but can be inferred from the data. Individuals are

allowed to be in one of  $K$  latent classes, each with characteristically distinct developmental profiles. Within a class, individual developmental trajectories are allowed to vary around this class profile. For each class  $k$ , continuous outcome variables  $Y$  are assumed to be normally distributed conditional on covariates  $x$ . The growth mixture model can be expressed as follows:

$$Y_{ik} = \nu_k + \Lambda_k \eta_{ik} + K_k x_{ik} + \varepsilon_{ik} \quad (1)$$

and

$$\eta_{ik} = \alpha_k + B_k \eta_{ik} + \Gamma_k x_{ik} + \zeta_{ik}. \quad (2)$$

Here,  $Y$  represents the repeated measures over fixed time points. The  $\eta_{ik}$  are random effects, and  $\Lambda_k$  represents time scores for the shape of the growth curves.  $K_k$  represents the effects of time-varying covariates, and  $\Gamma_k$  represents the effects of time-invariant covariates.  $\alpha_k$  represents the intercepts for  $\eta$  for latent class  $k$ . For the example we discuss,  $\alpha_k$  represents the different reading development trajectories for the different classes. The residual vectors  $\varepsilon_{ik}$  and  $\zeta_{ik}$  are assumed to have covariance matrices  $\Theta_k$  and  $\Psi_k$ , respectively.

The  $K$  trajectory classes are allowed to include variation in both intercepts and slopes in phonological awareness and word recognition. This framework, introduced by Muthén et al. (2002), Muthén and Muthén (2000), and Muthén and Shedden (1999), is much more wide ranging than the mixture models proposed by Nagin (1999), in which it is assumed that  $\Psi_k = 0$  and  $\Theta_k = \theta I$ . A model proposed by Muthén (2000) provides more flexibility by allowing class-to-class variation of the covariance matrices  $\Psi_k$  and  $\Theta_k$ . This modeling specification is particularly important when determining

the number of latent class trajectories. Depending on how the degree of class invariance is specified, different values for model fit criteria will be obtained.

Estimated posterior probabilities for each individual's class membership are derived as follows. Define the latent class membership indicators  $c_{ik}$  to be 1 if individual  $i$  belongs to class  $k$ , and 0 otherwise. Then

$$p_{ik} = P(c_{ik} = 1 | \mathbf{y}_i, \mathbf{x}_i) \propto P(c_{ik} = 1 | \mathbf{x}_i) f(Y_{ik} | \mathbf{x}_i). \quad (3)$$

In this study, the individual students were assigned to a class based on their highest estimated posterior probabilities. The posterior probabilities were computed for a given individual observation vector  $(\mathbf{y}_i, \mathbf{x}_i)$ . In other words, for a given model, individual students' posterior probabilities were computed as a function of the model estimates and the individuals' values on observed variables.

To understand the composition of classes and also to provide stability in class membership, Muthén (2000) introduced a multinomial logistic regression model to represent the relationship between  $c$  (latent class variable) and  $x$  (covariate). The multinomial logistic regression for predicting class membership with a covariate can be expressed as follows:

$$P(c_{ik} = 1 | x_i) = \exp(\beta_{0k} + \beta_{1k}x_i) / \sum_{c=1}^K \exp(\beta_{0c} + \beta_{1c}x_i), \quad (4)$$

for  $k = 1, \dots, K$  where we constrain  $\beta_{0K} = 0$  and  $\beta_{1K} = 0$ , and where  $P(c_{ik} = 1 | x_i)$  is the probability of being in class  $c_{ik}$  conditioned on covariate  $x$ ,  $\beta_{0k}$  is the class intercept, and  $\beta_{1k}$  is the regression coefficient for the  $k$ th class on  $x_i$ , the covariate. The parameter estimates in the model are most easily interpreted by computing predicted class membership probabilities as a function of the covariate.

One of the advantages of this general latent variable modeling framework is that one can systematically explore the influence of precursor skills on later development. For example, students with rapid development of phonological awareness in kindergarten, despite their low performance at the entry level, may continually show rapid development in word recognition as well. In contrast, students with a low entry level and slow development of phonological awareness may not improve much in word reading when they enter Grades 1 and 2. Despite the similarities in initial appearance between these two groups of students, the students with slow development may be qualitatively different from the students with rapid development of phonological awareness. Also, the differences in students' developmental trajectories in kindergarten may differentially influence their later reading development.

### Model Selection and Model Fit

In growth mixture modeling, determination of the optimal number of groups that best represents the data is part of the model selection procedure. The best method for determining the number of classes or groups is still a topic of controversy. For the present study, we considered two statistical indices as well as the overall interpretability of the model based on class counts and substantive theory for model selection. For comparison of nested models with

the same number of classes, the usual likelihood ratio chi-square difference can be used. However, when comparing models with different numbers of classes, the likelihood ratio test may no longer be applicable, and other information criteria must be used. Although McLachlan and Peel (2000) suggested assessing the number of modes of a distribution using the kernel method to estimate the density function, one drawback to this approach is that when classes or the components are not sufficiently separated, the mixture distribution can look unimodal, thus failing to detect the actual number of modes (classes). Also, if the data have a skewed distribution, using a normal mixture model will not be appropriate for capturing the number of classes (McLachlan & Peel, 2000). We used the Bayesian information criterion (BIC) to compare model fit between non-nested models. For a given model, BIC is calculated as follows:

$$\text{BIC} = -2 \log L + r \log n. \quad (5)$$

Here,  $L$  is the value of the model's maximized likelihood,  $n$  is the sample size, and  $r$  is the number of parameters in the model. To determine the optimal number of classes for the best representation of the data, one compares BIC values across the different models, with smaller BIC values indicating a better model fit. However, the overall model selection is guided not only by BIC values but also by entropy indices (described in the next paragraph) and the interpretability of the chosen model, because the BIC tends to favor models with fewer classes (Wiesner & Windle, 2004).

Model selection was also guided by examining the reliability of the classifications via the estimated posterior probabilities of class membership for each individual (Muthén, 2000). The precision of the classification can be assessed by how well the students are being classified into each class. A reliable classification will require the student to have posterior probabilities that are very high for belonging to a single class and very low for belonging to all of the other classes. These probabilities determine the most likely class for each student. For example, a student's estimated probabilities may be .80 for Class 1, .15 for Class 2, .05 for Class 3, and 0 for Classes 4 and 5. A reliable classification is linked to the precision of probabilities in differentiating class membership. To check for precision in classification, one summarizes the probabilities into average posterior probabilities. For example, if the average posterior probability of Class 1 is .90, then one can conclude that the students being classified into Class 1 have, on average, very high probabilities of Class 1 membership. In addition, the quality of the classification is also summarized by the entropy measure (Muthén, 2000). Entropy is expressed as follows:

$$E_K = 1 - \{ \sum_i \sum_k [-p_{ik} \log(p_{ik})] / n \log(K) \}. \quad (6)$$

The expression  $E_K$  is bounded between 0 and 1. An entropy measure close to 1 is considered to be evidence of good classification.

### Analysis Procedure

In the first analysis stage, the development of phonological awareness in kindergarten was examined separately from the development of word recognition, because these represent two distinct (although developmentally sequential) linguistic processes.

The number of classes, as well as the type of growth trajectories, may vary from one grade to the next depending on the different types of latent growth trajectories present. Consequently, a four-group model might be best for representing the developmental profiles in kindergarten, but a five- or six-group model may be more appropriate for representing development in Grades 1 and 2.

