Early Identification of Reading Difficulties Using Heterogeneous Developmental Trajectories

Christy K. Boscardin and Bengt O. Muthén

University of California, Los Angeles

David J. Francis

University of Houston

Eva L. Baker

University of California, Los Angeles

Abstract

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. 411 children in kindergarten through 2nd grade were administered measures of phonological awareness, word recognition, and rapid naming skills. The application of growth mixture models provides a more dynamic view of the learning process and the correlates that affect the rate of reading development. Growth mixture modeling was used to examine the presence of heterogeneous developmental patterns and served to identify one group of students with distinct developmental patterns who are most at risk for reading difficulties. The results indicate that precursor reading skills such as phonological awareness and rapid naming are highly predictive of later reading development 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 reading skills in subsequent years of the study.

Key words: Reading Development, Screening, Reading Skills, Achievement, Longitudinal Studies, Models

Early Identification of Reading Difficulties Using Heterogeneous Developmental Trajectories

Introduction

With mounting evidence for early detection as the key to remediation and prevention for later reading difficulties, investigation of effective strategies for earlier identification of students with reading difficulties seems to be the most imperative task at hand. According to research funded by the National Institute of Child Health and Human Development, if intervention is delayed until nine years of age (the time when most children with reading difficulties receive services), approximately 75% of children will continue to have difficulties learning to read throughout high school (Lyon, 1998). Satz and Fletcher (1988) have suggested that interventions are most effective if applied prior to the overt manifestation of disability. Schenck, Fitzimmons, Bullard, Taylor, and Satz (1980) also concluded that children at high risk who received intervention during kindergarten 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 recognized at an early age (de Hirsh, Jansky, & Langford, 1966; Strag, 1972). Additionally, according to Lyon (1998?), prevention and early intervention programs that combine good instruction in phonological awareness, phonics, fluency development, and reading comprehension strategies can increase reading skills in most children.

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

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, Fletcher, Shaywitz, Shaywitz, Torgesen, & Wood, 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 (Gunning, 1998; Francis, Fletcher, Shaywitz, Shaywitz, & Rourke, 1996a). Although the discrepancy-based method was the most widely used definition of reading difficulty, there were several conceptual and measurement problems that warranted an alternative method of identification of dyslexics and other poor readers (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996b; Shaywitz, Escobar, Shaywitz, Fletcher, & Makuch, 1992). Dissatisfaction with previous approaches to identification of children with reading difficulties have led to consideration of alternatives such as Response to Intervention (RTI) in the recent reauthorization of the Individuals with Disabilities Education Act (IDEA). The core concept of RTI is based on the idea that remedial intervention with continuous monitoring of student progress will contribute to the availability of data useful for identification. Based on the RTI paradigm, insufficient RTI will require special education and related services.

In light of a new paradigm shift in identification and reading disability, in this study, we offer a new approach to examining reading development through the application of a new longitudinal clustering technique called growth mixture modeling. The purpose of this study is 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. With the wider acceptance of emergent literacy theory, the detection of reading problems using static assessments is now being gradually replaced by the analysis of the individual growth of prerequisite reading skills as the proxy for learning development. Analysis of individual growth curves based on precursor skills can be used to form the theoretical basis for the classification of students with reading difficulties.

The analysis of individual developmental trajectories using precursor reading skills provides an alternative method of diagnosing and identifying students at risk for reading failure, when intervention efforts can be most effective. These individual growth curves not only yield earlier diagnosis of reading difficulties, but also provide contextual data for the further analysis of individual problems and differences (Francis et al., 1996b). Considering that growth and development are fundamental to the concept of learning, it seems only logical to consider longitudinal data as the primary source for the identification of learning 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, individual outcomes are represented using rates of growth over multiple time points. Using 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 has been introduced (Muthén, 2000; Muthén & Muthén, 2000; Muthén et al., 2002). This new technique, called growth mixture modeling, uses earlier developmental patterns to provide more reliable predictions for later development by taking into account the heterogeneity of student development. Muthén et al. (2002) report that growth mixture modeling allows for the identification of several different reading developmental profiles. Since 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. 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). In the current study, we will use growth mixture modeling to determine whether distinctive groups of students with various developmental profiles can be identified.

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 (Fletcher, Francis, Carlson, & Foorman, 2004; Schatschneider, Carlson, Francis, Foorman, & Fletcher, 2002; 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 not excluded. Measurements on precursor 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 in 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 have 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 some students from being followed into grade 2, and one school's decision not to participate in year 4. Consequently, as shown in Table 1, sample sizes in cohorts 4 and 5 are substantially smaller 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 ethnicities. The mean age for the kindergartners was 5.8 years (SD = 0.35), 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 are considered to be non-minority, while black, Hispanic, and other students are categorized as minority. The Hollingshead (1975) Four Factor Index of Social Status was used to obtain the following socioeconomic status (SES) breakdown: 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 this data.

Measures

A skills assessment battery including phonological awareness, rapid naming of letters, and word recognition was administered to each child in October, December, February, and April between kindergarten and grade 2. 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 non-words.

Phoneme segmentation. Subjects were instructed to listen to words and were asked to “tell me 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 (i.e., "say the word 'cup.' Now tell me what word would be left it 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 3 items were three- to five-phoneme two-syllable words for which the consonant to be deleted was in the middle (i.e., ti[g]er).

Sound categorization. The sound categorization task asked a child to select one of four words that did not share a phoneme with the rest (i.e., "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 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 asked a 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) have 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 is 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 non-words. This task is identical to blending phonemes into words except that nonwords are used in place of read 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]) (Schatschneider et al., 1999).

Each of the seven tasks consisted of 15 items scored as either correct or incorrect. Test-retest reliabilities for the seven tasks are (.85 to .90), (.86 to .92), and (.82 to .93) for kindergarten, grade 1, and grade 2, respectively (Carlson, 1998). For the analysis, the scores on these seven phonological awareness tasks were combined into one latent ability score. Instead of using raw total scores of phonological ability, Item Response Theory (IRT) scores based on estimates of each person's latent phonological trait were used to represent phonological awareness (Schatschneider et al., 2004).

Word Recognition. Skills in word recognition were assessed in Grades 1 and 2 by asking students to read aloud 50 words on 4x6-inch 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 matched for frequency of occurrence (Carroll, Davies, & Richman, 1971) and spanned first- through third-grade levels of difficulty. Instead of using raw total scores of phonological ability, IRT scores based on estimates of each person's latent trait on word recognition skills were used as an indicator for word recognition ability. The internal consistency estimate for the word list was .90. Concurrent and predictive validity for the word list were high, as evidenced by 0.80 correlations with the

Letter Word and Word Attack subtests of the Woodcock-Johnson Psycho-educational Battery – Revised (WJ-R; Woodcock & Johnson, 1989).

Rapid Naming. Denckla and Rudel's (1976) Rapid Automatized Naming Tests for Letters (RAN) was assessed in kindergarten. RAN letters were high-frequency lower-case 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 seconds 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). The scores obtained only from April data collection were log-transformed and included in the analysis[1].

Analysis

General growth mixture modeling (GGMM) introduced by Muthén and Shedden (1999), Muthén (2000), Muthén and Muthén (2000), Muthén and Muthén (2001), Muthén (2001), and Muthén et al. (2002), provides technical advantages over conventional growth models by allowing greater flexibility in model specifications and assumptions. The importance of growth modeling with mixture components has been explored by other researchers (Verbeke & LeSaffre, 1996; Nagin, 1999), however, recent work introduced by Muthén (2000), and Muthén and Shedden (1999) provides a much more flexible framework than previous models.

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 there may be two different groups of students with very different reading developmental trajectories, by using conventional growth modeling it would be difficult to detect the misspecification of the model since 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.

Using growth mixture models, the different subgroups in the model can be conceptualized as classifying the individuals into different collections of reading developmental 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 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. To test these hypotheses and to include the heterogeneity of subgroups in the model, a more general growth mixture model is now introduced.

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, 2001; Muthén, 2000; Muthén & Shedden, 1999). This technique goes beyond conventional growth modeling with random coefficients by using categorical latent variables to represent heterogeneity with respect to the growth shapes, concurrent outcomes, and later outcomes. Using categorical latent variables, for example, individuals with different reading developmental trajectories can be placed into different latent classes. These latent classes can represent multiple groups with different developmental trajectories where group membership is unknown but can be inferred from the data. The classes or heterogeneous subgroups are represented by a latent variable since the class memberships are unknown. 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:

$$Y_{ik} = \nu_k + \Lambda_k\, \eta_{ik} + K_k\, x_{ik} + \varepsilon_{ik}, \qquad (3)$$

and

$$\eta_{ik} = \alpha_k + B_k\, \eta_{ik} + \Gamma_k\, x_{ik} + \zeta_{ik}\,, \qquad (4)$$

Here, Y represents the repeated measures over fixed time points. The η_{ik} are random effects and Λ_k represent time scores for the shape of the growth curves. K_k represents the effects of time-varying covariates, and Γ_k represents the effects of time-invariant covariates. α_k represents the intercepts for η for latent class *k*. For example, current data α_k represents the different reading developmental trajectories for the different classes. The residual vectors, ε_{ik} and $\zeta_{ik,}$ are assumed to have covariance matrices Θ_k and Ψ_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 (2001), Muthén and Muthén (2000), and Muthén and Shedden (1999), is much more wide-ranging than the mixture models proposed by Nagin (1999), where it is assumed $\Psi_k = 0$ and $\Theta_{k} = \theta\, I$. In this case, the model is considered a fixed-effects model for each class, whereas the model introduced by Muthén, Khoo, Francis, and Boscardin (2002) is a more general and flexible form of growth mixture model. A model proposed by Muthén (2000) provides more flexibility by allowing class to class variation of the covariance matrices Ψ_k and Θ_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:

$$p_{ik} = P\,(c_{ik} = 1|\, y_i, x_i) \propto P\,(c_{ik} = 1|\, x_i)\, f\,(Y_{ik}\,|\, x_i) \qquad (5)$$

In this study, the individual students are assigned to a class based on their highest estimated posterior probabilities. The posterior probabilities are computed for a given individual observation vector (y_i, x_i). In other words, for a given model, individual students' posterior probabilities are 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 and Shedden (1999) 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:

$$P\,(c_{ik} = 1|x_i) = \exp(\beta_{0k} + \beta_{1k}x_i)/\sum_{c=1}^{K} \exp(\beta_{0c} + \beta_{1c}xi) \qquad (3)$$

for $k = 1, \ldots, K$ 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, β_{0K} is the class intercept, and β_{1K} is the regression coefficient for the kth class on xi, 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. On the other hand, 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 their developmental trajectories in kindergarten may differentially influence their later reading development. Growth mixture modeling allows for the examination of all these questions systematically without the restrictions of conventional growth modeling.

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) suggest 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 (McLachlan & Peel, 2000; Kass & Raftery, 1995). For a given model, BIC is calculated as follows:

$$\text{BIC} = -2 \log L + r \log n. \qquad (4)$$

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, BIC values across the different models will be compared, 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, since the BIC criterion 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 probabilities for belonging to all the other classes. These probabilities determine the most likely class for each student. For example, a student's estimated probabilities may be 0.80 for Class 1, 0.15 for Class 2, 0.05 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, the probabilities can be summarized into average posterior probabilities. For example, if the average posterior probability of Class 1 is 0.90, 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:

$$E_K = 1 - (\Sigma_i \Sigma_k(-p_{ik} \log(p_{ik}))/ n \log(K)) \qquad (5)$$

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, since 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, 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 which combined all three years of data. Although the developmental trajectories identified in the separate analyses 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 match 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 (Banden-Roche et al., 1997, cited in Muthén, 2000). Indication of a good model fit requires close alignment between the individual trajectories and the estimated mean trajectory.

Results

Initial Model Selection Using Two Separate Analyses

Models with between two and six classes were fit to the longitudinal phonological awareness data from the kindergarten portion of the data set. Table 3 presents the BIC values for these models and suggests that a four-class model gives the best fit. However, as the difference in the entropy values between the four-class and the five-class model was minimal, and the class proportions for the four-class model lack distinction between the students (Table 4), 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 performing group. PA 1 students started out initially with low performance in phonological awareness and exhibited no significant improvement throughout the entire kindergarten year. This group of students would be expected to be most at risk for developing reading difficulties in later grades. The estimated means and the standard errors for the five-class model are presented in Table 5.

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, the five-class model was found to be the best in terms of fit and interpretability. As shown in Table 7, WR 1 had the lowest intercept and slope. Students in WR 1 are thus characterized as the students who are at risk for reading difficulties. The estimated mean growth curves for the five-class model are shown in Figure 3.

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 three years. As an initial step in the analysis, we considered the model in which all students stayed in the same developmental trajectory throughout all three years. As shown in Table 8, PA 1 students corresponded to WR 1 growth trajectory, and PA 2 students were linked to WR 2 growth trajectory, and so on. 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 progress to a WR 3 profile based on word recognition development. By allowing change in class membership, one can determine which students stayed in the same developmental trajectories throughout three years and which students changed their class membership. The assumption is that if students do stay within the same class throughout the three grades, then the kindergarten classification should correspond directly to grade 1 and grade 2 classification. However, as illustrated by Figures 4, 5, and 6, the overlap between the kindergarten classification and grades 1 and 2 classification was minimal. For example, as Figure 4 illustrates, the estimated mean curves based on the WR 1 and WR2 do not fit the observed word recognition developmental trajectories of students who are classified into PA 1 and PA 2. Subsequently, we determined that not all PA 1 students correspond directly to WR 1 students. Instead, as expected, some students do move into other developmental trajectories in grades 1 and 2.

In the combined analysis, as shown by the path diagram in Figure 7, 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 rapid naming can be considered an etiologically distinct source of variance in reading outcomes (Petrill, Deater-Deckard, Thompson, DeThorne, & Schatschneider, 2006a). In addition to the inclusion of rapid naming, the growth parameters in grades 1 and 2 (word recognition intercept and word recognition growth) were regressed on the kindergarten phonological awareness intercept. 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. On the basis of class count, entropy, and model fit, the resulting ten-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 8 illustrates the estimated mean curves for phonological awareness and word recognition development.

In this final ten-class model, class 1 (PA 1 transition into WR 1) students again represent the students who are at risk for reading difficulties. Class 1 students' reading development trajectories stayed flat across all three years. The groups on the diagonal represent the students who stay 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 transition 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 is comprised of students who were originally classified as PA 3 in kindergarten who subsequently moved into WR 2 in grades 1 and 2, which were only a few students. The results also indicate 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.

Evaluation of Model Fit for the Ten-Class Model

In order to determine how well this ten-class model fits the data, a random sample of the observed individual trajectories was plotted against the model-estimated mean trajectories. Each individual student is assigned to his or her respective class based on their weighted individual class probabilities. A random sample of individual trajectories (observed, not estimated) was plotted 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 look homogeneous around the estimated mean curves (see Figures 9, 10, 11, 12, 13). Class 5 represents the students who are classified as PA 3 in kindergarten and remained in WR 3 in grades 1 and 2. Additionally, the entropy for the ten-class model was 0.76.

Differences in Ethnicity

To help characterize the Class 1 students, the relationships among gender, ethnicity, SES, and class membership were examined 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 ($p = 0.01$). SES and gender were not statistically significant.

Differences in Rapid Naming

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 kindergarten. Performance on rapid naming measured at the end of kindergarten was included in the model as a covariate for class membership. The significance of this covariate was examined using multinomial logistic regression. The logistic regression plot shown in Figure 14 illustrates the differential effect of rapid naming on the probability distribution function of the 10 classes. Figure 14 shows that the probability of belonging to Class 1 is much higher when performance on rapid naming is low. Conversely, as illustrated in Figure 14, as the score on the rapid naming increases the probability of an individual belonging into Class 1 decreases significantly. This finding qualitatively differentiates students who entered school with comparably poor phonemic awareness skills, but progressed to very different outcomes. Performance on rapid naming test seemed to be an important predictor of reading outcomes.

Discussion

The purpose of this study was to introduce growth mixture modeling as a new approach to the identification of heterogeneous reading developmental profiles with data from a three-year longitudinal study of reading precursors and outcomes. 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, students identified as the lowest-performing group in kindergarten stayed in the same developmental trajectory throughout the three 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, we can circumvent the problems associated with arbitrary classification of students as reading disabled.

As previous research has shown (Francis et al., 1994), 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 three-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. Re-conceptualizing 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 (DIBELS) to monitor reading progress, more than 40 states in Reading First schools are now currently using DIBELS to screen K-3 students for potential reading difficulties. Some states are also using the Phonological Awareness Literacy Screening Tests (PALS) 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 we consider the practicality of implementing the reading progress monitoring model, the minimum data requirement for reliable classification must also be taken into consideration. With only two assessment data points, it becomes difficult to obtain reliable estimates of the correlation between change and initial status 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 between class membership and rapid naming, the results suggest that rapid naming and phonological skills are good predictors of subsequent reading development, as previously shown in other studies (Petrill, Deater-Deckard, Thompson, DeThorne, & Schatschneider, 2006; Torgesen et al., 1990; Walsh, Price, & Gillingham, 1988). Identification of students with poor developmental profiles in reading is a multivariate problem and using the univariate approach is inadequate for studying the complexity of the problem. Performance on the rapid naming test in kindergarten was a key indicator for differentiating the different reading developmental profiles in this study, but other relevant measures 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-timepoint data. A growth mixture model provides an analytic tool capable of addressing identification and characterization of reading disability subtypes with longitudinal data. Although the results presented here may be considered merely an elaboration on the utility of underlying psychometric models, they do provide a basis for further discussion on the validity of IQ-achievement discrepancies for the diagnosis and remediation of reading difficulties.

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) state, "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) points out, reading instructional programs often are not linked to meet the needs of the characteristically different individual students. Identification of distinct developmental profiles suggests the potential for treatment to differentially impact performance in different groups. 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 suggests (Torgeson, Alexander, Wagner, Rashotte, Voeller, & Conway, 2001; Fletcher et al., 2005), early identification is key to successful rememdiation programs. Previous intervention research by Torgesen et al. (2001) reported that student gains in the accuracy of word-recognition skills with remediation did not correspond to gains in reading fluency. Especially with older children, fluency is a critical component for facilitating and developing reading comprehension strategies. As noted by Fletcher et al. (2005), students identified as having reading difficulties in third grade will need to read for eight hours a day for a year to overcome their deficits. Again, these studies underscore the need for early identification where prevention programs in early grades show parallel gains in both recognition and fluency (Torgesen, 2000). As pointed out by Lyon (1998), appropriate instruction carried out by expert teachers is not only the key to remediation, but also to prevention of reading difficulties. Previous intervention studies have shown that explicit and systematic instruction on the phonologic and orthographic connections in words, comprehension strategies, and increase opportunity to read for developing fluency have shown to be effective (Blachman, Schatschneider, Fletcher, Francis, Clonan, Shaywitz, & Shaywitz, 2004; Ehri, Nunes, Stahl, & Willows, 2001; Liberman & Liberman, 1990). 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. Given the increased emphasis on treatment, a systematic examination of differential treatment effect based on subtypes merits further investigation, although prior research on attribute-treatment intervention has met with limited success (Fletcher et al., 2005).