Once the number of classes was determined for phonological awareness and word recognition development separately based on the BIC values, the entropy indices, and the interpretability of the model, the final model combining all three grades was selected. The purpose of these two separate analyses in the first stage was to provide a basis for selecting the final model that combined all 3 years of data. Although the developmental trajectories identified in the separate analyses were informative, the overall goal of the study was to find the best representation of data using all three grades. The final model selection was then guided by class counts and the overall model fit.

For the final model, we checked the overall model fit and the quality of the classification by examining how closely the estimates matched the observed data. One way to check this is to compare the estimated mean curve with the observed trajectories of individuals or the observed mean curves based on individual estimated conditional class probabilities. For this technique, we assigned individuals to classes based on the estimated posterior probabilities and then compared the individual trajectories with the estimated mean trajectory (Bandeau-Roche, Miglioretti, Zeger, & Rathouz, as cited in Muthén, 2000). Indication of a good model fit requires close alignment between the individual trajectories and the estimated mean trajectory.

For the present study, we used all 12 time points to investigate reading development across three grade levels (kindergarten, first, and second). For kindergarten, the IRT scale scores on phonological awareness from four different time points (October, December, February, and April) represented reading development. For Grades 1 and 2, we used the IRT scale scores on word recognition spanning 2 years with eight time points to represent continual reading development.

## Results

### *Initial Model Selection Using Two Separate Analyses*

Prior to identification of developmental profiles representing all three grade levels, we conducted preliminary analyses for each precursor (phonological awareness) and reading skills (word recognition) separately. First, models with between two and six classes were fit to the longitudinal phonological awareness data from the kindergarten portion of the data. As Table 3 presents, although the four-class model had the lowest BIC value, upon further examination of the class proportions for the four-class model, we determined that the four-class model lacked distinction between the different classes and was difficult to interpret, as shown in Table 4. Given the minimal difference in the entropy values between the four-class and the five-class model, the five-class model was finally chosen for interpretability purposes.

The estimated mean growth curves representing the five different developmental profiles in kindergarten are shown in Figure 2. As Figure 2 illustrates, PA 1 students represented the lowest

Table 3  
*BIC and Entropy Values for Kindergarten Models*

Class	BIC	Entropy
2	1,380.66	0.83
3	1,369.44	0.75
4	1,358.15	0.74
5	1,370.02	0.70
6	1,382.06	0.73

Note. BIC = Bayesian information criterion.

performing group. The estimated means and the standard errors for intercepts and slopes for the five-class model are presented in Table 5. As shown in Table 5, the mean of PA 1 students' phonological scores at the end of kindergarten was about 1.55 *SD* below the mean, compared to about 1.07 *SD* below the mean for PA 2 students and 0.26 *SD* below the mean for PA 3. Both the PA 4 and PA 5 students performed higher, with 0.10 and 1.14 *SD* above the mean than the overall average for phonological awareness, which was 0.60 *SD* below the mean at the end of kindergarten. The slope estimate for PA 1 students was 0.03, representing flat trajectory, compared to 0.14 for PA 2, 0.40 for PA 3, 0.27 for PA 4, and 0.34 for PA 5. As shown in Figure 2 and Table 5, PA 1 students started out with low performance in phonological awareness and exhibited no significant improvement throughout the entire kindergarten year. Given the lack of growth during the year, this group of students would be expected to be most at risk for developing reading difficulties in later grades. The estimated means for the slope and intercept for the five-class model are presented in Table 5 and illustrated in Figure 3.

Next, a separate analysis was conducted to fit the word recognition data from Grades 1 and 2. As shown in Table 6, after evaluating different numbers of models using BIC and entropy indices, we found the five-class model to be the best in terms of fit and interpretability. There was a significant decrease in the BIC values between the four-class and the five-class models, however BIC values started to level off after the five-class model. Also, the difference in entropy was minimal across five-class, six-class, and seven-class models. The estimated mean growth curves for the five-class model are shown in Figure 3. After evaluating the BIC values and entropy indices, we determined that the five-class model was the best in terms of fit and interpretability.

As shown in Table 7 and Figure 3, WR 1 had the lowest intercept and slope in the five-class model. The average score for word recognition at the end of Grade 2 was 0.77 *SD* above the mean. For WR 1 students, the average score was 0.95 *SD* below the mean, which was significantly lower than the next low performing group (WR 2), which had an average score of 0.46 *SD* above the mean. Students in WR 1 were thus characterized as the students who were at risk for reading difficulties.

### *Final Model: Combining Models of Phonological Awareness and Word Recognition*

The selected five-class model for phonological awareness and the selected five-class model for word recognition were next combined to allow growth modeling for all 3 years. In the com-

Table 4  
*Class Counts and Proportion of Students in the Kindergarten Four-Class Model*

Variable	Class 1	Class 2	Class 3	Class 4
Class count	229	20	131	31
Proportion of total sample	56	5	32	7

bined analysis, as shown by the path diagram in Figure 4, development of phonological awareness in kindergarten was represented by Intercept 1 and Growth 1, development of word recognition was represented by Intercept 2 and Growth 2, classes were represented as latent variables C1 and C2, and rapid naming measured at the end of kindergarten was also added in the model as a covariate for class membership. Given previous research on phonological awareness and rapid naming, we believed that although there is significant overlap between these two skills, they contribute independently to reading and that rapid naming can be considered an etiologically distinct source of variance in reading outcomes (Petrill, Deater-Deckard, Thompson, DeThorne, & Schatschneider, 2006a).

As an initial step in the analysis, we considered the model in which all students stayed in the same developmental trajectory throughout all 3 years, as shown in Table 8. However, in the final model, to provide a more realistic representation of the data, students were allowed to change class membership during the transition from kindergarten (phonological awareness) to Grades 1 and 2 (word recognition). For example, a student classified as PA 1 based on phonological awareness development was allowed to

Table 5  
*Intercept and Slope for Five Phonological Awareness Development Profiles*

Profile	Intercept <i>M (SE)</i>	Slope <i>M (SE)</i>
PA 1	−1.55 (0.26)	0.03 (0.04)
PA 2	−1.07 (0.23)	0.14 (0.06)
PA 3	−0.26 (0.19)	0.40 (0.16)
PA 4	0.10 (0.15)	0.27 (0.04)
PA 5	1.14 (0.31)	0.34 (0.09)

*Note.* PA = phonological awareness.

progress to a WR 2 or to a WR 3 profile based on word recognition development. By allowing change in class membership, we were able to determine which students stayed in the same developmental trajectories throughout 3 years and which students changed their class membership. The assumption was that if students did stay within the same class throughout the three grades, then the kindergarten classification should correspond directly to Grade 1 and Grade 2 classification. To check for this assumption, we fitted the observed growth curves in Grades 1 and 2 to the estimated mean curves from the kindergarten classes. As illustrated in Figure 5, the observed growth curves in Grades 1 and 2 did not fit the estimated mean curve represented by the dark solid line. Subsequently, we determined that not all PA 1 students corresponded directly to WR 1 students, and not all PA 2 students corresponded directly to WR 2 students, and so on. Instead, as expected, some students did move into other developmental trajectories in Grades 1 and 2. To

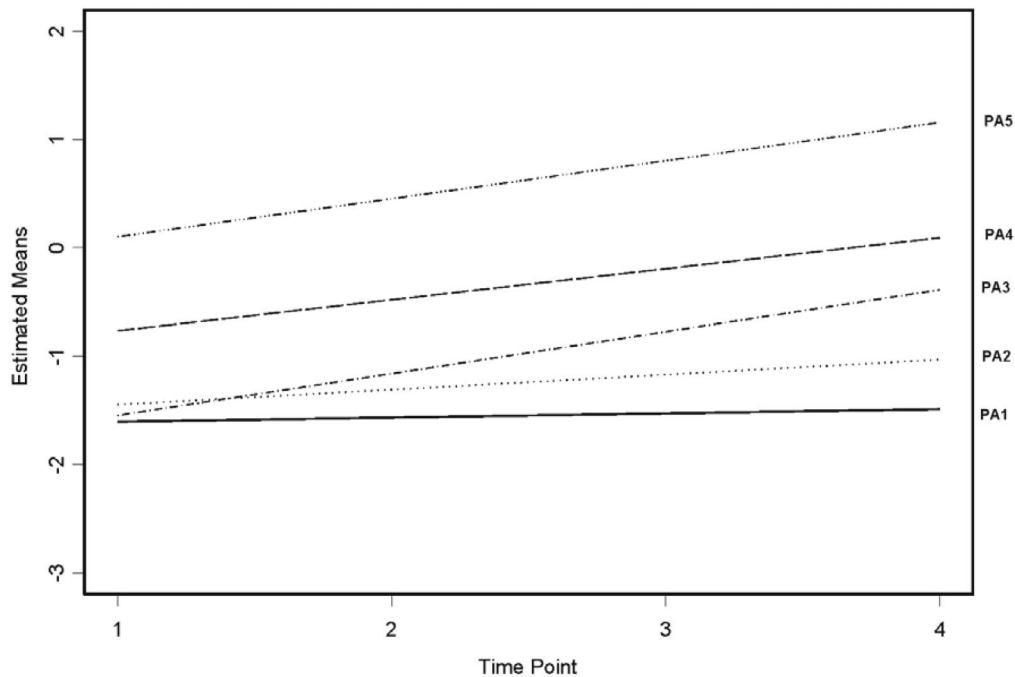


Figure 2. Estimated mean growth curves for the five-class model representing the phonological awareness (PA) in kindergarten.



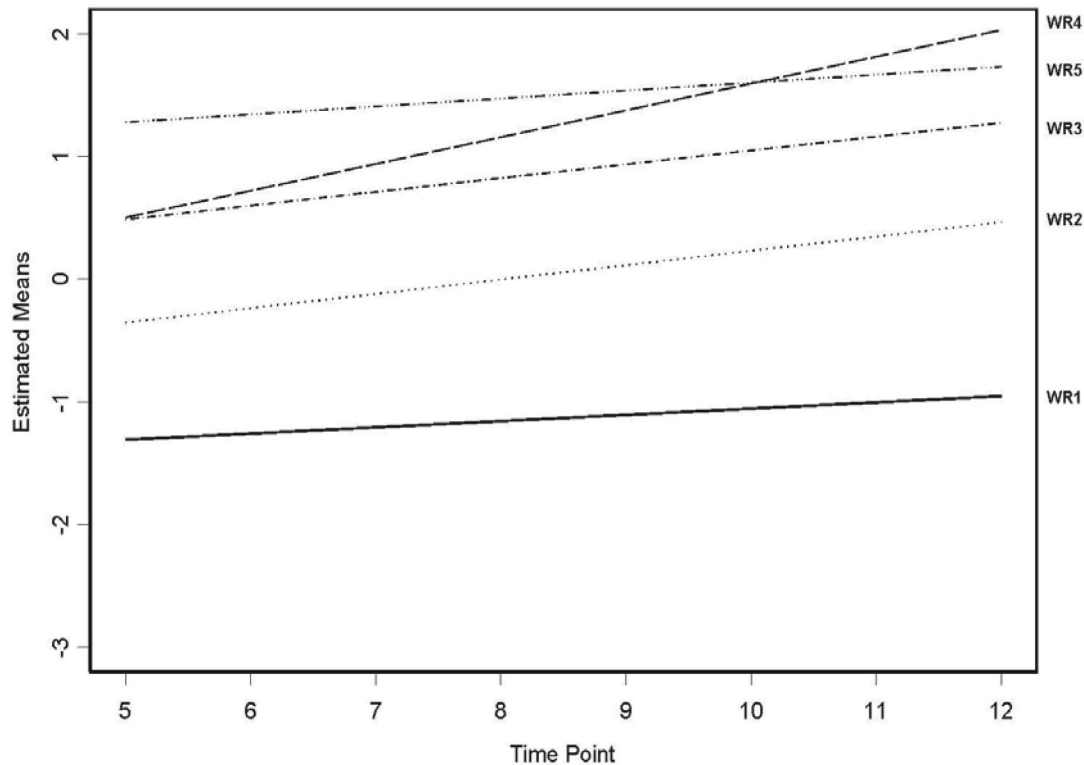


Figure 3. Estimated mean growth curves for the five-class model representing the word recognition (WR) development in first and second grades.

represent this transition from kindergarten to Grades 1 and 2, we allowed for transitional classes in the final model specification.

On the basis of class count, entropy, and model fit, the resulting 10-class model included several different transitional patterns in class membership from kindergarten to Grades 1 and 2. Table 9 describes the 10 different developmental profiles identified in the final model and the class counts based on their estimated class probabilities. The parameter estimates and their standard errors are shown in Table 10 and Figure 6 illustrates the estimated mean curves for phonological awareness and word recognition development. For example, in Figure 6, Class 1 represents the students with developmental profiles corresponding to PA 1 and WR 1, and Class 3 represents the students with developmental profiles corresponding to PA 2 and WR 3.

Table 6  
*BIC and Entropy Values for Grade 1 and 2 Models*

Model	BIC	Entropy
Two-class	2,260.08	0.76
Three-class	2,297.29	0.74
Four-class	2,273.16	0.73
Five-class	2,245.80	0.78
Six-class	2,236.49	0.79
Seven-class	2,233.74	0.82

Note. BIC = Bayesian information criterion.

As illustrated by Figure 6, in the final 10-class model, Class 1 (PA 1 transition into WR 1) students again represented the students who were at risk for reading difficulties because these students' trajectories remained flat. The Growth 1 estimates, which represented the slope for phonological awareness in the final model, ranged from 0.05 (representing flat trajectory) to 0.33 (significant growth). Similarly, for Growth 2, the estimates ranged from 0.03 to 0.27. As shown in Table 10, Class 1 students' reading development trajectories remained flat across all 3 years at 0.05 and 0.03 for Growth 1 and Growth 2, respectively. Class 2 students who, based on the slope and intercept estimates, started out initially at about the same level on phonological awareness skills as Class 1 students, continued to show growth on both phonological

Table 7  
*Intercept and Slope for Five Word Recognition Development Profiles*

Profile	Intercept <i>M (SE)</i>	Slope <i>M (SE)</i>
WR 1	-0.95 (0.10)	0.12 (0.02)
WR 2	0.46 (0.08)	0.27 (0.01)
WR 3	1.27 (0.08)	0.26 (0.01)
WR 4	2.03 (0.14)	0.51 (0.02)
WR 5	1.73 (0.07)	0.15 (0.01)

Note. WR = word recognition.

awareness (Growth 1 = 0.10) as well as word recognition (Growth 2 = 0.23) in contrast to Class 1 students, who showed essentially no growth (Growth 1 = 0.05 and Growth 2 = 0.03) during the same time period. As shown in Table 9, the groups on the diagonal represent the students who remained within the same class across the three grades. The students transitioning into different classes from their original classification in kindergarten are represented by Class 3 (PA 2 into WR 3), Class 4 (PA 3 into WR 2), Class 6 (PA 3 into WR 4), Class 7 (PA 4 into WR 3), and Class 9 (PA 4 into WR 5). The proportion of students that transitioned from one developmental profile to another is shown by the italicized entries in Table 9.