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 simply reflects that minorities are at increased risk for reading problems. A recent report by Snow et al. (1998) to the National Research Council suggests that, "children from poor families and children of African American and Hispanic descent are at much greater risk of poor reading outcomes." Teasing apart the relationship of external factors and reading achievement is complicated by inadequate indicators of SES. One of the limitations of this 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 is insufficient.

As pointed out by one of the reviewers, one of the limitations of our findings is the relatively restricted measure of word recognition and lack of reading comprehension measure. However, we found that the correlation is quite high between our word recognition measure and the word attack measure, as measured by the Woodcock Johnson Word Attack (Woodcock & Johnson, 1989). At the end of grade 1, the correlation is .78 between our word reading measure and Woodcock word attack. At the end of grade 2, the correlation is .71. Given the importance of reading comprehension skills in reading development, replication of this study with inclusion of a comprehensive word reading measure as well as reading comprehension is warranted.

Another important factor omitted in the analysis the consideration of school level variability. Given there were only three schools in the data, systematic exploration of school effect on student reading outcome 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. Also, given the recent findings in behavioral genetics studies (Byrne, Delaland, Fielding-Barnsley, Quain, Samuelsson, & Hoien, 2002; Harlaar, Spinath, Dale, & Plomin, 2005; Petrill, Deater-Deckard, Thompson, DeThorne, & Schatschneider, 2006b), it will be important to consider environmental factors as well as genetic differences in early readers when examining differential treatment effects. For successful remediation, it is important to consider which other factors help characterize the different reading development profiles in future studies.

Whether growth mixture modeling can be harnessed to improve identification of at-risk students and assist teachers in targeting these students with appropriate instruction awaits future research, which has only recently become possible due to the increased emphasis on assessment and data-driven instruction. The feasibility of implementing an ongoing early reading assessment program into practice remains a question for further research.

References

Bandeen-Roche, K., Miglioretti, D.L., Zeger, S.L., & Rathouz, P.J. (1997). Latent variable regression for multiple discrete outcomes. *Journal of the American Statistical Association, 92,* 1375-1386.

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, MD, USA: Paul H. Brookes Publishing Co.

Blachman, B., Schatschneider, C., Fletcher, J., Francis, D., Clonan, S., Shaywitz, B., & Shaywitz, S. (2004). Effects of Intensive Reading Remediation for Second and Third Graders and a 1-year Follow-up. *Journal of Educational Psychology, 96,* 444- 461.

Bryk, A.S., Raudenbush S.W. (1992). Hierarchical linear models. Newbury Park, CA: Sage Publications.

Byrne, B., Delaland, C., Fielding-Barnsley, R., Quain, P., Samuelsson, S., Hoien, T., et al. (2002). Longitudinal twin study of early reading development in three countries: Preliminary results. Annals of Dyslexia, 52, 49-74.

Carlson, C.D. (1998). Socioeconomic Status and Reading Achievement: The Mediating Role of Home Processes and Pre-Reading Skills. Unpublished Dissertation, University of Houston, Texas.

Carroll, J., Davies, P., & Richman, B. (1971). *The American heritage word frequency book.* Boston, MA: Houghton Mifflin.

Catts, H.W. (1991). Early Identification of reading disabilities. *Topics in Language Disorders, 12,* 1-16.

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.

Ehri, L.C., Nunes, S.R., Stahl, S.A., & Willows, D.M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Education Research, 71,* 393-447.

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,* 3, 506-522.

Foorman, B.R., Francis, D.J., Fletcher, J.M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to lead: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90,* 37-55.

Francis, D.J., Fletcher, J.M., Shaywitz, B.A., Shaywitz, S.E., & Rourke, B. (1996a) 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., Shaywitz, S.E., Stuebing, K.K., Shaywitz, B.A., & Fletcher, J.M. (1996b). Developmental lag versus deficit models of reading disability: a longitudinal, individual growth curves analysis. *Journal of Educational Psychology, 88,* 3-17.

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. 2958). Baltimore, MD, USA: Paul H. Brookes Publishing Co.

Gunning, T.G. (1998). Assessing and Correcting Reading and Writing Difficulties. Boston: Allyn and 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(4), 373-384.

Haynes, M., & Jenkins, J. (1986). Reading instruction in special education resource rooms. *American Educational Research Journal, 23*, 161-190.

Hertzog, C. & Shaie, K.W. (1988). Stability and change in adult intelligence: Simultaneous analysis of longitudinal means and covariance structures. *Psychology and Aging, 3,* 122-130.

Jennrich, R.I. & Schluchter, M.D. (1986). Unbalanced repeated measures models with structured covariance matrices. *Biometrics, 42,* 805-820.

Kass, R. & Raftery, A.E. (1995). Bayes factors. *Journal of the American Statistical Association, 90,* 773-795.

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

Liberman, I.Y. & Liberman, A.M. (1990). Whole language vs. code emphasis: Underlying assumptions and their implications for reading instruction. *Annals of Dyslexia, 40*, 51-76.

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, D.C. 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: A Wiley Interscience Publication.

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). Lawrence Erlbaum Associates.

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. Washington, D.C.: APA.

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. (Ed.), Multilevel Modeling: Methodological Advances, Issues, and Applications (in press). Mahaw, NJ.: Lawrence Erlbaum Associates.

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. & Muthén, L. (2001). Mplus User's Guide. Los Angeles, CA: Muthén & Muthén.

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 Research Council. (1998). Prevention of Reading Difficulties in Young Children. C.E. Snow, M.S. Burns, & P. Griffin, (Eds.). Washington, DC: National Academy Press.

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

Report on the 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. NIH Pub. No. 00-4754. Washington, DC:U.S. Department of Health and Human Services, Public Health Service, National Institute of Health, National Institute of Child Health and Human Development.

Satz, P. & Fletcher, J.M. (1988). Early identification of learning disabled children: An old problem revisited. *Journal of Consulting and Clinical Psychology, 56,* 824829.

Schatschneider, D., 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.

Schenck, B., Fitzimmons, 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, B.A. & Shaywitz, S.E. (1994). Measuring and analyzing change. In G.R. Lyon (Ed.), Frames of Reference for the Assessment of Learning Disabilities: New Views on Measurement Issues (pp. 59-68). Baltimore, MD, USA: Paul H. Brookes Publishing Co.

Shaywitz, S.E., Escobar, M.D., Shaywitz, B.A., Fletcher, J.M., & Makuch, R. (1992). Distribution and temporal stability of dyslexia in an epidemiolgical 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: Longwood Division, Allyn and Bacon.

Strag, G.A. (1972). Comparative Behavioral Ratings of Parents with Severe Mentally Retarded, Special Learning Disability, and Normal Children. Journal of Learning Disabilities, 5, 10.

Torgesen, J.K. & Wagner, R.K. (2002). Predicting reading ability. Journal of School Psychology, 40, 1-26.

Torgeson, 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. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research & Practice, 15*, 55-64.

Torgesen, J., Wagner, R., 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 Tests 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, 1, 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, 5, 431-442.

Woodcock, R.W., & Johnson, M. B. (1989). *Woodcock-Johnson Psychoeducational Battery-Revised.* Allen, TX: DLM Teaching Resources.

Footnote

[1]For the rapid naming variable, we combined the speed measure, RNL_S (# correct/# of seconds) with total correct (RNL_TR) to create TRNL (rapid naming) where TRNL =log_2(RNL__TR/RNL__S_ + .1).

Table 1

Cohort Structure for the Larger Study

Cohort	Year 1	Year 2	Year 3	Year 4
1	K (n = 182)			
2	1 (n = 182)	2 (n = 159)		
3	K (n = 183)	1 (n = 170)	2 (n = 133)	
4		K (n = 210)	1 (n = 158)	2 (n = 91)
5			K (n = 189)	1 (n = 109)

Table 2

Descriptive Statistics for the Measures Used in the Study

Variables	M	SD
Phonological Awareness at time 1	-1.22	.59
Phonological Awareness at time 2	-1.01	.65
Phonological Awareness at time 3	-.79	.74
Phonological Awareness at time 4	-.60	.82
Word Recognition at time 1	-.98	.84
Word Recognition at time 2	-.68	.85
Word Recognition at time 3	-.40	.90
Word Recognition at time 4	-.12	.90
Word Recognition at time 5	.26	.76
Word Recognition at time 6	.46	.77
Word Recognition at time 7	.60	.77
Word Recognition at time 8	.77	.77
Rapid Naming at time 4 (transformed)	-.26	.74

Table 3

BIC and Entropy Values for Kindergarten Models

Class	BIC	Entropy
2	1380.66	0.83
3	1369.44	0.75
4	1358.15	0.74
5	1370.02	0.7
6	1382.06	0.73

Table 4

Class counts and Proportion of Students in the Kindergarten Four-class Model

	Class 1	Class 2	Class 3	Class 4
Class Counts	229	20	131	31
Proportion of Total Sample	56	5	32	7