As Table 9 shows, there were only 8 students grouped into Class 4. This class comprised students who were originally classified as PA 3 in kindergarten and who subsequently moved into WR 2 in Grades 1 and 2, which were only a few students. The results also indicated that none of the students transitioned into or out of Class 1. This finding suggests that students in Class 1 were very homogeneous and were indeed the students who we considered to be most at risk for reading difficulties in later grades.

We also investigated the relationship between the growth parameters in kindergarten and the growth parameters in Grades 1 and 2. Only the relationship between word recognition growth and phonological awareness intercept was shown to be statistically significant. In other words, the rate of development in Grades 1 and 2 was directly related to the status of phonological awareness at the end of kindergarten.

#### Evaluation of Model Fit for the 10-Class Model

In order to determine how well this 10-class model fit the data, we plotted a random sample of the observed individual trajectories against the model-estimated mean trajectories. Each individual student was assigned to his or her respective class based on the student's weighted individual class probabilities. A random sample of individual trajectories (observed, not estimated) was plotted

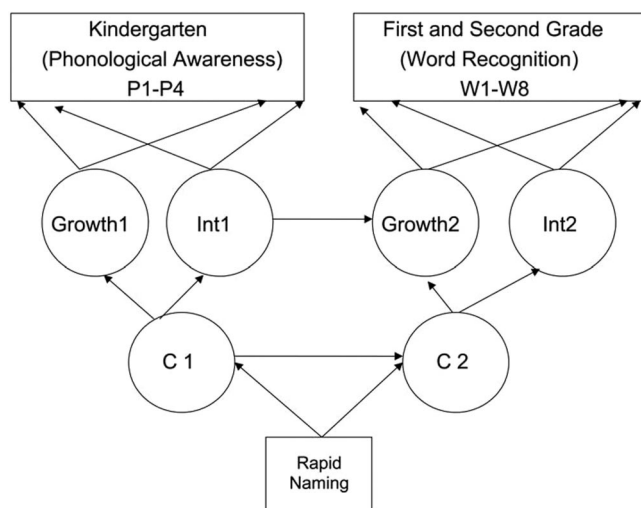


Figure 4. Path diagram for phonological awareness (P) and word recognition (W) combined. Int = intercept; C = class.

Table 8  
Five-Class Model Specification

Kindergarten	Grades 1 and 2				
	WR 1	WR 2	WR 3	WR 4	WR 5
PA 1	Class 1				
PA 2		Class 2			
PA 3			Class 3		
PA 4				Class 4	
PA 5					Class 5

Note. WR = word recognition; PA = phonological awareness.

against the estimated mean trajectories for each class for comparison. Except for signs of minor discrepancy in Class 5, all of the observed individual trajectories looked homogeneous around the estimated mean curves (see Figure 7). Class 5 represented the students who were classified as PA 3 in kindergarten and remained in WR 3 in Grades 1 and 2. Additionally, the entropy for the 10-class model was 0.76, which signified moderate to high clarity in the classification.

#### Differences in Rapid Naming

Performance on rapid naming, measured at the end of kindergarten, was included in the model as a covariate for class membership. As explained in the Method section, the significance of this covariate was examined by including a multinomial logistic regression component in the final model. The logistic regression plot shown in Figure 8 illustrates the differential effect of rapid naming on the probability distribution function of the 10 classes. Figure 8 shows that the probability of belonging to Class 1 was much higher when performance on rapid naming was low. Conversely, as illustrated in Figure 8, as the score on rapid naming increased, the probability of an individual belonging to Class 1 decreased significantly. This finding qualitatively differentiated students who entered school with comparably poor phonological awareness (e.g., PA 1 and PA 2 students) but progressed to very different outcomes. The results indicate that students who exhibited no significant improvement in phonological awareness in kindergarten were also the students with the lowest rapid naming skills at the end of kindergarten.

#### Differences in Ethnicity

To help characterize the Class 1 students, we examined the relationships among gender, ethnicity, SES, and class membership using the chi-square test. There was a statistically significant difference in the proportion of minority students in Class 1 compared to other classes. The proportion of minority students in Class 1 was 51% compared to only 31% in other classes ( $p = .01$ ). SES and gender were not statistically significant ( $ps = 1.0$  and  $.34$ , respectively).

#### Discussion

The purpose of this study was to introduce growth mixture modeling as a new approach to the identification of heterogeneous

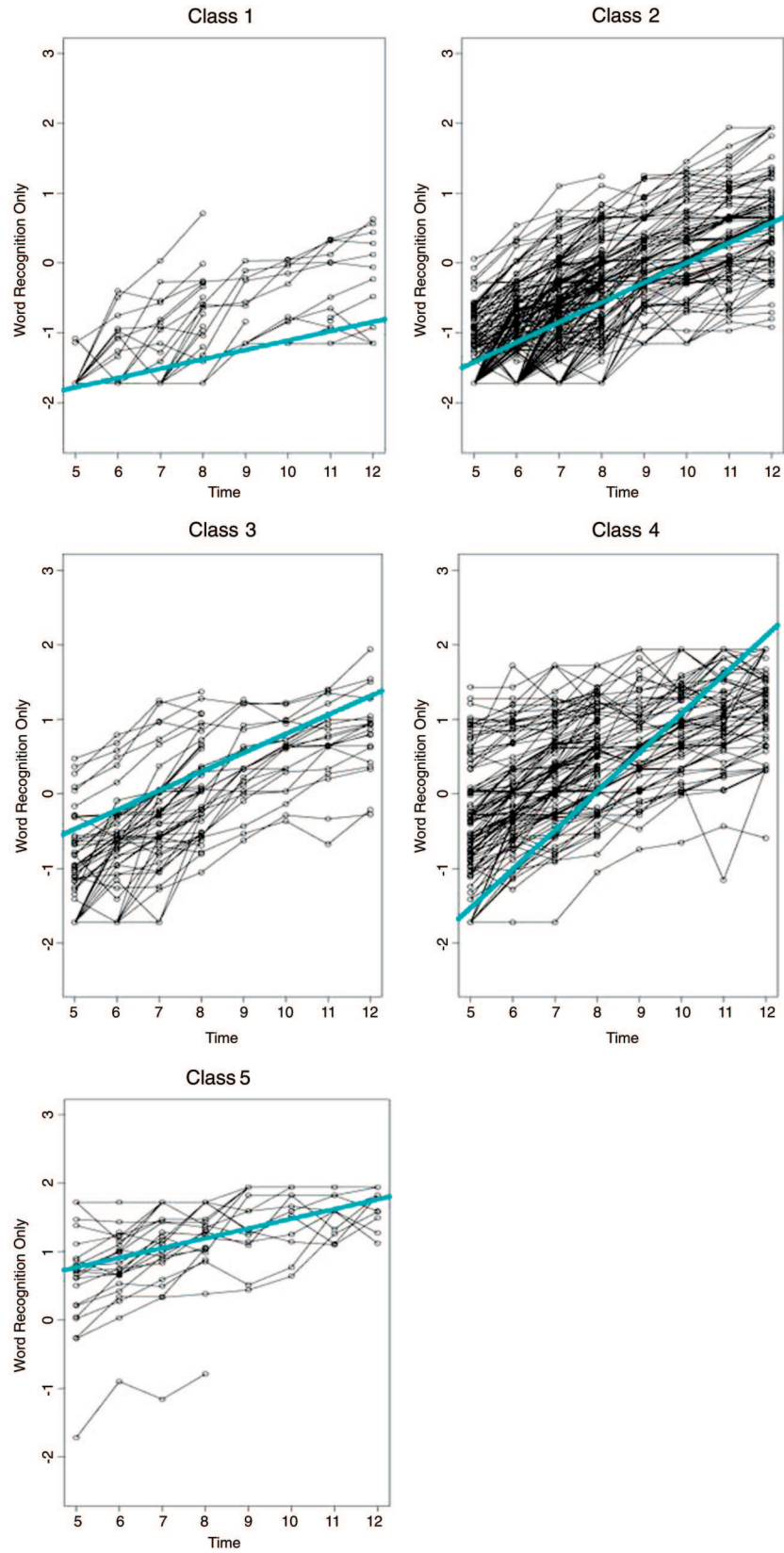


Figure 5. Word recognition development based on kindergarten classification.

Table 9  
*Class Specifications and Class Counts for the 10-Class Model*

Kindergarten	Grades 1 and 2				
	WR 1	WR 2	WR 3	WR 4	WR 5
PA 1	Class 1: 45 (11%)				
PA 2		Class 2: 63 (15%)	Class 3: 77 (19%)		
PA 3		Class 4: 8 (2%)	Class 5: 56 (14%)	Class 6: 36 (9%)	
PA 4			Class 7: 20 (5%)	Class 8: 8 (14%)	Class 9: 30 (7%)
PA 5					Class 10: 18 (4%)

*Note.* Values represent *n* and percentage for each class. The italicized entries represent the five classes that changed class membership between grades. WR = word recognition; PA = phonological awareness.

reading development profiles with data from a 3-year longitudinal study of reading precursors (e.g., phonological awareness) and outcomes (e.g., word recognition). Using this technique, we have identified a group of students with a distinct developmental pattern who are most at risk for reading difficulties. Although further studies are required to validate this group of students as potentially reading disabled, we have empirically identified a group of students with reading difficulties using a new approach.

Application of growth mixture modeling in this study highlights two important issues related to reading development research. First, as shown in previous research, this study empirically demonstrated the multidimensional continuity of the distribution of reading ability (Shaywitz et al., 1992). The results from this study indicate that developmental profiles identified in kindergarten are directly associated with reading development in Grades 1 and 2. The students identified as having difficulties acquiring phonological awareness skills in kindergarten exhibited slower developmental patterns in word recognition skills in subsequent years of the study. Specifically, although students in the lowest performing trajectory class were allowed to change membership with potential for improvement, in fact, nearly all of the students identified as the lowest performing group in kindergarten stayed in the same developmental trajectory throughout the 3 years. The use of growth mixture models to identify and classify students with reading difficulties minimizes anomalies and unfairness that are consequences of using an arbitrary cutoff for classification purposes. Using growth mixture models, researchers can circumvent the

problems associated with arbitrary classification of students as reading disabled.

Previous research has shown (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1994), and the results from this study support the notion, that reading difficulties are best characterized by deficits in prerequisite skills that lead to deficits in reading development, rather than by a lag in reading development. Identification of a group of students with persistent deficits over the 3-year period suggests that unless the students acquire the necessary prerequisite skills, they will continue to lag behind. This finding underscores the need for early identification and interventions specifically targeting deficit skills. Reconceptualizing the identification of reading difficulties using longitudinal measures stimulates further questions regarding the implications of early assessment practices and suggests possible directions for future research in this area. Although there has been considerable debate surrounding the validity of using the Dynamic Indicators of Basic Early Literacy Skills to monitor reading progress, more than 40 states in Reading First schools are now currently using the Dynamic Indicators of Basic Early Literacy Skills to screen Kindergarten–Grade 3 students for potential reading difficulties. Some states are also using the Phonological Awareness Literacy Screening Tests as an alternative tool. Given the lack of consensus on the most appropriate measures for monitoring reading progress, implementation of the proposed identification model will require further research in the area of assessment development. In addition, as researchers consider the practicality of implementing the reading progress monitoring model, the minimum data requirement for reliable classification must also be taken into consideration. As previous research has shown (Rogosa, 1988; Willett, 1988), it becomes difficult to obtain reliable estimates of the correlation between change and initial status with only two assessment data points due to measurement error in initial status.

Second, the findings suggest that the students with reading difficulties may in fact consist of various subgroups or subtypes, each with distinct developmental profiles manifesting from differences not only in outcomes but also possibly in etiology. As shown in the development of phonological awareness in kindergarten, although Class 1 and Class 3 students looked similar in terms of their initial status, rates of development differed significantly between these two groups of students, and, ultimately, this difference in the rate of development manifested a greater gap in students' reading development. Given the significant relationship

Table 10  
*Estimated Means and (Standard Errors) of Intercepts and Growth Parameters for the 10-Class Model*

Class	Intercept 1	Growth 1	Intercept 2	Growth 2
1	−1.50 (0.07)	0.05 (0.02)	−1.05 (0.09)	0.03 (0.04)
2	−1.15 (0.06)	0.10 (0.02)	0.25 (0.14)	0.23 (0.03)
3	−1.15 (0.06)	0.10 (0.02)	1.02 (0.09)	0.27 (0.02)
4	−0.60 (0.22)	0.29 (0.05)	0.25 (0.14)	0.23 (0.03)
5	−0.60 (0.22)	0.29 (0.05)	1.02 (0.09)	0.27 (0.02)
6	−0.60 (0.22)	0.29 (0.05)	1.52 (0.08)	0.26 (0.01)
7	0.15 (0.07)	0.30 (0.02)	1.02 (0.09)	0.27 (0.02)
8	0.15 (0.07)	0.30 (0.02)	1.52 (0.08)	0.26 (0.01)
9	0.15 (0.07)	0.30 (0.02)	1.82 (0.05)	0.17 (0.02)
10	1.09 (0.20)	0.33 (0.05)	1.82 (0.05)	0.17 (0.02)