Table 5

Intercept and Slope for Five Phonological Awareness Development Profiles

	Intercept	Slope
PA1	-1.55 (0.26)	0.03 (0.04)
PA2	-1.07 (0.23)	0.14 (0.06)
PA3	-0.26 (0.19)	0.40 (0.16)
PA4	0.10 (0.15)	0.27 (0.04)
PA5	1.14 (0.31)	0.34 (0.09)

Table 6

BIC and Entropy Values for Grade 1 and 2 Models

	BIC	Entropy
Two-class	2260.08	0.76
Three-class	2297.29	0.74
Four-class	2273.16	0.73
Five-class	2245.8	0.78
Six-class	2236.49	0.79
Seven-class	2233.74	0.82

Table 7

Intercept and Slope for Five Word Recognition Development Profiles

	Intercept	Slope
WR1	-0.95 (0.10)	0.12 (0.02)
WR2	0.46 (0.08)	0.27 (0.01)
WR3	1.27 (0.08)	0.26 (0.01)
WR4	2.03 (0.14)	0.51 (0.02)
WR5	1.73 (0.07)	0.15 (0.01)

Table 8

Five-class Model Specification

	Grade 1 and 2				
Kindergarten	WR1	WR2	WR3	WR4	WR5
PA1	Class 1				
PA2		Class 2			
PA3			Class 3		
PA4				Class 4	
PA5					Class 5

Table 9

Class Specifications and Class Counts for the Ten-class Model

	First and Second Grade				
Kindergarten	WR1	WR2	WR3	WR4	WR5
PA1	Class 1				
	45 (11%)				
PA2		Class 2	*Class 3*		
		63 (15%)	*77 (19%)*		
PA3		*Class 4*	Class 5	*Class 6*	
		8 (2%)	56 (14%)	*36 (9%)*	
PA4			*Class 7*	Class 8	*Class 9*
			20 (5%)	8 (14%)	*30 (7%)*
PA5					Class 10
					18 (4%)

Note. The italicized entries represent the five classes that change class membership between grades.

Table 10

Estimated Means and Standard Errors of Intercepts and Growth Parameters for the Ten-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 Captions

Figure 1. Individual growth trajectories for word recognition development.

Figure 2. Estimated mean growth curves for the five-class model representing the phonemic awareness in kindergarten.

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

Figure 4. Class 1 and Class 2: word recognition development based on kindergarten classification.

Figure 5. Class 3 and Class 4: word recognition development based on kindergarten classification.

Figure 6. Class 5: word recognition development based on kindergarten classification.

Figure 7. Path diagram for phonemic awareness and word recognition combined.

Figure 8. 10-class model: estimated mean growth curves for phonemic awareness and word recognition.

Figure 9. Class 1 and Class 2: observed individual growth trajectories with estimated mean growth trajectories for the 10-class model.

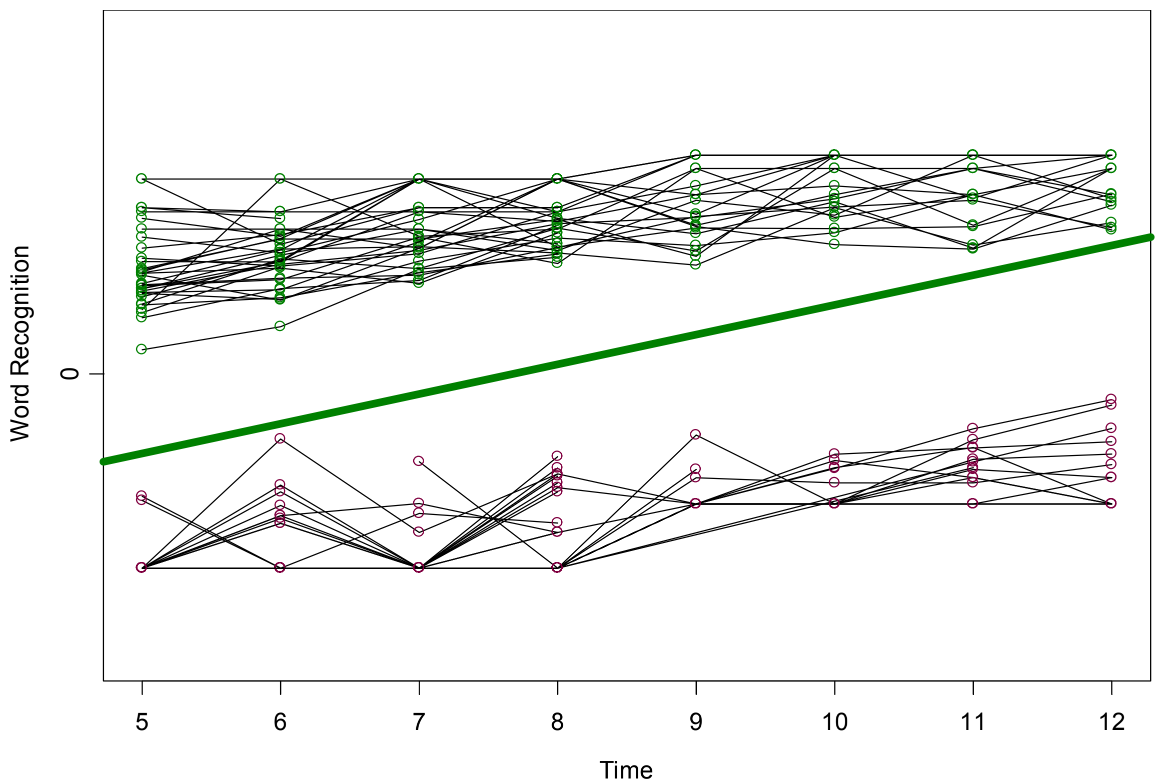
Figure 10. Class 3 and Class 4: observed individual growth trajectories with estimated mean growth trajectories for the 10-class model.

Figure 11. Class 5 and Class 6: observed individual growth trajectories with estimated mean growth trajectories for the 10-class model.

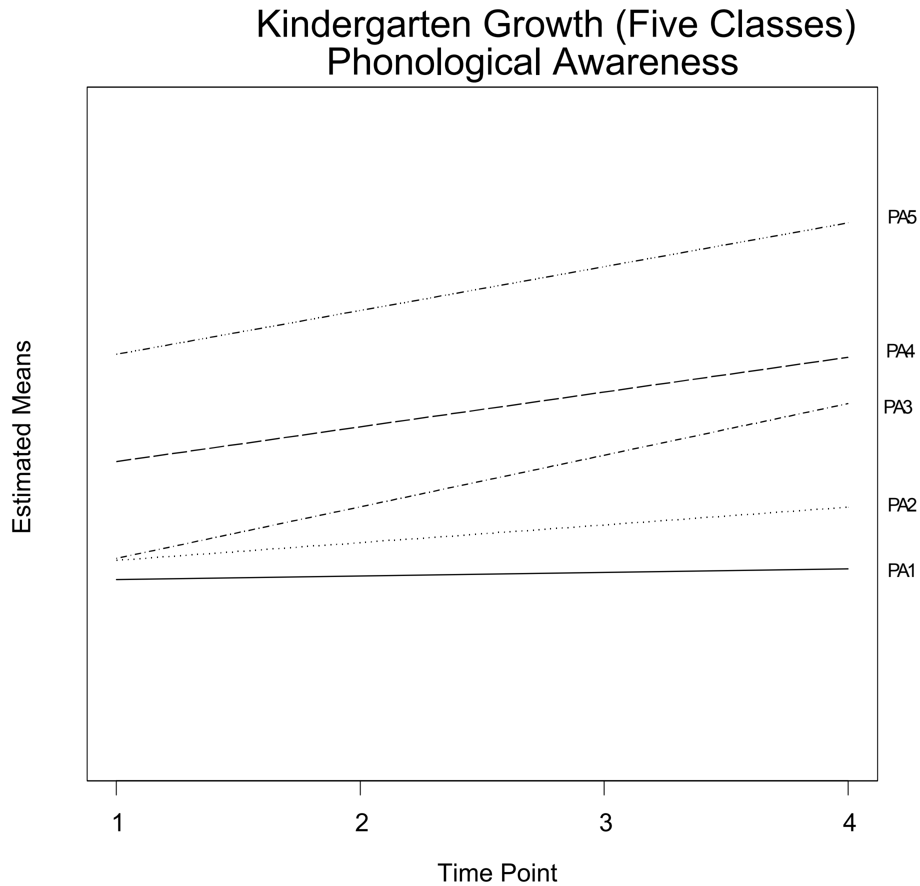
Figure 12. Class 7 and Class 8: observed individual growth trajectories with estimated mean growth trajectories for the 10-class model.