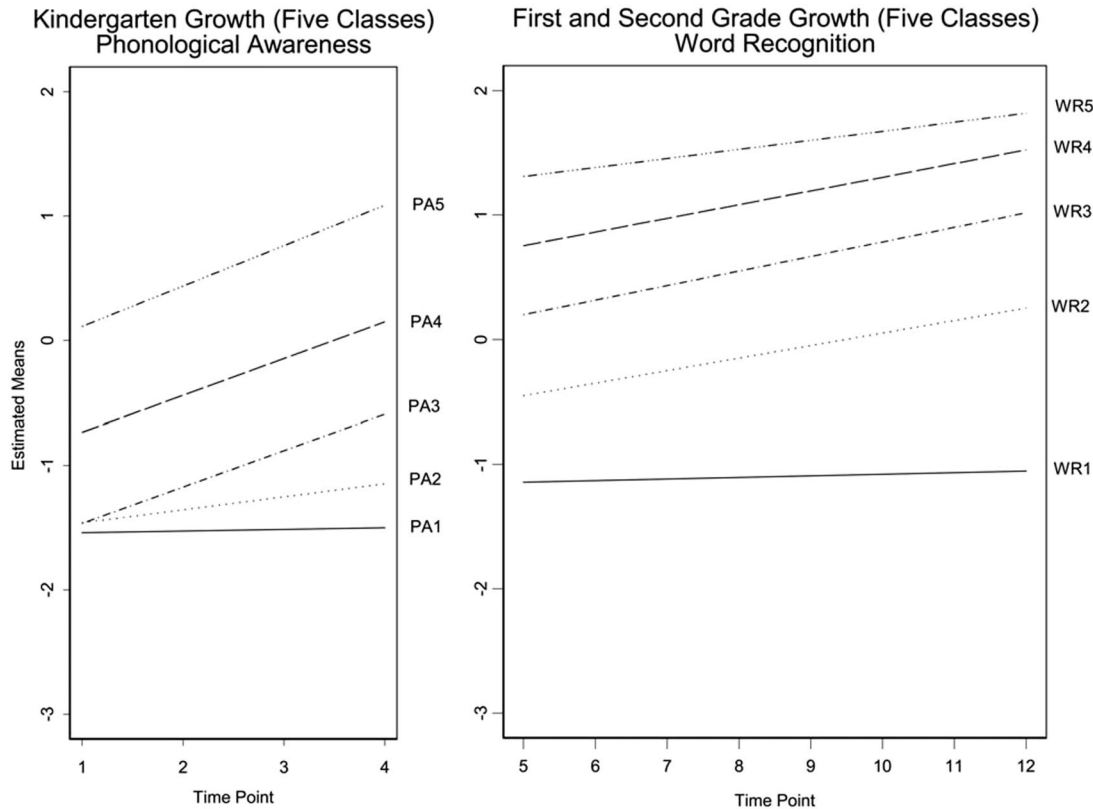


Figure 6. The 10-class model: estimated mean growth curves for phonological awareness (PA) and word recognition (WR).

between class membership and rapid naming, the results suggest that rapid naming and phonological skills are good predictors of subsequent reading development, as was previously shown in other studies (Petrill, Deater-Deckard, Thompson, DeThorne, & Schatschneider, 2006b; Torgesen, Wagner, Simmons, & Laughon, 1990; Walsh, Price, & Gillingham, 1988). Identification of students with poor developmental profiles in reading is a multivariate problem, and a univariate approach is inadequate for studying the complexity of the problem (Fletcher, Francis, Rourke, Shaywitz, & Shaywitz, 1992; Francis et al., 2005). Performance on the rapid naming test in kindergarten was a key indicator for differentiating the different reading development profiles in this study, but other relevant measures such as expressive vocabulary, verbal memory, and story recall should be explored and studied in detail. Although previous studies on reading disability subtypes have used multivariate clustering techniques to identify subtypes, these have been limited to single time point data (Morris et al., 1998). A growth mixture model may provide an alternative approach to identifying reading disability subtypes.

With different developmental profiles of reading, deficits in specific areas can be easily identified, and appropriate instructional strategies can target specific problem areas. As Berninger and Abbott (1994) stated, "The diagnosis of learning disability often is not tied to well-specified deficits clearly linked to instructional interventions" (pp. 166). Consequently, as a large-scale field study

conducted by Haynes and Jenkins (1986) revealed, reading instructional programs often are not linked to meet the needs of the characteristically different individual students. Identification of distinct developmental profiles may provide ways to differentiate treatment based on different treatment responses. With an effective identification system linked to appropriate remediation strategies, students will be able to receive supplementary instruction that is appropriately targeted for maximum benefit. As previous research has suggested (Fletcher, Francis, Morris, & Lyon, 2005; Torgesen et al., 2001), early identification is key to successful remediation programs. It is also important to recognize that, in accordance with the RTI model, identification is only the first stage in the intervention process. For successful remediation and prevention, early screening and appropriate intervention have to be followed up with progress monitoring.

This study also found that the percentage of minority students in Class 1 was higher than in other classes. This increased representation of minorities in Class 1 reflects the fact that minorities are at increased risk for reading problems. A report from the National Research Council (1998) suggested that "children from poor families and children of African American and Hispanic descent are at much greater risk of poor reading outcomes" (p. 27). Teasing apart the relationship of external factors and reading achievement is complicated by inadequate indicators of SES such as the self-report data that we used in our study. One of the limitations of this

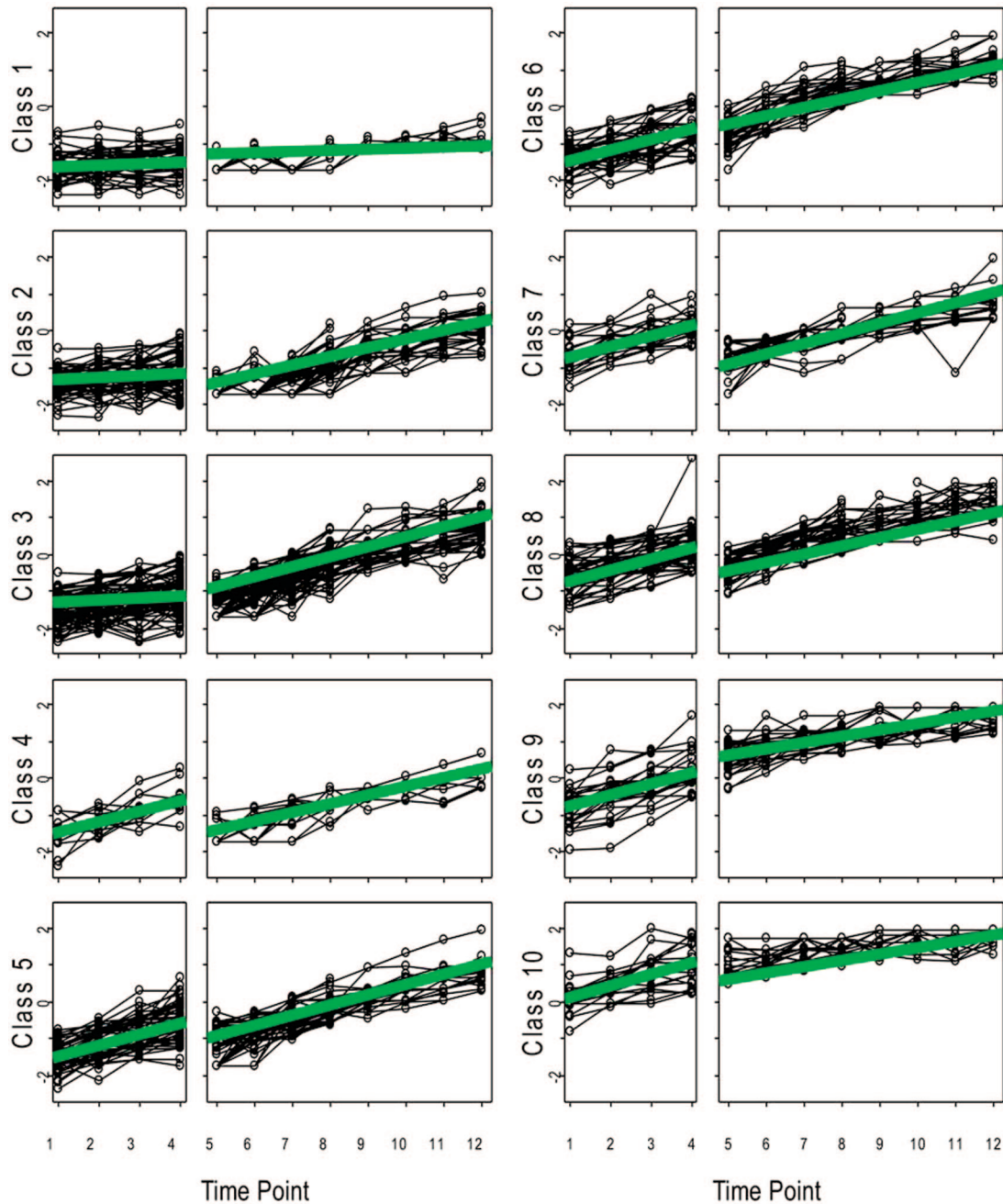


Figure 7. Observed individual growth trajectories with estimated mean growth trajectories for the 10-class model.

study is that, although the identification of Class 1 students was associated with minority status, given the limited information on student background, the characterization of students with potential reading failure was insufficient.