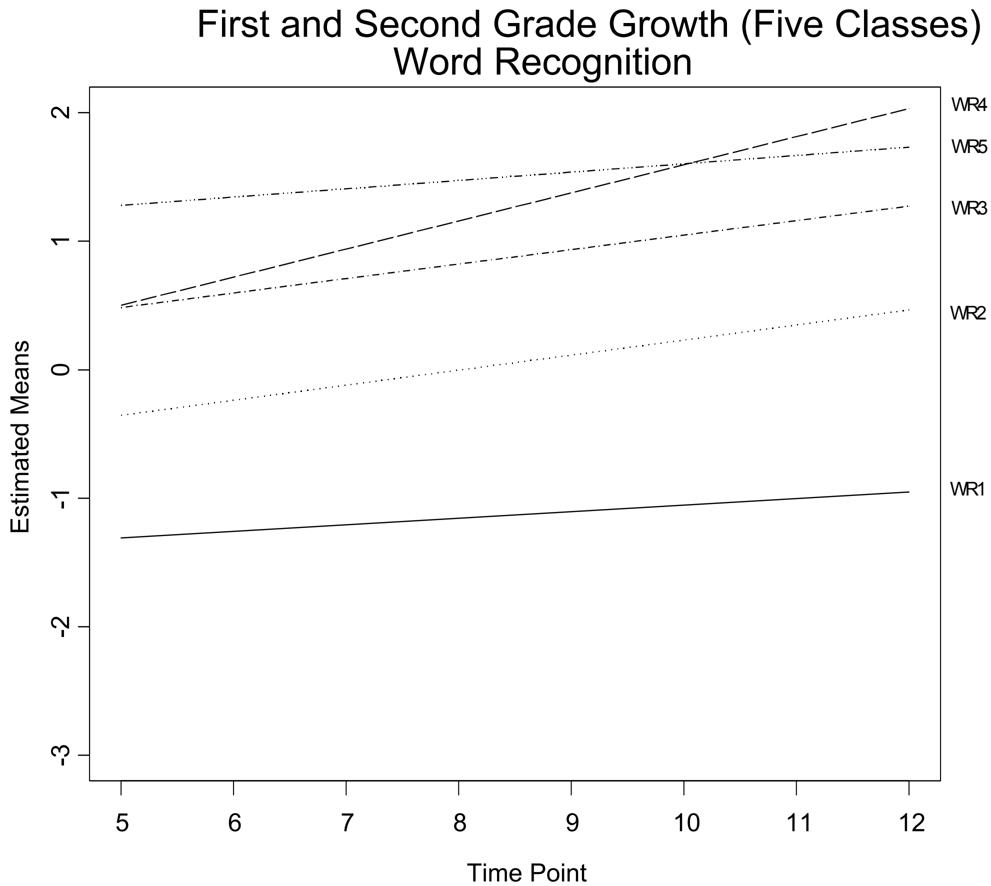
Figure 13. Class 9 and Class 10: observed individual growth trajectories with estimated mean growth trajectories for the 10-class model.

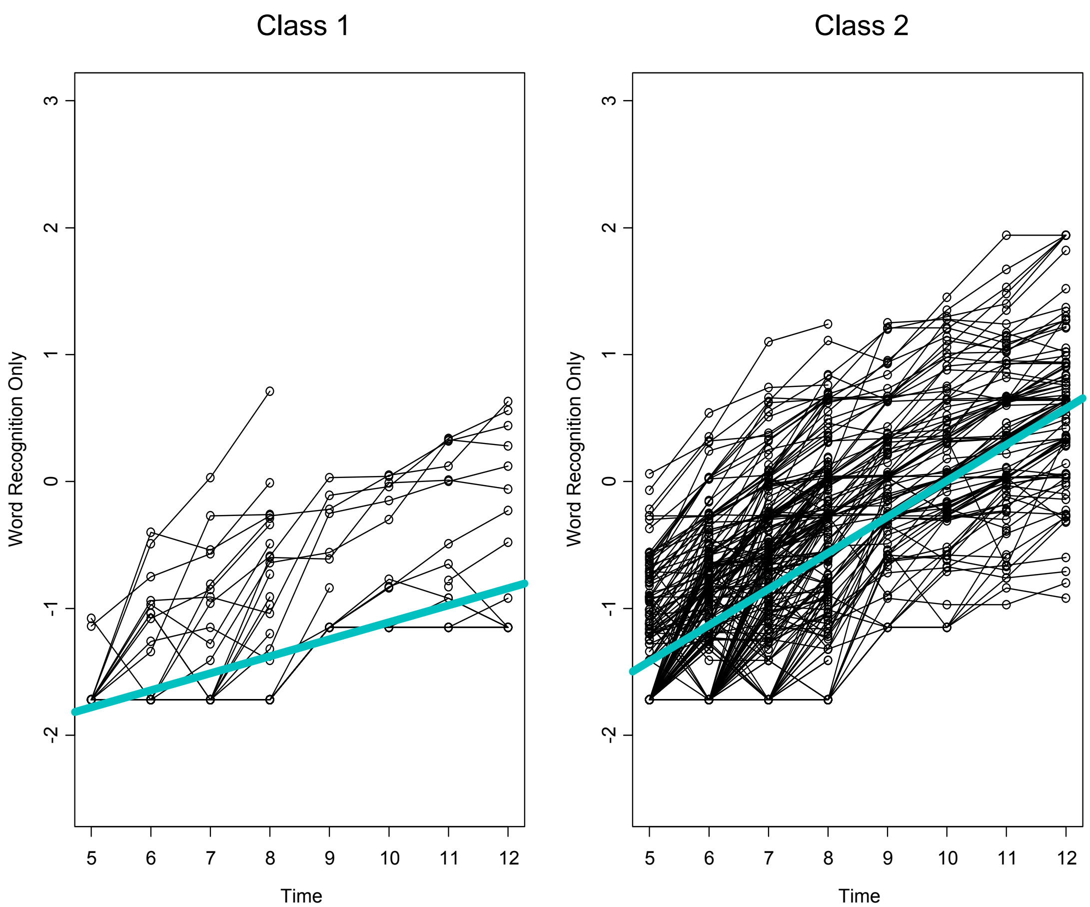
Figure 14. Multinomial logistic regression of class regressed on rapid naming scores.

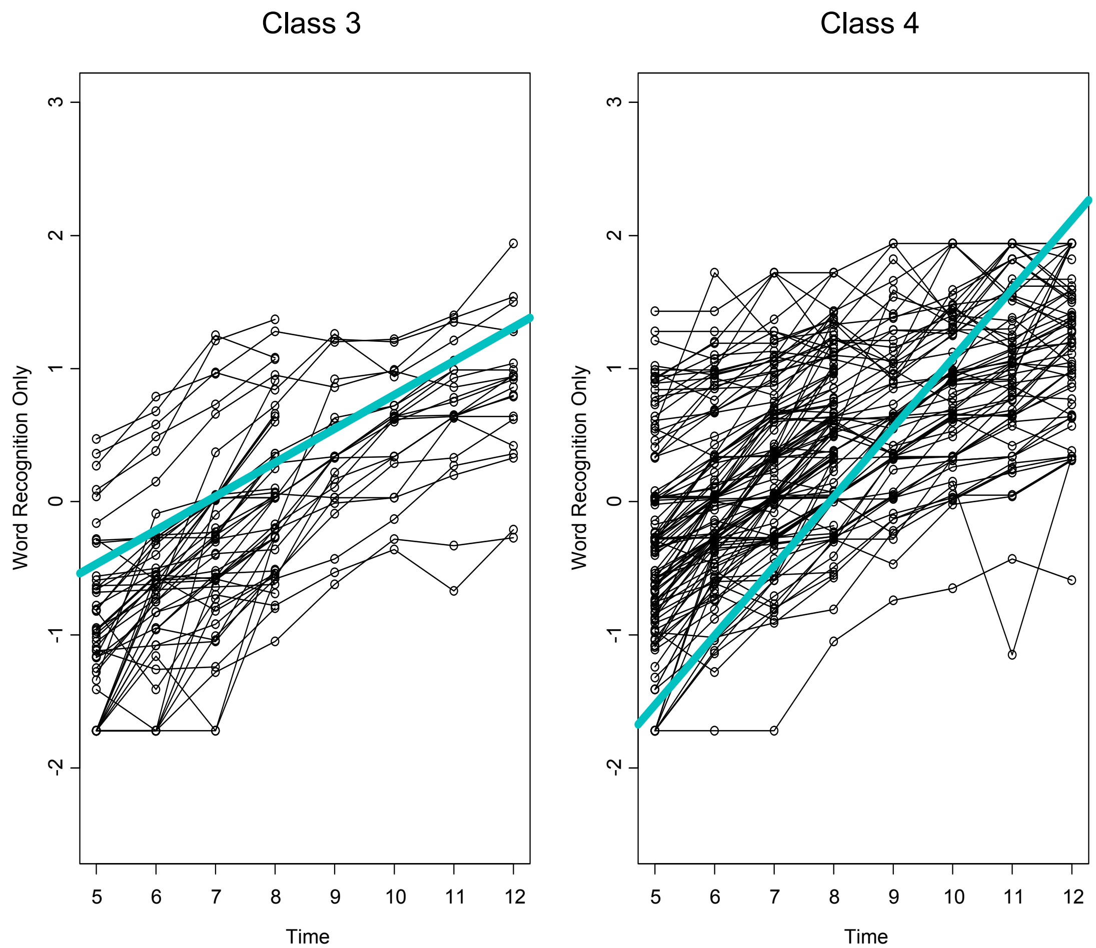


Word Recognition
0
5
6
7
8
9
10
11
12
Time

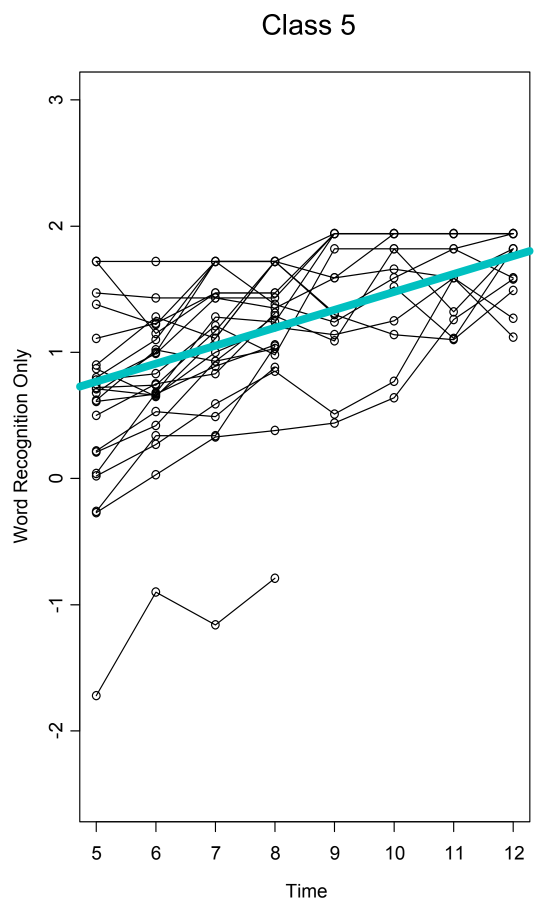

Kindergarten Growth (Five Classes)
Phonological Awareness
Estimated Means
Time Point
1
2
3
4
PA5
PA4
PA3
PA2
PA1