Another limitation of our findings is the relatively restricted measure of word recognition and the lack of a reading comprehension measure. We found that the correlation was quite high between our word recognition measure and the word attack

measure, as measured by the Word Attack subtest of the Woodcock–Johnson Psychoeducational Battery–Revised (Woodcock & Johnson, 1989). At the end of Grade 1, the correlation was .78 between our word reading recognition measure and Woodcock–Johnson Word Attack. At the end of Grade 2, the correlation was .71. Given the importance of reading comprehension skills in reading development, replication of this study with the inclusion of a comprehensive word reading

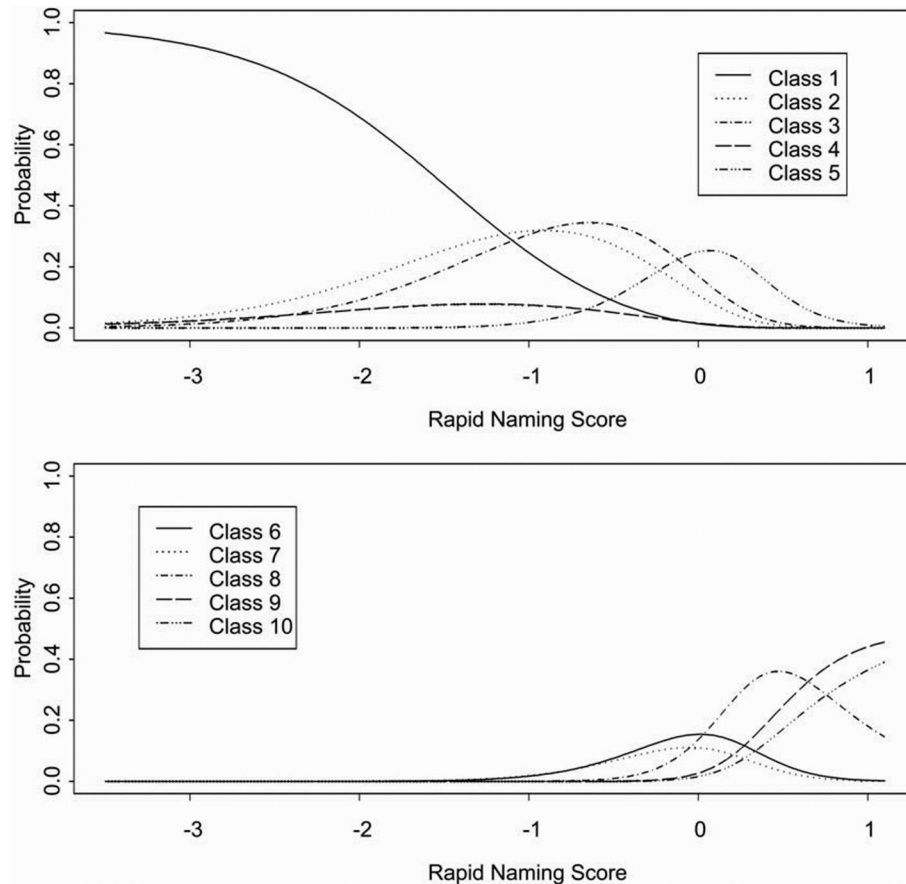


Figure 8. Multinomial logistic regression of class regressed on rapid naming score.

measure that includes nonsense words as well as reading comprehension is warranted.

Another important factor omitted in the analysis was school-level variability. Given there were only three schools in the study, systematic exploration of school effects on student reading outcomes was not possible with this current study sample. The fact that all three schools came from one single district with a common approach to reading instruction should minimize the potential school-level variability due to instruction.

Finally, given the recent findings in behavioral genetics studies (Byrne et al., 2002; Harlaar, Spinath, Dale, & Plomin, 2005; Petrill et al., 2006b), it will be important to consider environmental factors as well as genetic differences in early readers when examining differential treatment effects. Petrill et al. (2006b) found that environmental influences were substantial in explaining the individual variances in phonological awareness, however rapid naming seemed to be significantly influenced by genetic variance. For successful remediation, it is important for future studies to consider various individual and environmental factors to help characterize the different reading development profiles.

## References

- Berninger, V. W., & Abbott, R. D. (1994). Redefining learning disabilities: Moving beyond aptitude-achievement discrepancies to failure to respond to validated treatment protocols. In R. G. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 163–185). Baltimore: Brookes.
- Bradley, R., Danielson, R.L., & Hallahan, D.P. (2002). *Identification of learning disabilities: Research to practice*. Mahwah, NJ: Erlbaum.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models*. Newbury Park, CA: Sage.
- Byrne, B., Delaland, C., Fielding-Barnsley, R., Quain, P., Samuelsson, S., Høien, T., et al. (2002). Longitudinal twin study of early reading development in three countries: Preliminary results. *Annals of Dyslexia*, 52, 49–74.
- Carroll, J., Davies, P., & Richman, B. (1971). *The American heritage word frequency book*. Boston: Houghton Mifflin.
- Catts, H. W., & Kamhi, A. G. (1999). *Language and reading disabilities*. Needham Heights, MA: Allyn & Bacon.
- Compton, D. L., Fuchs, D., Fuchs, L. S., & Bryant, J. D. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology*, 98, 394–409.
- De Hirsh, K., Jansky, J., & Langford, W. (1966). *Predicting reading failure*. New York: Harper & Row.
- Denckla, M., & Rudel, R. (1976). Rapid automatized naming (RAN): Dyslexia differentiated from other learning disabilities. *Neuropsychologia*, 14, 471–479.
- Fletcher, J. M., Coulter, W. A., Reschly, D. J., & Vaughn, S. (2004). Alternative approaches to the definition and identification of learning

- disabilities: Some questions and answers. *Annals of Dyslexia*, 54, 304–331.
- Fletcher, J. M., Francis, D. J., Morris, R. D., & Lyon, G. R. (2005). Evidence-based assessment of learning disabilities in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, 34, 506–522.
- Fletcher, J. M., Francis, D. J., Rourke, B. P., Shaywitz, S. E., & Shaywitz, B. A. (1992). The validity of discrepancy-based definitions of reading disabilities. *Journal of Learning Disabilities*, 25, 555–561.
- Fletcher, J. M., & Satz, P. (1984). Test-based versus teacher-based predictions of academic achievement: A three-year longitudinal follow-up. *Journal of Pediatric Psychology*, 9, 193–203.
- Francis, D. J., Fletcher, J. M., Shaywitz, B. A., Shaywitz, S. E., & Rourke, B. (1996). Defining learning and language disabilities: Conceptual and psychometric issues with the use of IQ tests. *Language, Speech, and Hearing Services in Schools*, 27, 132–143.
- Francis, D. J., Fletcher, J. M., Stuebing, K. K., Lyon, G. R., Shaywitz, B. A., & Shaywitz, S. E. (2005). Psychometric approaches to the identification of LD: IQ and achievement scores are not enough. *Journal of Learning Disabilities*, 38, 98–108.
- Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1994). The measurement of change: Assessing behavior over time and within a developmental context. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 29–58). Baltimore: Brookes.
- Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology*, 88, 3–17.
- Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research and Practice*, 18, 157–171.
- Gunning, T. G. (1998). *Assessing and correcting reading and writing difficulties*. Boston: Allyn & Bacon.
- Harlaar, N., Spinath, F. M., Dale, P. S., & Plomin, R. (2005). Genetic influences on word recognition abilities and disabilities: A study of 7 year old twins. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 46, 373–384.
- Haynes, M., & Jenkins, J. (1986). Reading instruction in special education resource rooms. *American Educational Research Journal*, 23, 161–190.
- Hollingshead, A. B. (1975). *Four-Factor Index of Social Status*. New Haven, CT: Yale University Press.
- Jenkins, J. R., & O'Connor, R. E. (2002). Early identification and intervention for young children with reading/learning disabilities. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 99–149). Mahwah, NJ: Erlbaum.
- Jennrich, R. I., & Schluchter, M. D. (1986). Unbalanced repeated measures models with structured covariance matrices. *Biometrics*, 42, 805–820.
- Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, 78, 243–255.
- Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 38, 963–974.
- Liberman, I. Y., Shankweiler, D., & Liberman, A. M. (1989). The alphabetic principle and learning to read. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 1–33). Ann Arbor: University of Michigan Press.
- Lindamood, C. H., & Lindamood, P. C. (1979). *Lindamood Auditory Conceptualization Test*. Austin, TX: PRO-ED.
- Lindstrom, M. J., & Bates, D. M. (1988). Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. *Journal of the American Statistical Association*, 83, 1014–1022.
- Lyon, G. R. (1998). *Overview of reading and literacy initiatives*. Washington, DC: National Institute of Child Health and Human Development.
- Lyon, G. R., Fletcher, J. M., Shaywitz, S. E., Shaywitz, B. A., Torgesen, J. K., & Wood, F. B. (2001). Rethinking learning disabilities. In C. E. Finn, Jr., R. A. J. Rotherham, & C. R. Hokanson, Jr. (Eds.), *Rethinking special education for a new century* (pp. 259–287). Washington, DC: Thomas B. Fordham Foundation and Progressive Policy Institute.
- McLachlan, G., & Peel, D. (2000). *Finite mixture models*. New York: Wiley Interscience.
- Morris, R. D., Steubing, K. K., Fletcher, J. M., Shaywitz, S. E., Lyon, G. R., Shankweiler, D. P., et al. (1998). Subtypes of reading disability: Variability around a phonological core. *Journal of Educational Psychology*, 90, 347–373.
- Muthén, B. (2000). Latent variable mixture modeling. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: New developments and techniques* (pp. 1–33). Mahwah, NJ: Erlbaum.
- Muthén, B. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In L. M. Collins & A. Sayer (Eds.), *New methods for the analysis of change* (pp. 289–322). Washington, DC: American Psychological Association.
- Muthén, B., Khoo, S. T., Francis, D., & Boscardin, C. (2002). Analysis of reading skills development from kindergarten through first grade: An application of growth mixture modeling to sequential processes. In S. R. 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.
- Nagin, D. S. (1999). Analyzing developmental trajectories: A semi-parametric, group based approach. *Psychological Methods*, 4, 139–157.
- National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute of Child Health and Human Development.
- National Research Council. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Parrilla, R. K., Kirby, J. R., & McQuarrie, L. (2004). Articulation rate, naming speed, verbal short-term memory, and phonological awareness: Longitudinal predictors of early reading development? *Scientific Studies of Reading*, 8, 3–26.
- Petrill, S. A., Deater-Deckard, K., Thompson, L. A., DeThorne, L. S., & Schatschneider, C. (2006a). Genetic and environmental effects of serial naming and phonological awareness on early reading outcomes. *Journal of Educational Psychology*, 98, 112–121.
- Petrill, S. A., Deater-Deckard, K., Thompson, L. A., DeThorne, L. S., & Schatschneider, C. (2006b). Reading skills in early readers: Genetic and shared environmental influences. *Journal of Learning Disabilities*, 39, 48–55.
- Rogosa, D. R. (1988). Myths about longitudinal research. In K. W. Schaie, R. T. Campbell, W. Meredith, & S. C. Rawlings (Eds.), *Methodological issues in aging research* (pp. 171–210). New York: Springer-Verlag.
- Satz, P., & Fletcher, J. M. (1988). Early identification of learning disabled children: An old problem revisited. *Journal of Consulting and Clinical Psychology*, 56, 824–829.
- Scarborough, H. S. (1998). Predicting the future achievement of second graders with reading disabilities: Contributions of phonemic awareness,



- verbal memory, rapid serial naming, and IQ. *Annals of Dyslexia*, 48, 115–136.
- Schatschneider, C., Carlson, C. D., Francis, D. J., Foorman, B. R., & Fletcher, J. M. (2002). Relationship of rapid automatized naming and phonological awareness in early reading development: Implications for the double-deficit hypotheses. *Journal of Learning Disabilities*, 35, 245–256.
- Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology*, 96, 265–282.
- Schatschneider, C., Francis, D. J., Foorman, B. R., & Fletcher, J. M. (1999). The dimensionality of phonological awareness: An application of item response theory. *Journal of Educational Psychology*, 91, 439–449.
- Schenck, B., Fitzsimmons, J., Bullard, P. C., Taylor, H. G., & Satz, P. (1980). A prevention model for children at risk for reading failure. In R. M. Knights & D. J. Bakker (Eds.), *Treatment of hyperactive and learning disordered children* (pp. 31–48). Baltimore: University Park Press.
- Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Distribution and temporal stability of dyslexia in an epidemiological sample of 414 children followed longitudinally. *New England Journal of Medicine*, 326, 145–150.
- Slavin, R. E. (1994). *Preventing early school failure: Research, policy and practice*. Needham Heights, MA: Allyn & Bacon.
- Speece, D. L. (2005). Hitting the moving target known as reading development: Some thoughts on screening children for secondary interventions. *Journal of Learning Disabilities*, 38, 487–493.
- Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology*, 86, 24–53.
- Strag, G. A. (1972). Comparative behavioral ratings of parents with severe mentally retarded, special learning disability, and normal children. *Journal of Learning Disabilities*, 5, 52–56.
- Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K. S., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities*, 34, 33–58.
- Torgesen, J. K., & Wagner, R. K. (2002). Predicting reading ability. *Journal of School Psychology*, 40, 1–26.
- Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Burgess, S., & Hecht, S. (1997). Contributions of phonological awareness and rapid automatic naming ability to the growth of word-reading skills in second to fifth-grade children. *Scientific Studies of Reading*, 1, 161–185.
- Torgesen, J. K., Wagner, R. K., Simmons, K., & Laughon, P. (1990). Identifying phonological coding problems in disabled readers: Naming, counting, or span measures? *Learning Disabilities Quarterly*, 13, 236–243.
- Verbeke, G., & LeSaffre, E. (1996). A linear mixed-effects model with heterogeneity in the random effects population. *Journal of the American Statistical Association*, 91, 217–221.
- Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive Test of Phonological Processes*. Austin, TX: PRO-ED.
- Walsh, D. J., Price, G. G., & Gillingham, M. G. (1988). The critical but transitory importance of letter naming. *Reading Research Quarterly*, 23, 108–122.
- Wiesner, M., & Windle, M. (2004). Assessing covariates of adolescent delinquency trajectories: A latent growth mixture modeling approach. *Journal of Youth and Adolescence*, 22, 431–442.
- Willett, J. B. (1988). Questions and answers in the measurement of change. *Review of Research in Education*, 15, 345–422.
- Wolf, M., & Bowers, P. G. (2000). Naming-speed deficits in developmental reading disabilities: An introduction to the special series on the double-deficit hypothesis. *Journal of Learning Disabilities*, 33, 322–333.
- Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock-Johnson Psychoeducational Battery-Revised*. Allen, TX: DLM Teaching Resources.

Received February 2, 2006

Revision received June 19, 2007

Accepted July 3, 2007 ■