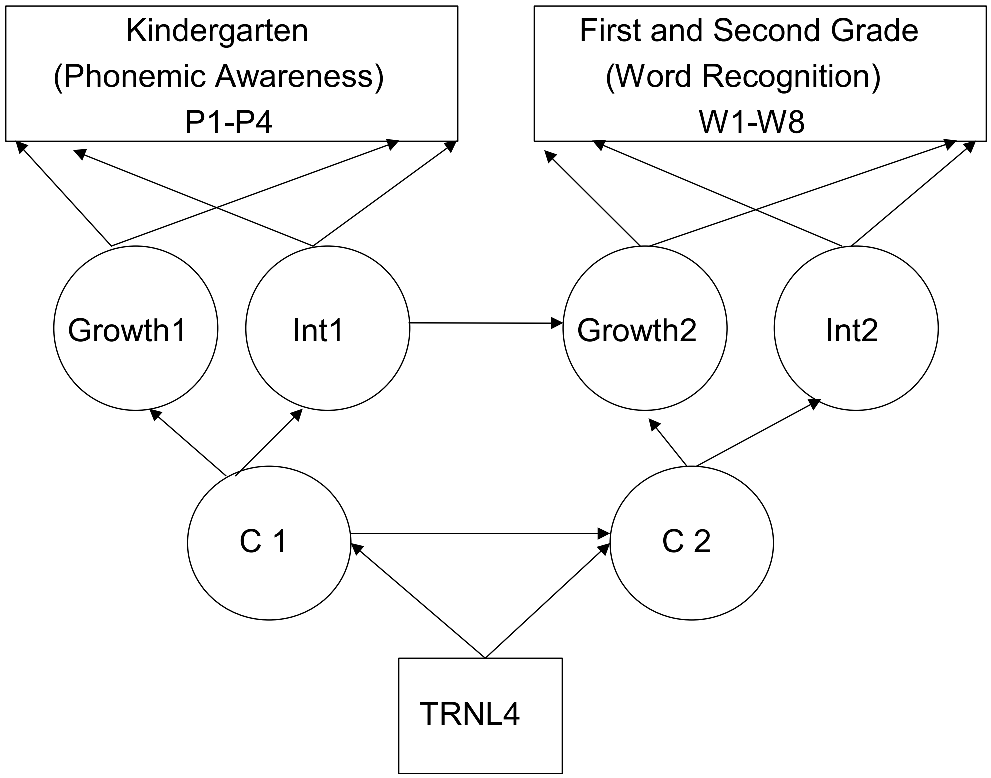

First and Second Grade Growth (Five Classes)
Word Recognition
Estimated Means
Time Point
WR4
WR5
WR3
WR2
WR1

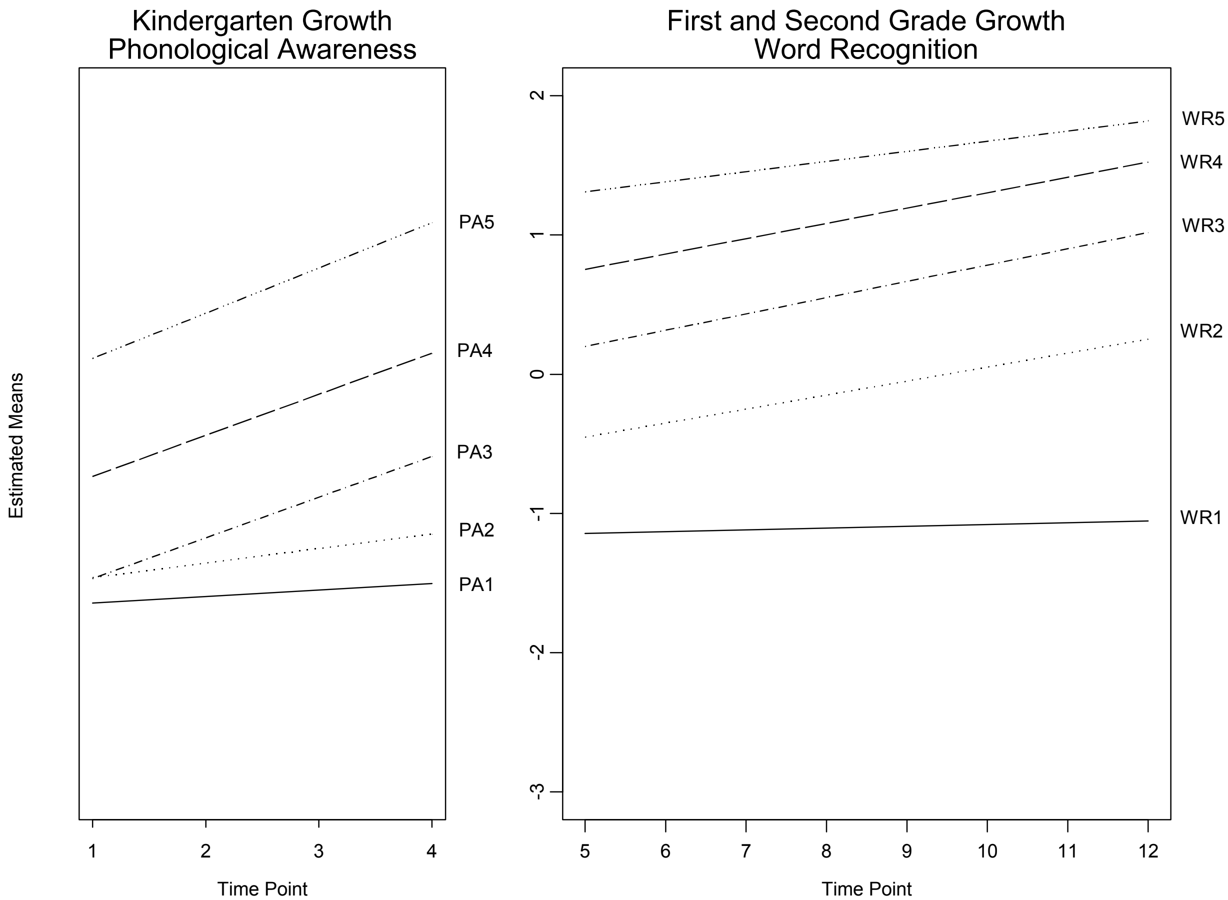

Class 1
Class 2
Word Recognition Only
Time
3
2
1
0
-1
-2
5
6
7
8
9
10
11
12

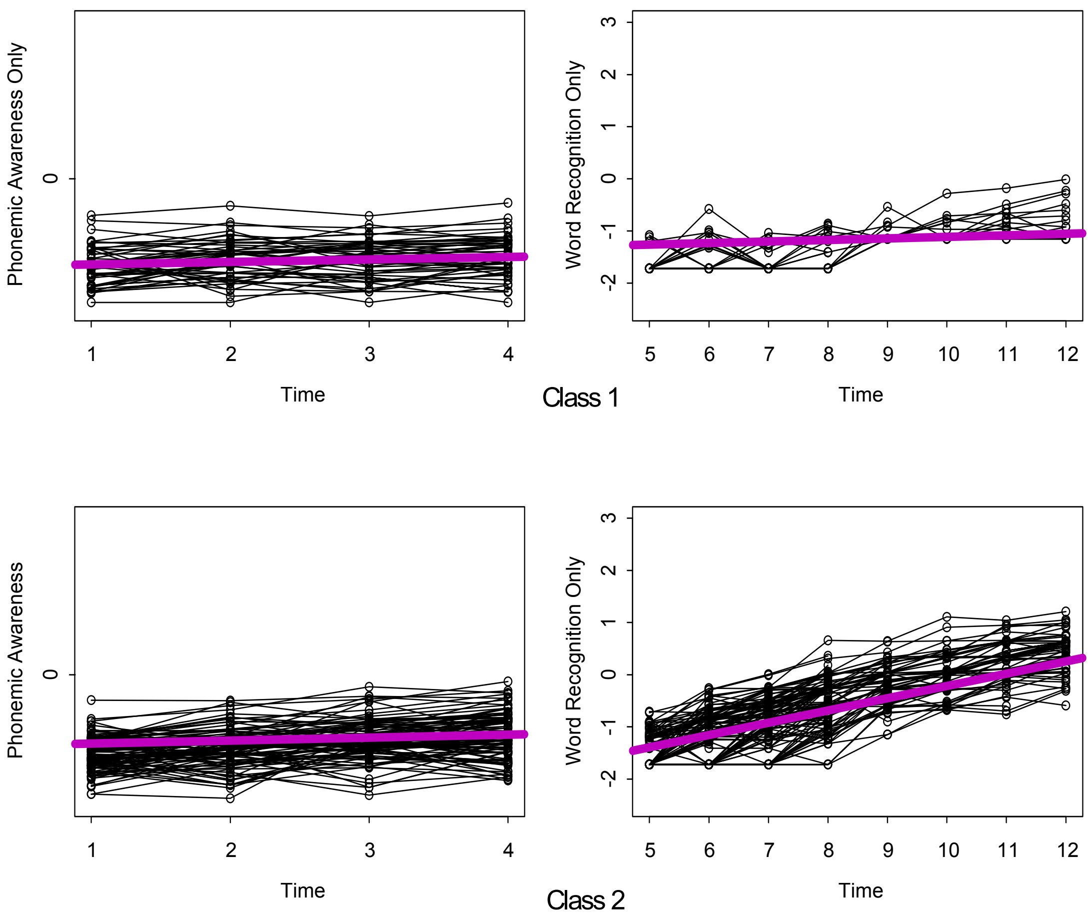


Class 3
Class 4
Word Recognition Only
Time
3
2
1
0
-1
-2
5
6
7
8
9
10
11
12

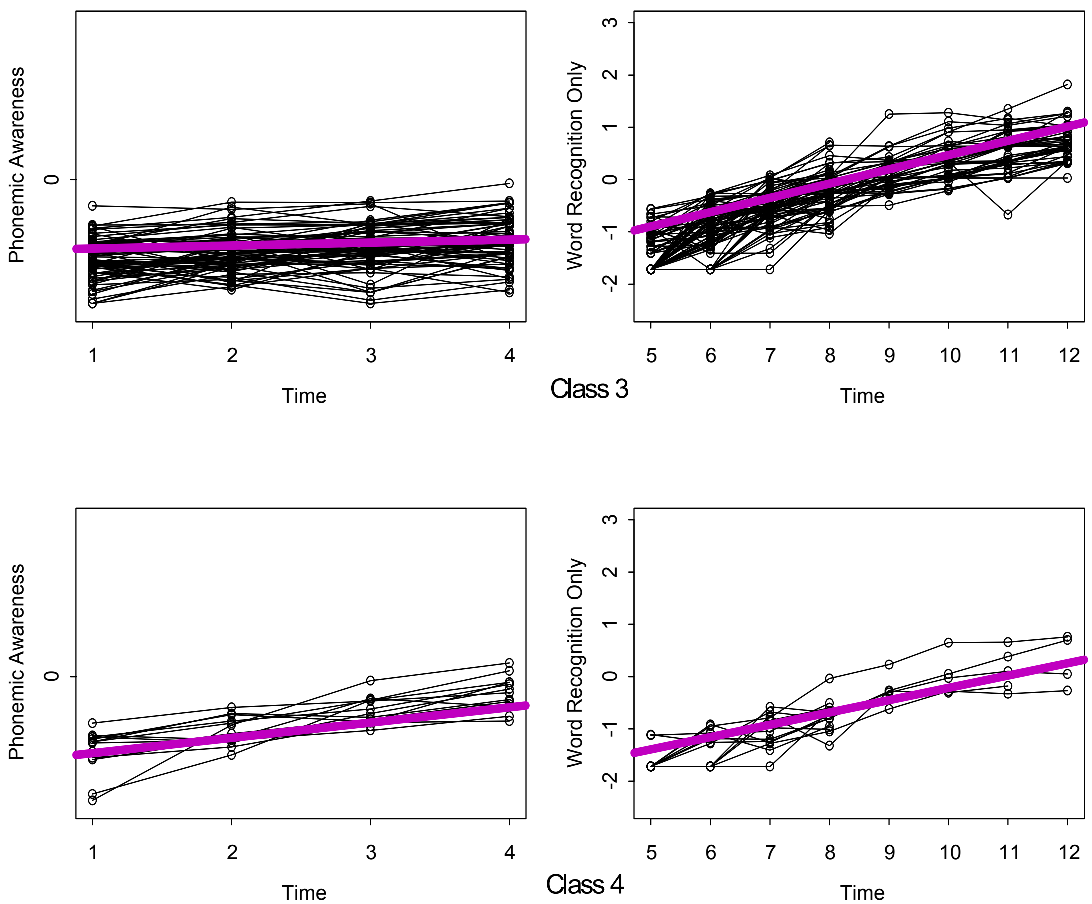

Class 5
Word Recognition Only
Time
3
2
1
0
-1
-2
5
6
7
8
9
10
11
12

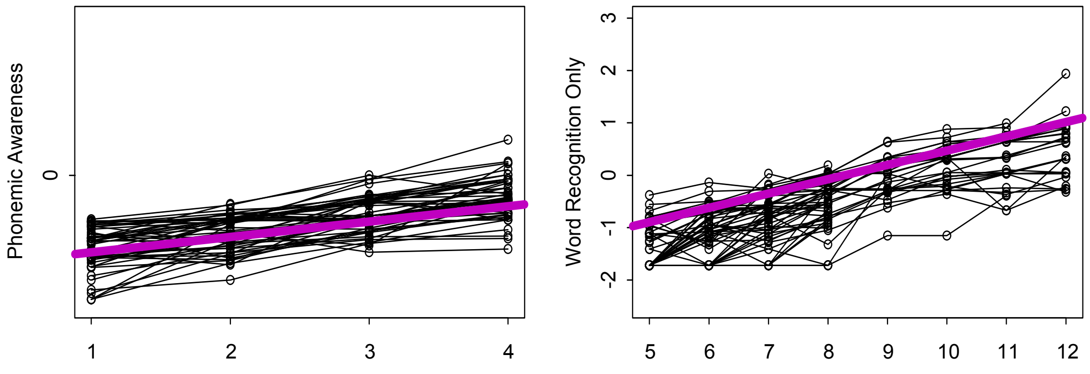

Kindergarten
(Phonemic Awareness)
P1-P4
First and Second Grade
(Word Recognition)
W1-W8
Growth1
Int1
Growth2
Int2
C 1
C 2
TRNL4


Kindergarten Growth
Phonological Awareness
Estimated Means
PA5
PA4
PA3
PA2
PA1
1
2
3
4
Time Point
First and Second Grade Growth
Word Recognition
2
1
0
-1
-2
-3
WR5
WR4
WR3
WR2
WR1
5
6
7
8
9
10
11
12
Time Point

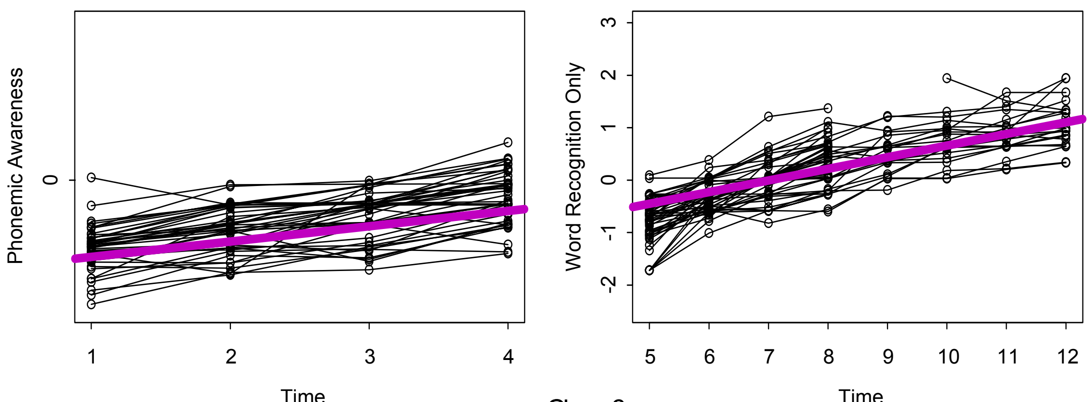


Phonemic Awareness Only
0
1
2
3
4
Time
Word Recognition Only
3
2
1
0
-1
-2
5
6
7
8
9
10
11
12
Time
Class 1
Phonemic Awareness
0
1
2
3
4
Time
Word Recognition Only
3
2
1
0
-1
-2
5
6
7
8
9
10
11
12
Time
Class 2

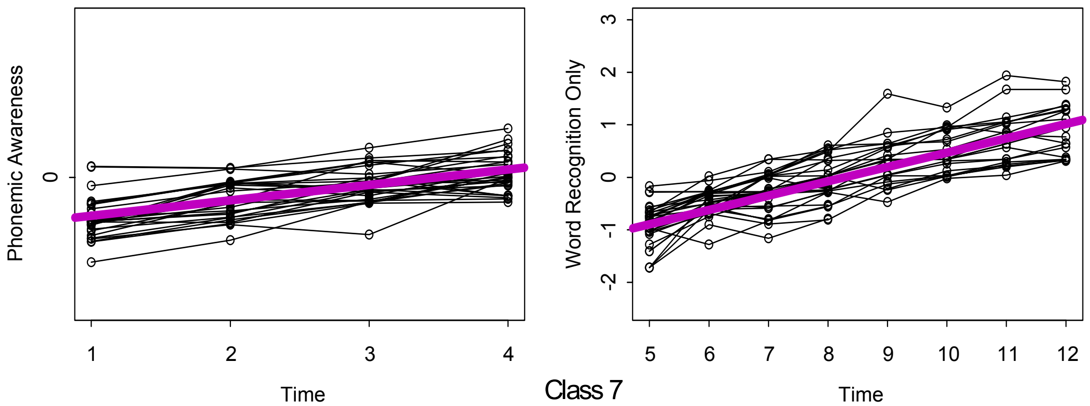

Phonemic Awareness
0
1
2
3
4
Time
Word Recognition Only
3
2
1
0
-1
-2
5
6
7
8
9
10
11
12
Time
Class 3
Phonemic Awareness
0
1
2
3
4
Time
Word Recognition Only
3
2
1
0
-1
-2
5
6
7
8
9
10
11
12
Time
Class 4



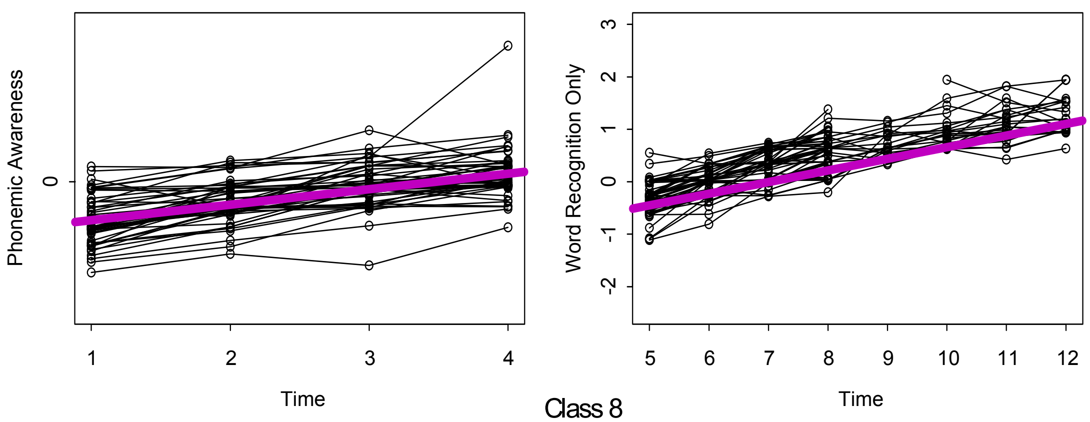
Class 5



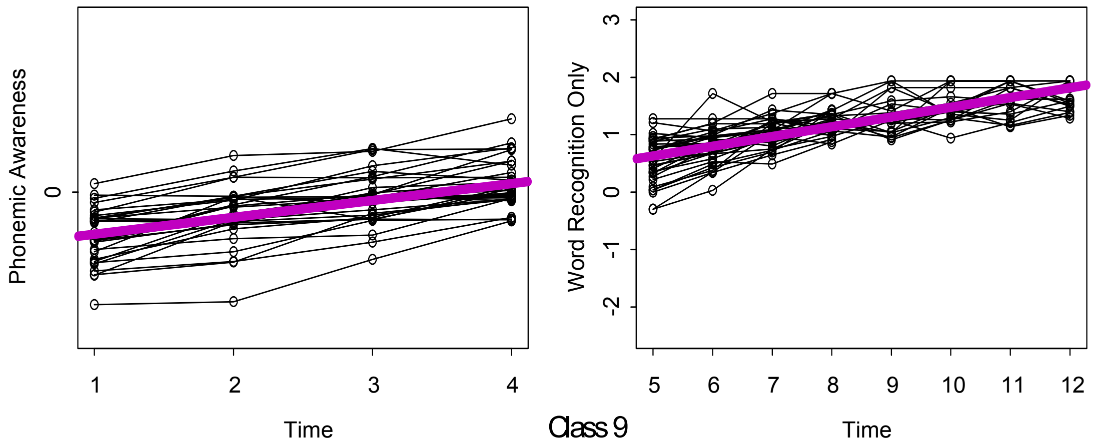
Class 6



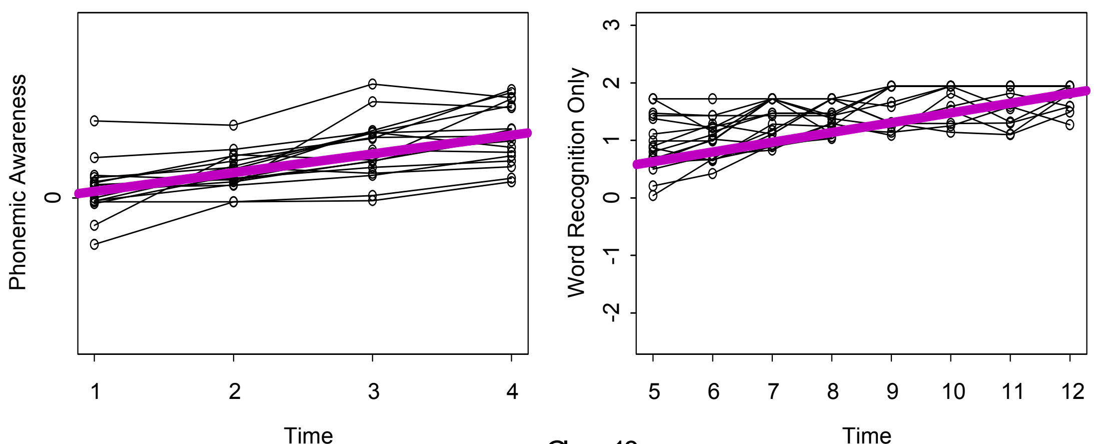
Class 7



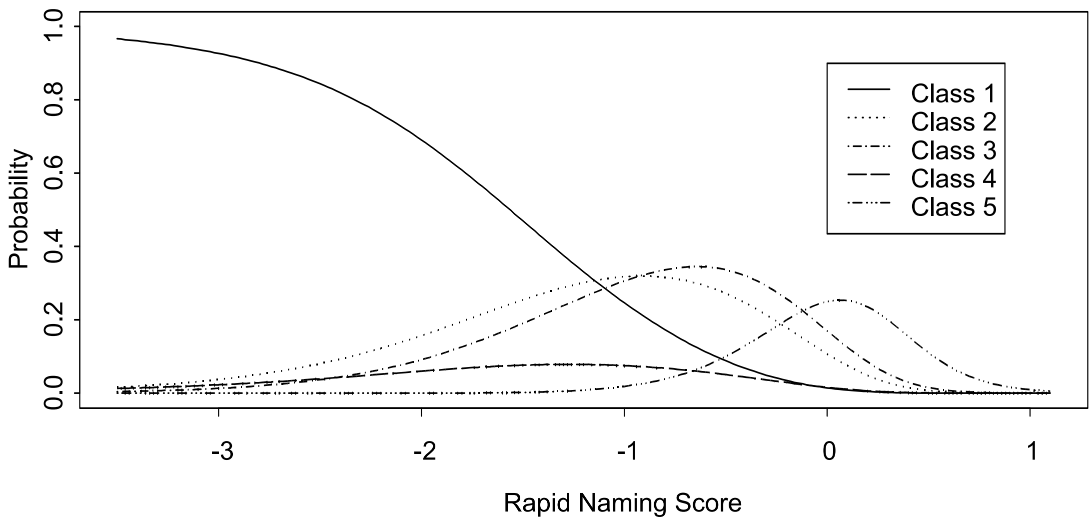
Class 8



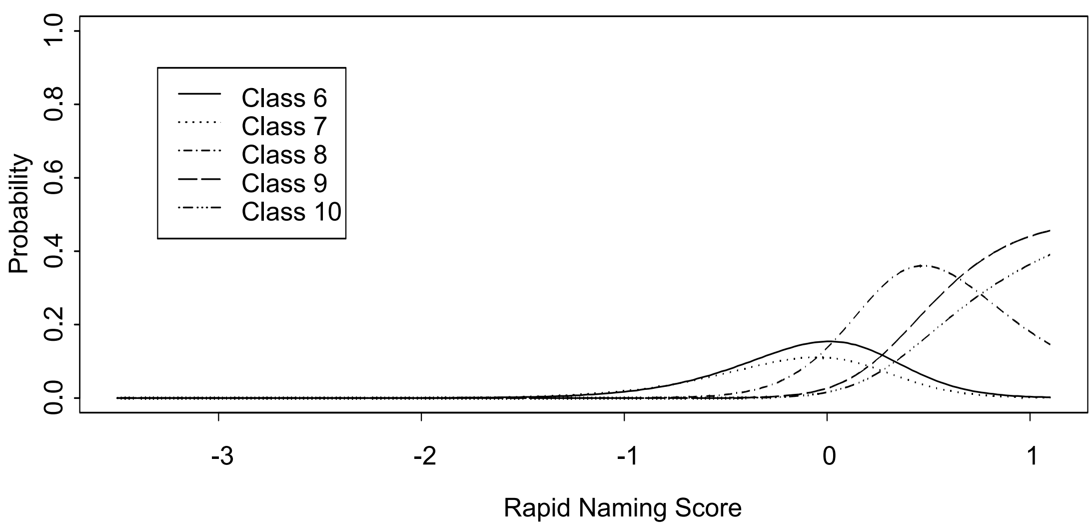
Class 9



Class 10



Probability
1.0
0.8
0.6
0.4
0.2
0.0
-3
-2
-1
0
1
Rapid Naming Score
Class 1
Class 2
Class 3
Class 4
Class 5



Probability
1.0
0.8
0.6
0.4
0.2
0.0
-3
-2
-1
0
1
Rapid Naming Score
Class 6
Class 7
Class 8
Class 9
Class 10